BLOOD BIOMARKERS IN

CHRONIC OBSTRUCTIVE PULMONARY DISEASE (COPD)

by

Minhee Jin

B.H.Sc., The University of Western Ontario, 2009
B.Tech F.Sc., The British Columbia Institute of Technology, 2013
M.Sc., The University of Strathclyde, 2014

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Experimental Medicine)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

July 2019

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

Blood Biomarkers in Chronic Obstructive Pulmonary Disease (COPD)

submitted by Minhee Jin in partial fulfillment of the requirements for

the degree of Master of Science

in Experimental Medicine

Examining Committee:

Dr. Don D. Sin
Supervisor

Dr. Janice Leung
Supervisory Committee Member

Dr. Tillie-Louise Hackett
Supervisory Committee Member

Dr. Bradley Quon
External Examiner

Abstract

Chronic obstructive pulmonary disease(COPD) is an age-related disease and demonstrates many biological hallmarks of aging including telomere shortening and epigenetic alterations. This study examined the relationship between telomere length, reflecting replicative senescence, and Alu and LINE-1(L1) methylation, reflecting one of the major epigenetic changes of peripheral blood leukocytes, and clinical outcomes, including lung function, health status, rate of exacerbations, and risk of mortality with different COPD stages.

Using quantitative polymerase chain reaction, we measured the absolute telomere length(aTL) of DNA extracted from blood samples of 576 participants with moderate-to-severe COPD treated with either azithromycin or placebo in the Macrolide Azithromycin for Prevention of Exacerbations of COPD(MACRO) study. Using bisulfite pyrosequencing, we measured the percentage(%) Alu and L1 methylation of DNA extracted from blood samples of 495 participants with acute-exacerbating-COPD(AECOPD) in the COPD Rapid Transition Program(RTP), and 373 participants with non-COPD in the Canadian Cohort of Obstructive Lung Disease(CanCOLD).

Participants with shorter aTL (defined as below the median cutoff value for aTL) had worse health status as defined by higher St. George's Respiratory Questionnaire(SGRQ) scores (P=0.034). In the placebo arm, the rate of exacerbations (P=0.002) and the risk of mortality (P=0.015) were significantly higher in the shorter telomere group than in the longer telomere group. % Alu methylation was significantly related to FEV_1% predicted, FEV_1 in liters(L) and FEV_1/FVC ratio (P<0.0001), and when the cohort was divided into quartiles, the higher the quartile the better the FEV_1 % predicted and FEV_1(L) (P<0.0001). % Alu methylation was also significantly related to the decline of FEV_1% predicted (%/year), (P=0.004) and FEV_1/FVC ratio (%/year)(P=0.007) as

well as to total and sub-domain SGRQ scores (P<0.0001 except the Impact score P=0.041). % L1 methylation was significantly different between different states of COPD: stable-, non- & AE-COPD, where the non-COPD participants had the highest percentage of methylation and AECOPD participants demonstrated the lowest (P<0.0001).

Together these data suggest that senescence blood biomarkers in COPD patients are associated with poor health outcomes. These relatively simple blood measurements may represent a potential source of clinically translatable biomarkers for identifying patients at high risk of morbidity and mortality in COPD.

Lay Summary

Chronic Obstructive Pulmonary Disease (COPD) is a progressive lung disease which is caused by airway inflammation and damage to the air sacs in the lungs. COPD is characterized by shortness of breath, wheezing and coughing which often lower the quality of life by interfering with physical and social activities. The risk of COPD increases rapidly with aging, especially after 40 years of age. This study investigated measurable indicators in the blood to observe the relationships not only with the standardized lung function measures, which are used to diagnose COPD but also with the patient-related outcomes such as health-status and mortality in patients with COPD. Findings from this study could help physicians to forecast disease progression and assist with tailoring treatment for patients with COPD in the future.

Preface

Chapter 1 has been published. Minhee Jin, Lee EC, Ra SW, Fishbane N, Tam S, Criner GJ, Woodruff PG, Lazarus SC, Albert R, Connett JE, Han MK, Martinez FJ, Aaron SD, Reed RM, Man SFP, Leung JM, Sin DD. (2018) Relationship of Absolute Telomere Length with Quality of Life, Exacerbations, and Mortality in COPD. Chest. 154(2):266-273. Work presented in this chapter was conducted at the Centre for Heart Lung Innovation UBC and St. Paul's Hospital (HLI) located in St. Paul's hospital. EC. Lee and I conducted all the experiments and wrote most of the manuscript. SW. Ra and N. Fishbane helped performing statistical analysis and interpretations. GJ. Criner, PG. Woodruff, SC. Lazarus, R. Albert, JE. Connett, MK. Han, FJ. Martinez, SD. Aaron, RM. Reed, SFP. Man, JM. Leung and DD. Sin contributed to critical revision of the manuscript for important statistical, intellectual, clinical content and approved the final version. Data and leukocyte DNA samples used in the study presented in Chapter 1 were collected from Macrolide Azithromycin for Prevention of Exacerbation of COPD (MACRO) study. The biomarker component of the MACRO study was approved by the UBC Providence Health Care Ethics Committee (approval No. H11000786). The study "Telomere measurement" under Method was originally outlined by O'Callaghan NJ and Fenech M (2011) A quantitative PCR method for measuring absolute telomere length. Biol Proced Online.13:3.

Data and leukocyte DNA samples used in the study presented in Chapter 2 were collected from the Canadian Cohort of Obstructive Lung Disease (CanCOLD) and COPD Rapid Transition Program (RTP). The biomarker component of the CanCOLD and RTP were approved by the UBC Providence Health Care Ethics Committee (approval No.H11-00786 and H16-02270, respectively). Part of the experiment in Chapter 2, "pyrosequencing" was done at the Centre for

Molecular Medicine and Therapeutics by Dr. J. McIsaac. All other work presented in this chapter was conducted at the HLI. I was responsible for the design, sample preparation and data analysis. Y. Yang, A. Li and Dr. F. Stuart helped performing statistical analysis. Dr. D. Sin provided intellectual content and supervised the project.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AECOPD	Acute exacerbation of chronic obstructive pulmonary disease
aTL	absolute telomere length
BMI	Body mass index
C	Cytosine
CanCOLD	Canadian chronic obstructive lung disease
CD4T	CD4+T cells
CD8T	CD8+T cells
CI	Confidence interval
COLD	Canadian COPD prevalence
COPD	Chronic obstructive pulmonary disease
CpG	Cytosine-phosphate-guanine
CRP	C-reactive protein
CV	Cardio vascular
DM	Diabetes mellitus
DNA	Deoxyribonucleic acid
EWAS	Epigenome-wide association studies
FEV1	Forced expiratory volume in 1 second
FVC	Forced vital capacity
GERD	Gastroesophageal reflux disease
Gran	Granulocytes
HR	Hazard ratio
ICS	Inhaled corticosteroids
L	Liter
L1	LINE-1
LABA	Long-acting beta2 agonists
LAMA	Long-acting muscarinic antagonists
LINEs	Long interspersed nuclear elements
MACRO	Macrolide Azithromycin for Prevention of Exacerbations of COPD
Mono	Monocytes
NK	Natural killer cells
Nrf2	Nuclear factor erythroid 2-related factor
PCR	Polymerase chain reaction
RR	Risk ratio
RTP	Rapid transition program
SAM	*S*-adenosylmethionine
SD	Standard deviation
SGRQ	St. George's respiratory questionnaire

SINEs	Small interspersed nuclear elements
SNP	Single nucleotide polymorphism
Sp1	Transcription factor specificity protein 1
SPH	St. Paul's hospital
VIF	Variance inflation factor

Acknowledgements

I express my sincere gratitude and appreciation to all who have been there for me throughout the little details and processes of my project. First of all, I would like to thank my supervisor, Dr. Don Sin, for your provision of the valuable opportunity and guidance. I would also like to thank supervisory committee members, Dr. Janice Leung and Dr. Tillie-Louise Hackett for invaluable and insightful feedback which have been essential for my project progressing forward.

Thanks to my dearest colleague/ co-author/ friend/ advisor, Grace Eun-Chong Lee for your sincere support and tremendous help to execute and publish my Chapter 1. Thanks to Dr. Seungwon Ra for the help on statistical analyses for my Chapter 1 as well. Thanks to Dr. Julia MacIsaac for your technical support on experiments and guidance for Chapter 2. I would also like to thank Yolanda Yang, Annie Li and Maen Obeidat for your immense knowledge and guidance on statistical analyses for Chapter 2.

I give my thanks to all members of Sin Laboratory for being amazing individuals as well as being tremendous help academically and morally. Thank you, Dr. Michaela Nam, Dr. Fernando Studart, Dr. David Jaw, Dr. Anthony Tam, Dr. Chung Cheung and Dr. Tony Yang for your advice and guidance. I would also like to thank my fellow lab mates Kimia Shahangian for being such a cheerful deskmate, and Julia Yang, Nancy Yang and Dr. Yuji Cho for making my lab-life joyful. All the members of the Centre for Heart Lung Innovation, I give my deep appreciation to you as well.

Last, but not least, thank you to my family and friends for their unconditional support and love. Nothing would have been possible without you all.

Dedication

To my beloved family: Duk-yeon, Anna, Angela, Joseph, Sean, Joshua, Choco, Sandy, Rosa and Christian. And also…to whom actually try to read this through till the end.

Introduction

Previous studies have suggested that COPD may be a disease of accelerated aging for a variety of reasons including a close relationship with senescence-related disorders such as osteoporosis and dementia, as well as its exponential increase in prevalence beyond 50 years of age [1-4]. One important biomarker of replicative senescence is telomere length. Telomeres are specific repetitive DNA sequences that buttress terminal ends of chromosomes to maintain their integrity by limiting DNA attrition during replication and thus preventing harmful DNA events such non-homologous recombination. [5, 6] With each replication cycle, there is progressive shortening of telomeres, [7] which over time results in cell senescence and/or apoptosis when replicative (Hayflick) limits are reached [8, 9]. COPD is associated with oxidative stress, which becomes amplified during exacerbations [10, 11] and oxidative stress enhances telomere shortening [7].

Another important biomarker of epigenetic alteration is DNA methylation [1, 12-14]. DNA methylation is a major epigenetic mechanism that is involved in gene regulation and cell differentiation by adding methyl (CH_3) group to 5-carbon of the cytosine typically at cytosine-phosphate-guanine (CpG) sites [15-17]. Previously, specific CpG islands were recognized as predictor methylation sites of both biological and chronological age [18] and it is now well established that aging and other stressors such as microbial infections can lead to general hypomethylation at these sites [13, 18-20]. Thus, DNA hypomethylation has been used as a biomarker of biological aging in many age-related disease conditions such as cancer [20-22], cardiovascular disease [23, 24], neurological diseases [25-27] and total mortality [28, 29].

Chapter 1: The Relationship of Absolute Telomere Length with Quality of Life, Exacerbations, and Mortality in Chronic Obstructive Pulmonary Disease (COPD)

1.1 Telomere Shortening in COPD

We have shown previously that telomere lengths of peripheral leukocytes correlate to those in lung tissue, suggesting that telomeres in blood may be considered a biomarker of replicative senescence in the lungs [30]. Importantly, we and other groups have shown that telomeres are shorter in peripheral leukocytes of COPD patients (compared to healthy participants) [1, 31-33] and that short telomeres are associated with common comorbidities of COPD such as cardiovascular disease [34-36] and cancer [37, 38]. However, there is little information on whether telomere lengths from peripheral blood are associated with important patient-related outcomes in COPD such as health status, exacerbations or mortality. Using data from the Macrolide Azithromycin for Prevention of Exacerbations of COPD (MACRO) study, [39] we determined the relationship between blood telomeres and these patient-related outcomes in COPD patients.

1.2 Methods

1.2.1 Samples and Patient Related Outcomes

We used data and leukocyte DNA samples collected from participants with COPD in the MACRO study; the details of which have been previously published [39]. Briefly, MACRO recruited participants from 17 sites across 12 academic health centres in the United States. A total of 1,142 participants with COPD were randomized to azithromycin or placebo in addition to standard care

for 12 months. The inclusion criteria were: a clinical diagnosis of COPD, age of 40 years or older, a smoking history of at least 10 pack-years, a ratio of post-bronchodilator FEV_1/FVC less than 70% with a post-bronchodilator FEV_1 of less than 80%, and either hospital admission or emergency room presentation for acute exacerbation of COPD, as well as continuous usage of supplemental oxygen or treatment with systemic glucocorticoids within the preceding year. Exclusion criteria included asthma, hearing impairment, and an acute exacerbation of COPD within the 4 weeks prior to enrollment. Clinical outcomes including COPD exacerbations, health status, and all-cause mortality were assessed at recruitment and through follow-up.

We utilized the St. George's Respiratory Questionnaire (SGRQ)[40] to evaluate health status of the participants. SGRQ has three sub-domains: Symptoms, Activity and Impact. The *Symptom* score measures the impact of respiratory symptoms, their frequency and severity on health status; The *Activity* score measures the physical impairment of daily activities and the *Impact* score measures psychosocial disturbances due to the disease. SGRQ scores were collected at baseline and scores range from 0 to 100, with higher scores indicating worse status. An exacerbation was defined as "a complex of respiratory symptoms (increased or new onset) of more than one of the following: cough, sputum, wheezing, dyspnea, or chest tightness with a duration of at least 3 days requiring treatment with antibiotics or systemic steroids [41]". The participants reported whether a COPD exacerbation had occurred at each clinic visit and during phone follow-ups. The rates of exacerbation were determined by dividing the number of acute exacerbations by the person-years of follow-up. Owing to differential enrollment dates, participants were followed for up to 3.5 years.

1.2.2 Telomere Measurement and Statistical Analysis

We collected blood from 597 participants who agreed to participate in DNA biobanking. Each DNA sample was extracted from peripheral blood buffy coat, which was quantified using NanoDrop 8000 spectrophotometer (Thermo Scientific, Wilmington, USA) and diluted down to the concentration of 5ng/µl in EDTA buffer solution; aTL in each DNA sample was measured using a quantitative PCR (qPCR) method as outlined by O'Callaghan and Fenech [42]. In this method, a standard curve is generated from the fluorescent signals given by a series of known concentrations of telomere oligomer (TEL) DNA [(TTAGGG)x14]. The concentration of each testing sample was predicted by plotting the sample's fluorescence signal on a standard curve. To serve as a reference gene, a concentration of a single copy gene (36B4) DNA in each sample was measured using the same method. Telomere length was calculated as the ratio of the telomere DNA length from the standard curve to the 36B4 DNA length. Each sample was run in triplicate on Hard-Shell PCR plate 384-well, using CFX 384 Touch Real-Time PCR Detection System (Bio-rad). DNA from a short telomere cell line (HEK293) and a long telomere cell line (K562) were plated in each cycle to control inter-experimental variations [43].

The telomere lengths of each sample were normalized to a reference cell line control sample, which was evaluated on each plate [42]. Telomere length measurements (kbp/genome) were natural log-transformed to achieve normality. For analytic purposes, we divided the participants into four quartiles according to their telomere length. We then used the Jonckheere-Terpstra trend test to compare continuous variables and the Cochran-Armitage test to compare dichotomous variables across the quartiles. In order to evaluate the rate of exacerbation and mortality, aTLs were divided into two groups using the median value as the cut-off, with values above the median defined as longer telomere length and those below as shorter telomere length. SPSS version 21.0 for Windows

(IBM Corp., Armonk, NY, USA) and R Studio version 3 were used for statistical analyses. Two-sided p-values <0.05 were considered significant.

1.2.2.1 aTL and Quality of Life

A linear regression model was constructed to examine the relationship between aTL and SGRQ. To create the multivariable linear regression models with aTL, significant variables from univariable analyses (p<0.1) and variables with biological plausibility were included in the model using a stepwise selection process. The covariables included age, gender, race, body mass index (BMI), current smoking status, total SGRQ score, FEV_1 % predicted and FVC % predicted at baseline.

1.2.2.2 aTL and the Exacerbation Rate

A negative binomial model was used to estimate the annualized rate of exacerbations. The model considered the terms for treatment, baseline COPD exacerbation history (i.e. systemic steroids and/or antibiotics use), smoking status at screening, severity of airflow limitation, and aTL (e.g. shorter or longer) as fixed effects and age and baseline dyspnea grade as covariables. The offset variable ln (exposure time in years) was used. Formal testing for effect modification of Azithromycin (AZ) therapy by aTL group on the exacerbations was performed by including a multiplicative interaction term in a fully adjusted negative binomial model.

1.2.2.3 aTL and Mortality

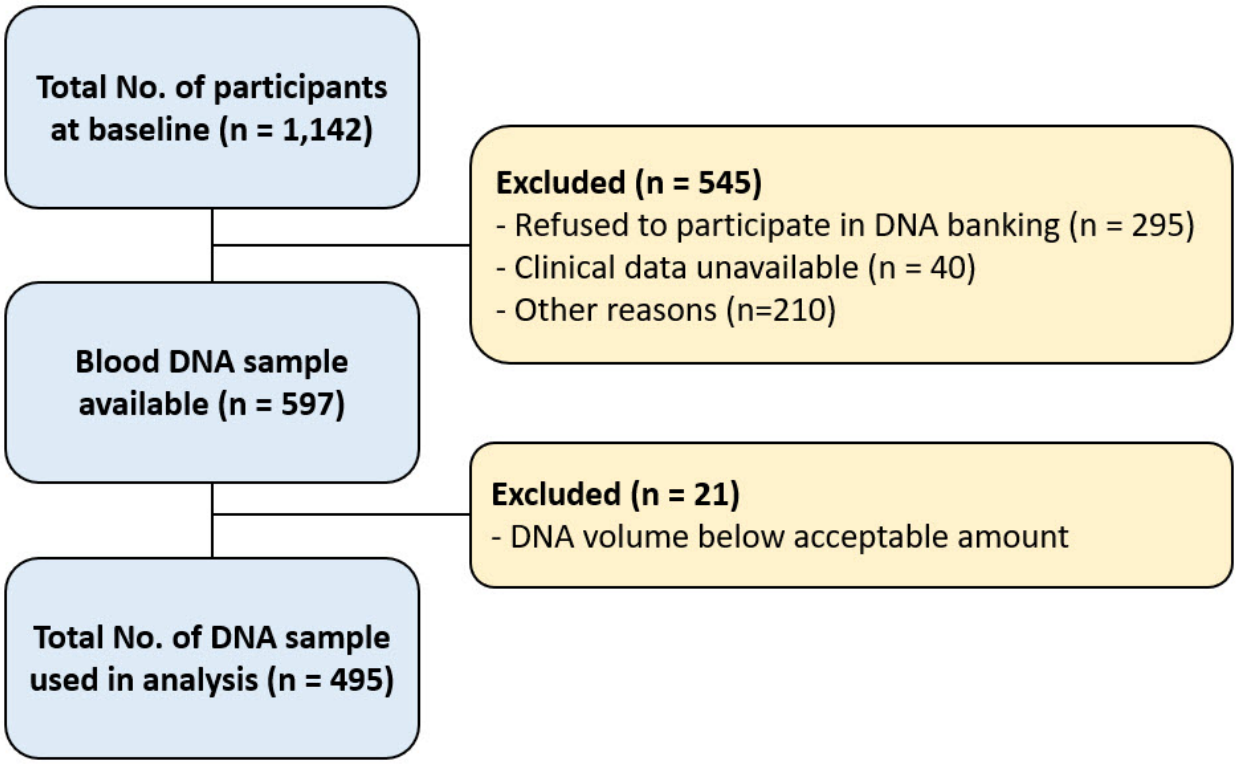
After stratifying the data according to treatment, time to death was compared between shorter and longer aTL groups (based on the median value of aTL) using Kaplan-Meier curves and log-rank

test. A Cox proportional hazards regression was performed to adjust for potential confounders including age, BMI, severity of airflow obstruction, dyspnea grade, smoking status and oxygen use at screening, the rate of exacerbation, and aTL. Formal testing for effect modification of AZ therapy by aTL on mortality was performed by including a multiplicative interaction term in a fully adjusted Cox proportional hazards model.

1.3 Results

A flow diagram of the present study is presented in **Figure 1**. We excluded from analysis participants who did not have DNA collected or had missing clinical data, leaving 576 participants for this study. A summary of the demographic and clinical characteristics of MACRO participants who did and did not have telomeric measurements can be found in the supplement (Appendix A). The mean age of participants at the baseline visit was 66.0 (SD = 8.6) years and 55.6% of the cohort were males. **Table 1** provides clinical and demographic data of participants according to leukocyte telomere length in quartiles. There was a significant trend across aTL for sex (0.027**),** age (p<0.001), the activity score of SGRQ (p=0.006), and FVC % predicted at baseline (p=0.006). The relationship of telomere length with the various domains of SGRQ (activity, symptom and impact) scores are shown in Appendix B.1, B.2 and B.3**,** respectively**.** The relationship of telomere length with FEV_1 and FVC are shown in Appendix C and D**,** respectively. A summary of the comorbidities data Appendix E and the cell proportion data Appendix F can be found in the supplement.

Figure 1 Number of participants excluded from analysis and reasons for exclusion.



A total of 1,142 participants were recruited for the Macrolide Azithromycin for Prevention of Exacerbations of COPD study for the randomization. Of 597 samples of DNA extraction received at the center, 21 samples were excluded because of the low volume after the dilution. A total of 576 samples were generated in this study.

Table 1 Clinical and demographic characteristics of Macrolide Azithromycin for Prevention of Exacerbations of COPD study participants by quartiles of leukocyte telomere length.

	All	Telomere Quartiles (kbp/genome)				
		Q1 (<112.4)	Q2 (112.4-137.5)	Q3 (137.5-171.8)	Q4 (>171.8)	P-trend
No. of participants	576	144	144	144	144	–
Male sex	319 (55.4)	90 (62.5)	80 (55.6)	78 (54.2)	71 (49.3)	**0.027**
Azithromycin group	303 (52.6)	77 (53.5)	77 (53.5)	63 (47.9)	80 (55.6)	**0.062**
Age	66 ± 8.5)	68.7 ± 8.0	67.2 ± 7.9	64.3 ± 8.8	63.8 ± 8.5	**<0.001**
White	488 (84.7)	131 (91.0)	122 (84.7)	113 (78.5)	122 (84.7)	**0.062**
Smoking (pack-years)	57.3 ± 31.1	59.1 ± 33.4	57.9 ± 29.5	54.5 ± 31.0	57.6 ± 30.6	0.402
Current smoking status	107 (18.6)	20 (13.9)	27 (18.8)	28 (19.4)	32 (22.2)	**0.077**
Chronic bronchitis	255 (44.6)	66 (46.5)	67 (47.2)	65 (45.1)	57 (39.6)	0.221
Cancer history	91 (15.8)	23 (16.0)	27 (18.8)	23 (16.0)	18 (12.5)	0.332
BMI (kg/m^2)	27.5 ± 6.0	27.3 ± 6.1	27.2 ± 5.9	27.2 ± 5.8	28.2 ± 6.0	0.203
SGRQ, total score	47.9 ± 15.8	50.2 ± 14.3	47.4 ± 16.0	47.7 ± 15.6	46.4 ± 16.9	**0.050**
Activity score	67.8 ± 19.6	72.4 ± 17.0	66.7 ± 18.9	67.0 ± 20.4	65.1 ± 21.0	**0.006**
Symptom score	59.6 ± 19.4	60.3 ± 19.0	59.8 ± 18.8	60.2 ± 18.7	58.0 ± 21.1	0.503
Impact score	32.9 ± 17.4	34.3 ± 16.2	32.6 ± 17.9	32.8 ± 17.2	32.0 ± 18.5	0.227
Oxygen use	357 (62.0)	98 (68.1)	88 (61.1)	81 (56.3)	90 (62.5)	0.234
CRP (mg/L)	5.1 ± 4.0	5.7 ± 4.4	4.9 ± 3.8	4.9 ± 4.1	5.0 ± 3.6	0.212
FEV_1/FVC %	42.3 ± 12.4	42.5 ± 13.3	42.2 ± 12.0	42.5 ± 12.2	42.1 ± 12.2	0.888
FEV_1, % predicted	39.6 ± 15.7	38.0 ± 16.1	40.4 ± 15.0	40.3 ± 16.1	39.8 ± 15.4	0.307
FVC, % predicted	70.5 ± 18.7	66.1 ± 16.3	72.3 ± 20.1	72.0 ± 19.2	71.8 ± 18.3	**0.006**
LAMA	365 (63.4)	85 (59.0)	96 (66.7)	92 (63.9)	92 (63.9)	0.598
LABA	423 (73.4)	105 (72.9)	109 (75.7)	107 (74.3)	102 (70.8)	0.813
ICS	450 (78.1)	112 (77.8)	112 (77.8)	118 (81.9)	108 (75.0)	0.558

Study participants were divided into four groups based on telomere length, with the first quartile being the shortest and the fourth quartile being the longest. Mean ± SD valudes are given for normally distributed variables, whereasdichotomous data are given as counts (% of column totals). Jonckheere-Terpstra trend test was used for continous variables, and the Cochran-Armitage test was used for dichotomous variables for *P* trend.

1.3.1 aTL and Quality of Life

In the multivariable model, all variables except for age, SGRQ, and FVC% predicted, were not significantly related to aTL and were dropped from the final analysis. Participants with shorter telomere length had worse health status defined by higher SGRQ scores (β =-0.09, p=0.034, **Table 2**).

Table 2 Multiple linear regression analysis using stepwise selection with absolute telomere length.

	Coefficient (95% CI)	Standardized β	*P* value	VIF
(intercept)	5.354 [4.965, 5.744]			
Age	-0.009 [-0.013, -0.006]	-0.211	**<0.001**	1.033
Male sex	0.039 [-0.024, 0.102]	0.053	0.223	1.118
White	0.016 [-0.068, 0.100]	0.016	0.709	1.051
BMI (kg/m^2)	0.003 [-0.002, 0.008]	0.049	0.258	1.141
Current smoking status	0.016 [-0.069 to 0.100]	0.016	0.718	1.217
SGRQ total score	-0.002 [-0.004, 0.000]	-0.090	**0.034**	1.075
FEV_1, % predicted	-0.002 [-0.005, 0.001]	-0.090	0.140	2.211
FVC, % predicted	0.002 [0.000, 0.003]	0.092	**0.027**	1.042

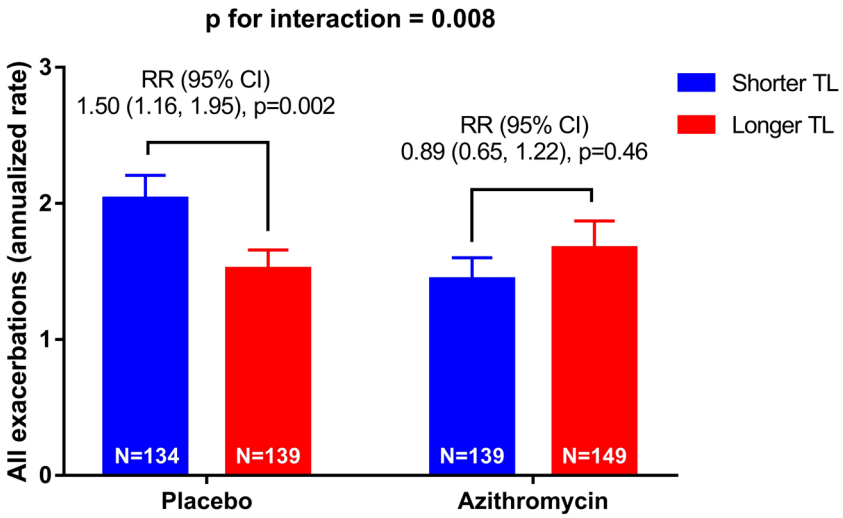
A stepwise regression was analyzed with potential variables with biological plausibility or that showed an association with the outcome of interest (absolute telomere length) to the level of P<0.1 in the univariable analysis: age, sex, race, BMI, current smoking status, total SGRQ score, FEV_1% predicted, and FVC % predicted at baseline. The final variables that had a significant relationship with absolute telomere length were age, total SGRQ score, and FVC % predicted.

1.3.2 aTL and the Exacerbation rate

A total of 890 exacerbations occurred among the 576 COPD patients during follow-up. The annualized rates of exacerbations were not statistically different between shorter and longer aTL

groups (1.71±2.03 vs 1.60±2.11; p=0.417 by Mann-Whitney test). However, after stratifying the data by treatment, the rate of exacerbation was higher in the subjects with shorter aTL in the placebo use group. (p=0.002; rate ratio from negative binomial analysis, 1.50; 95% CI, 1.16 to1.95; **Figure 2.**). A multiplicative interaction term was used in the negative binomial model which demonstrated a significant interaction between treatment and aTL group (p=0.008).

Figure 2 Rate of exacerbation according to leukocyte TL in placebo and azithromycin treatment groups of the Macrolide Azithromycin for Prevention of Exacerbation of COPD study.



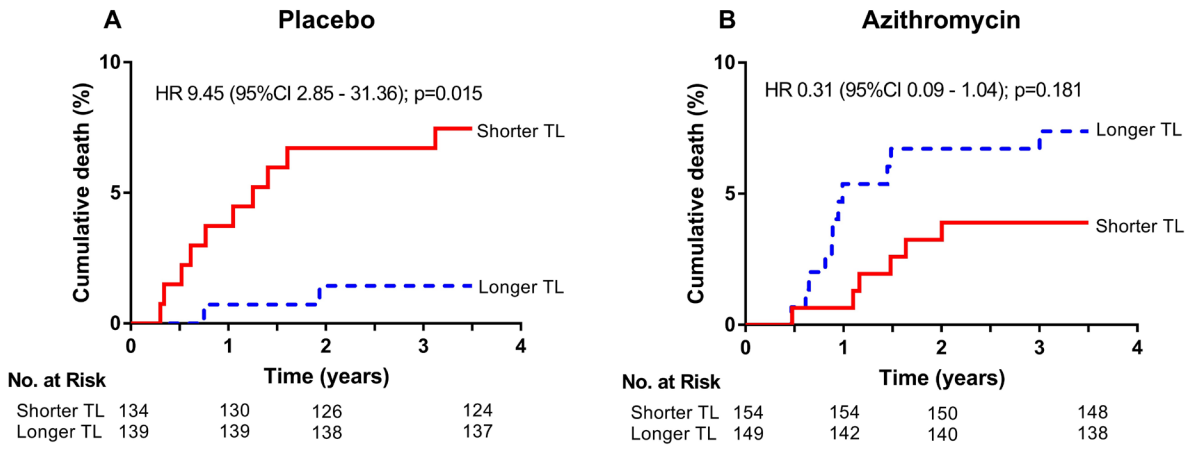
Participants were divided into short vs long telomere groups using the median value (137.5kb/genome). P values were generated using a negative binomial model, and only placebo groups showed significance (P = 0.002). P for interaction is 0.008. Error bars represent the SEM.

1.3.3 aTL and Mortality

When divided into two groups using the median value (137.5kb/gene), those with shorter telomeres experienced a significantly higher risk of total mortality (HR, 9.45; 95 % CI, 2.85-31.36; *p=0.015*)

compared to those with longer telomeres in the placebo group (**Figure 3A**) but not in the azithromycin arm of the study (HR, 0.31; 95 % CI, 0.09-1.04; *p*=0.181; **Figure 3B**). The relationship between aTL and mortality was significant for all participants when aTL was considered as continuous variable (HR, 12.31; 95 % CI, 2.23 - 67.97*; p=0.004*). A multiplicative interaction term was used in the Cox proportional hazard model which demonstrated a significant interaction between treatment and aTL group (*p=0.017*)

Figure 3 A, B, Kaplan-Meier survival curves for all-cause mortality according to TL in peripheral blood in (A) placebo and (B) azithromycin groups in the Macrolide Azithromycin for Prevention of Exacerbations of COPD study.



Participants were divided into short vs long telomere groups using the median value (137.5kb/genome) as the threshold. P values were generated using the log-rank test.

1.4 Discussion

The most important finding from the present study was that short telomere lengths in peripheral leukocytes were associated with reduced quality of life, and a higher risk of exacerbation and death

in patients with moderate to severe COPD. Together, these data support the notion that COPD is a systemic disease of accelerated aging and that replicative senescence, denoted by peripheral blood telomeres, is associated with poor health outcomes in COPD.

There were strengths to the present study. First, we measured absolute, rather than relative, telomere length of peripheral leukocytes, which likely increased the precision of the measurements and most importantly enabled cross-comparisons with other studies. For instance, using population-based data from the CanCOLD (Canadian Chronic Obstructive Lung Disease) Study, we demonstrated previously that the expected median telomere length of peripheral leukocytes for a 66 year old is approximately 150 kbp/genome [44]. In the current study (where the mean age of the cohort was 66 years), the median value was only 137 kbp/genome, which would be the expected value for someone in their 80's, suggesting that on average MACRO participants were biologically much older than their chronological age. Consistent with data from previous studies, we also found a significant relationship of telomere length with age and sex, which is consistent with data from previous studies that have also demonstrated these relationships [43, 45]. Second, in the present study, we related telomere lengths to health outcomes that are of direct importance to patients: exacerbations and health status. Patients in the lowest quartile of telomeres had significantly worse health status than those in the highest quartile, which was best represented by the "activity score" domain of SGRQ. When we excluded data with log aTL less than 4 and outliers defined by the Robust regression and Outlier removal (ROUT) method at 1% false discovery rate[46], all results remained unchanged except for the relationship between aTL and symptoms domain of SGRQ, which became weakly significant (R^2=0.00568; p=0.0406). In the subgroup analysis stratified by sex however, the significant relationship with total score of SGRQ persisted

only in the female population (multivariate regression, p=0.012). Interestingly, patients in the lowest quartile had significantly lower FVC values raising the possibility that reduced telomeres may be a biomarker of "restrictive" physiology of COPD patients. As we did not image or perform lung volume measurement of the lungs of the MACRO participants, we could not determine the precise etiology of reduced FVC. This may reflect suboptimal effort, or represent a large volume of trapped gas, volume derecruitment [47], or a mix of emphysematous and (interstitial) fibrotic changes in the same lung, which may be observed in 8% of smokers with COPD [48]. Irrespective of the etiology, the positive relationship between telomere length and FVC value has been reported in several other studies [49, 50]. Some possible mechanisms include shared genetic predisposition to shorter telomere length and reduced FVC, and inflammation due to exposure to noxious materials (i.e. cigarette smoke) and oxidative stress [51-53]. We also found that the rate of exacerbations was significantly higher in patients with shorter versus longer telomeres in the placebo but not the azithromycin arm of MACRO. While the exact reason for this unknown, we posit that by reducing the overall exacerbation rate, the use of azithromycin abolished the differential rates between those with and without short telomeres. Similar findings were also noted for mortality where those with shorter telomeres had over 9-fold increase in the risk of mortality in the placebo but not in the azithromycin group. However, these data should be interpreted with caution, as the number of deaths during the study period was relatively small, which increases the risk of a type 2 statistical error.

The mechanism by which short telomeres are associated with poor patient outcomes is not entirely clear. It is now well established that replicative senescence results in a change of cellular phenotype to a pro-inflammatory state, a process that has been referred to as senescence-associated secretory

phenotype (SASP). Cells that acquire SASP become metabolically active and secrete a wide-array of different pro-inflammatory proteins [31, 54]. While it was beyond the purview of the present study to determine the exact pathways by which reduced telomeres may contribute to poor health outcomes, previous studies have suggested that a pro-inflammatory phenotype in COPD is associated with increased risk of exacerbations and mortality [55, 56]. Interestingly, despite this, we did not find a significant relationship between telomere length and plasma c-reactive protein (CRP) levels in the current study, which is consistent with previous reports of other COPD cohorts [32, 43]. While CRP is a very sensitive biomarker of systemic inflammation, it lacks specificity to the lungs. A previous study suggests that the relationship of blood telomeres with inflammation improves significantly when a more lung-specific biomarker is used [32]. Additional studies will be needed to validate this hypothesis. Another plausible explanation for the relationship between short telomeres and poor clinical outcomes is that short telomeres may be a biomarker of increased environmental exposures (e.g. from cigarette smoking) over lifetime, leading to greater oxidative stress, [57] which in turn contributes to poor outcomes.

There were also some limitations to the present study. First, this cohort did not include healthy control participants, and only included those with moderate to severe COPD, who were at increased risk of exacerbations were included. Thus, we could not evaluate the relationship of aTL to health outcomes across the full range of subjects from healthy controls to patients across all GOLD stages of COPD [58]. Second, this study measured telomeres of peripheral leukocytes rather than those from lung tissue. This might have led to the underestimation of the effects of telomeres on health outcomes as somatic tissues have generally longer telomeres than leukocytes [59]. However, we have shown previously that the telomere lengths of peripheral leukocytes

significantly correlate with those in lung tissue [30] and most of the previously published studies that have examined telomere length in COPD have used peripheral leukocytes to reflect the telomere lengths in other tissues [1, 31-33, 43]. Third, we used DNA samples extracted from buffy coat to measure the leukocyte telomere length. Peripheral blood tissue is made up of different cell types (e.g. lymphocytes, monocytes and etc) which have different telomere lengths [60]. Although there are significant differences in the telomere lengths across different inflammatory cells, the absolute differences are very small and nearly negligible compared with the inter-individual variation of telomeres [61]. Consistent with this notion, we found that adjustment for cell proportions in our dataset did not materially alter any of the reported relationships (data not shown; available from authors). Lastly, as noted previously, the overall mortality rate during the study period in MACRO was low and as such the relationship between telomeres and death across the two treatment groups should be interpreted cautiously.

.

Chapter 2: The Relationship of DNA Methylation with Lung Function, Quality of Life and Mortality in Chronic Obstructive Pulmonary Disease (COPD)

2.1 DNA Methylation in COPD

Although there is considerable amount of literature on DNA methylation related to cancer and neurogenerative diseases, there is a marked scarcity of such studies in COPD. Moreover, the few studies that have evaluated DNA methylation in COPD have yielded inconsistent data. Epigenome-wide association studies (EWAS) on lung function have revealed some specific CpG sites that were associated with COPD; however, replication of these "hits" in external cohorts has been generally lacking. [62-66]. Some have been poorly designed and analyzed; Marioni et al. found that most of the CpG sites that were significantly related to FEV_1 became insignificant after the adjustment for age, sex, height and smoking status [67].

More focused studies of *global* methylation have produced more consistent results. One common method is to measure repetitive DNA sequences dispersed in the genome [21, 66, 68]. For example, Lange et al., investigated the relationship between lung function and methylation of repetitive DNA elements called Alu and LINE-1 using pyrosequencing. Alu and LINE-1 hypomethylation was significantly associated with lower lung function and with more rapid lung function decline, respectively [69]. Lange et al. also showed that hypermethylation of Alu but not

LINE-1 was significantly related to a lower odds of developing COPD [69]. A major limitation of this study was that it included only male participants and only 16% of the cohort had COPD.

2.1.1 Transposable Elements

Approximately 45% of the human genome consists of the repetitive DNA repeat sequences called transposable elements (TE) or transposons [70, 71]. Transposons have the ability to move around in the genome [70, 71]. It is believed that over time, transposons have contributed to human evolution and diversity by disrupting coding regions or inducing splice signaling, causing DNA rearrangement. For this reason, in the normal human genome, transposons are silenced through DNA methylation to prevent potentially harmful consequences from these effects [70-75]. More than 97% of transposons are retrotransposons which are created through a copy-and-paste mechanism by RNA intermediates, followed by reverse transcription, thus leading to two copies of theses genomic sequences and thereby increasing the size of the genome [73, 76-78]. The remaining are transposons, which are derived from a cut-and-paste mechanism and as such do not increase the size of the genome following transposition [73, 76-78].

2.1.1.1.1 Alu and LINE-1

Both Alu and LINE-1 (L1) are active retrotransposons which take up more than 30% of the human genome [21, 70, 71]. The precise role of Alu and L1 elements in human is not fully elucidated, the most plausible one, however, is their evolutionary role contributing to formation of new genes or gene combinations as they have been accumulated in human genome over millions of years and contributed to human diversity [79-81]. Alu and L1are the most abundant transposons of the small interspersed nuclear elements (SINEs) and the long interspersed nuclear elements (LINEs), respectively [21, 70, 71, 73, 75, 82]. There are approximately 1.4 million Alu and 0.5 million L1 copies throughout the genome, and they are usually highly methylated; because their sequences are CpG rich, their insertions in the genome introduce new methylation sites in those genomic

locations [21, 73, 77, 83, 84]. Given the ubiquitous presence of Alu and L1 in the genome, some have used these repetitive elements as a surrogate marker of global DNA methylation of the genome [21, 83-86]. Importantly, hypomethylation of Alu and/or L1 has been associated with higher genomic instability and alterations in or deregulation of gene expression [77, 87].

In this study, we used DNA methylation pattern of Alu and L1 in the peripheral blood of individuals with or at risk of chronic obstructive pulmonary disease (COPD) as a biomarker for risk of disease progression and poor clinical outcomes in COPD. Specifically, we determined the association of Alu and L1 methylation in peripheral blood, a surrogate marker of global DNA methylation, with lung function measures among individuals with COPD and control subjects (i.e. those without COPD) and their clinical outcomes including mortality and health status as measured by the St. George's Respiratory Questionnaire (SGRQ) scores. We hypothesized that hypomethylation of Alu and L1 elements would be increased in patients with COPD and associated with lower lung function, poorer health status and an increased risk of mortality.

2.2 Methods

2.2.1 Sampling Cohorts

We used two different cohorts for this study: the Canadian Cohort of Obstructive Lung Disease (CanCOLD) representing individuals without COPD or with stable-COPD and the COPD Rapid Transition Program (RTP) representing patients with COPD who were hospitalized for an acute exacerbation (AECOPD). The full details of RTP [88] and CanCOLD [89, 90] have been previously published. A brief description is provided below.

2.2.1.1 Canadian Cohort of Obstructive Lung Disease (CanCOLD)

The Canadian Cohort of Obstructive Lung Disease (CanCOLD) is a population-based study which consists of more than 1,800 individuals who were 40+ years of age and were randomly selected from the Canadian COPD prevalence study (COLD). COLD was initially recruited from 9 sites across 6 provinces in Canada to better understand the prevalence of COPD and its risk factors. Pre- and post-bronchodilator spirometry data, respiratory questionnaires scores such as St. George Respiratory Questionnaire (SGRQ), and blood samples were collected at baseline, at 1.5 years and 3 years of study enrollment. Approximately one third of the CanCOLD participants demonstrated COPD by spirometry (post-bronchodilator FEV_1/FVC ratio of less than 70%) and the rest demonstrated normal lung function (post-bronchodilator FEV_1/FVC ratio $\geq$70% with FEV_1>80% of predicted).

2.2.1.2 COPD Rapid Transition Program (RTP)

The COPD Rapid Transition Program (RTP) is a study with the primary purpose of identifying blood biomarkers that can diagnose acute exacerbations (AECOPD). RTP consists of a total of 568 patients who were over 19 years of age who were admitted to St. Paul's Hospital (SPH) or Vancouver General Hospital in Vancouver, Canada, for the management of AECOPD, or were seen at the COPD clinic at SPH for chronic COPD management. The latter group had stable symptoms and had been free of AECOPD for at least 8 weeks. Participants who were under 19 years of age and experiencing during outpatient COPD clinic visits were excluded. Blood samples were collected from exacerbating patients on days 1, 3 and 7 of the hospitalization, then on days 30 and day 90 post-hospitalization, and from non-exacerbating patients when they had their COPD

clinic visit at SPH. Clinical outcomes including co-morbidities and mortality were assessed at recruitment and through follow up.

2.2.2 Patient Related Outcomes

Health status of participants was assessed using the St. George's Respiratory Questionnaire (SGRQ) [40]. SGRQ was administered at baseline and scores ranged from 0 to 100, with higher scores indicating worse health status. Due to differential enrollment dates, RTP mortality data was standardized at 1 year of follow-up. Lung function decline for each participant from CanCOLD was calculated by taking into account the baseline, visit 1,2, and 3 (i.e. 4 time points across 8 years of follow-up).

2.2.3 Bisulfite Pyrosequencing

DNA was extracted from peripheral whole blood or buffy coat samples from participants in RTP and CanCOLD using the QIAamp® DNA Blood Mini Kit (Qiagen) following manufacturer's recommendations. The extracted DNA was quantified using a NanoDrop™ 8000 spectrophotometer (Thermo Fisher Scientific). Techniques for bisulfite pyrosequencing have been previously published in detail [21, 87]. Briefly, bisulfite conversion was carried out using EZ-96 DNA Methylation™ Kits (Zymo Research). This process converts unmethylated cytosines to uracils while leaving the methylated cytosines intact, distinguishing the two states in subsequent assays. The bisulfite converted DNA was then amplified by polymerase chain reaction (PCR) using the HotStar Taq® DNA Polymerase (Qiagen) kit and the bisulfite converted PCR pyrosequencing was performed using PyroMark Q96 Vacuum Prep Workstation (Qiagen) and the

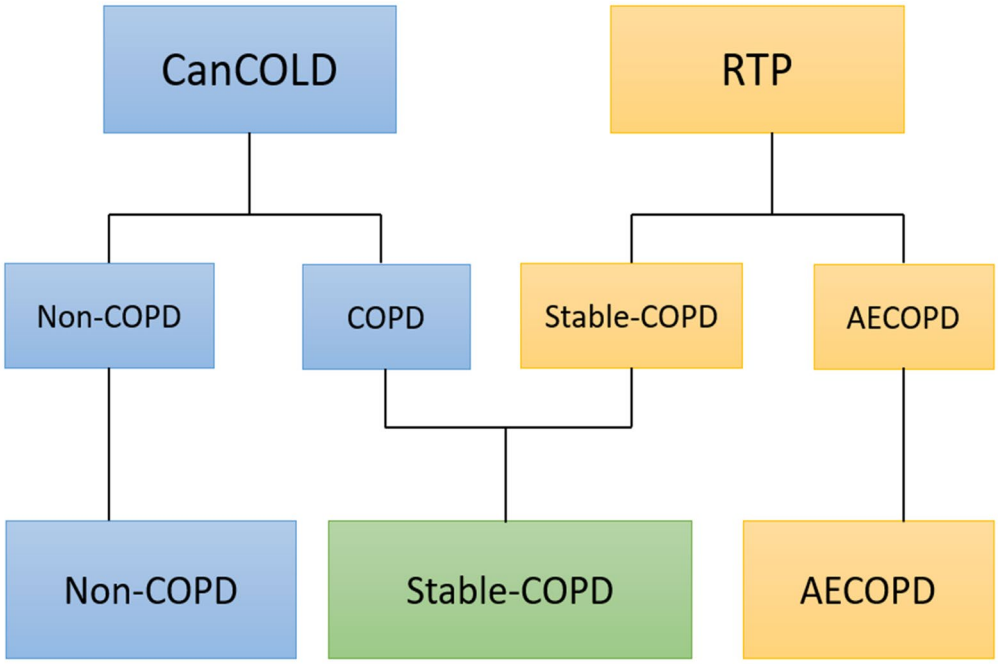
PyroMark Q96 ID pyrosequencer (Qiagen). Percentage of DNA methylation was determined using the pyrosequencing analysis software, PyroMark CpG Software (Qiagen).

2.2.4 Statistical Analysis

SPSS version 25 for Windows (IBM Corp), R version 3 (R Foundation) and Prism version 7 (GraphPad) were used for statistical analyses. Two-sided P-values <0.05 were considered significant. A multivariable linear regression model with a backward stepwise selection process was used to examine the relationship between % methylation of each repetitive element and lung function measures (FEV_1 in liters, FEV_1 % predicted and FEV_1/FVC ratio) or SGRQ scores. Covariates with biological plausibility that were included in the model were: age, sex, race, BMI, current smoking status and pack-years. % Alu methylation was divided into quartiles and analyzed in the same manner as above with FEV_1 in liters and FEV_1 % predicted. A multivariable linear regression that included the previously mentioned covariates as well as baseline lung function was used to determine the relationship between % methylation of each repetitive element and lung function decline, which included four separate lung function measures over 8 years in CanCOLD. The slopes of each lung function measure were first estimated using a linear regression of lung function on time. The relationship between % methylation of each repetitive element and lung function decline were also analyzed with a linear mixed effects (LME) model. To examine the relationship between the % methylation of each repetitive element and mortality rate in RTP, time to death was compared between higher and lower % methylation groups (based on the median value of the % methylation) using Kaplan-Meier curves, and a log-rank test was used for statistical testing. Lastly, the % methylation of each repetitive element from both CanCOLD and RTP was pooled and participants were categorized into three different COPD states: non-COPD, AECOPD

and stable-COPD. Stable-COPD included stable-COPD control participants from RTP and spirometric COPD participants (post-bronchodilator FEV_1/FVC ratio of less than 70%) from CanCOLD (**Figure 4**). The participants were then categorized into current-smokers vs. former- and never-smokers as smoking status was a co-variate that was significantly related to % methylation. The difference of % methylation between COPD states within each smoking status was analyzed using a Kruskal-Wallis H test and each COPD status was compared to one other with Dunn's multiple comparison tests.

Figure 4 Categorization of different states of COPD from pooled study cohorts: CanCOLD and RTP.

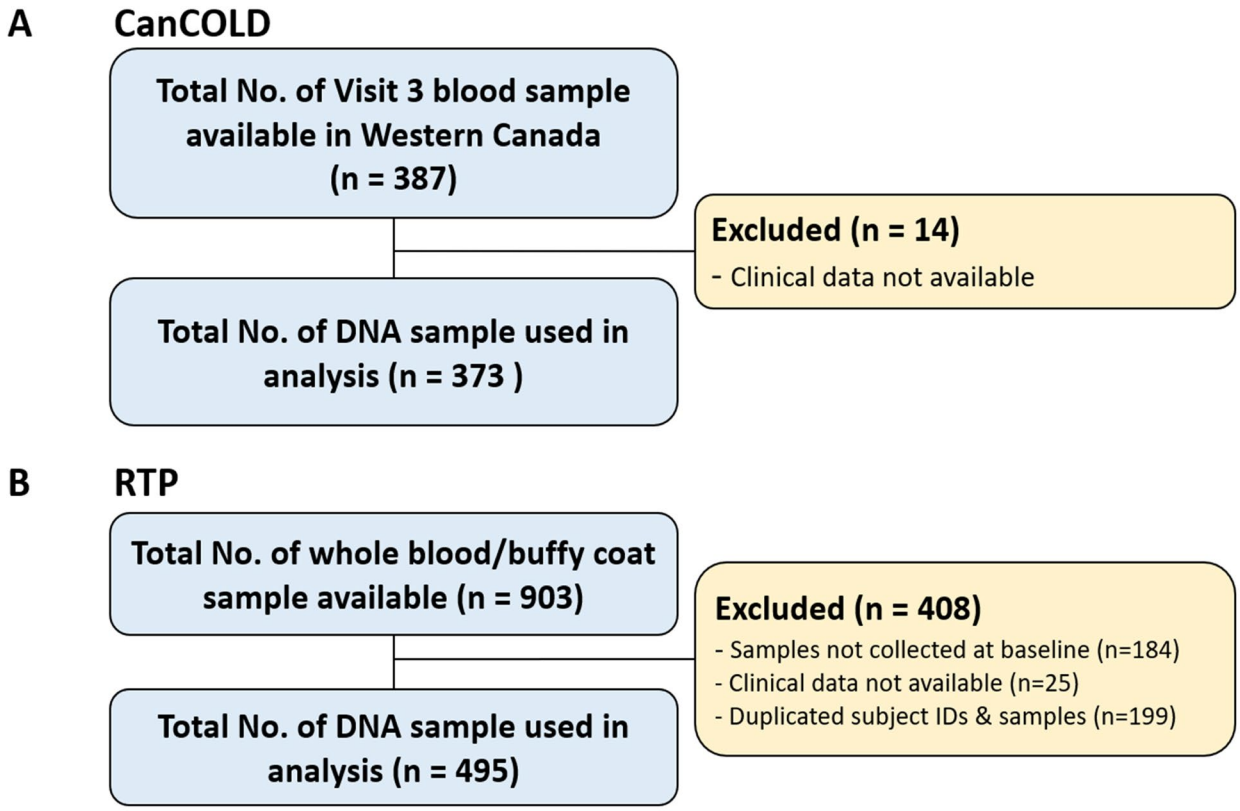


% methylation of each Alu and L1 repetitive element from both CanCOLD and RTP was pooled and participants were categorized into three different COPD states.: non-COPD, AECOPD and stable-COPD. Stable-COPD includes stable-COPD control participants from RTP and physician-diagnosed-COPD participants from CanCOLD.

2.3 Results

A flow diagram of this study is presented in **Figure 5A** and **B.** Participants who did not have clinical data available in CanCOLD or RTP were excluded from the analysis. Duplicated samples from RTP were also excluded from the final analysis.

Figure 5 A, B, Number of participants excluded from analysis and reason for exclusion in (A) Canadian Cohort of Obstructive Lung Disease (CanCOLD) and (B) COPD Rapid Transition Program (RTP).



387 blood samples were available for CanCOLD and 14 samples were excluded because clinical data were not available. A total of 373 CanCOLD samples were used in this analysis. 903 blood/buffy coat samples were available for RTP and 408 samples were excluded because samples were not taken at baseline, clinical data were not available or samples were duplicated. A total of 495 RTP samples were used in this study.

2.3.1 CanCOLD

The mean age of participants from CanCOLD was 69.72±9.99 years and 56.57% of the cohort were men. **Table 3** provides clinical demographic data of participant including % methylation of Alu and L1.

Table 3 Clinical and demographic characteristics of Canadian Cohort of Obstructive Lung Disease study participants.

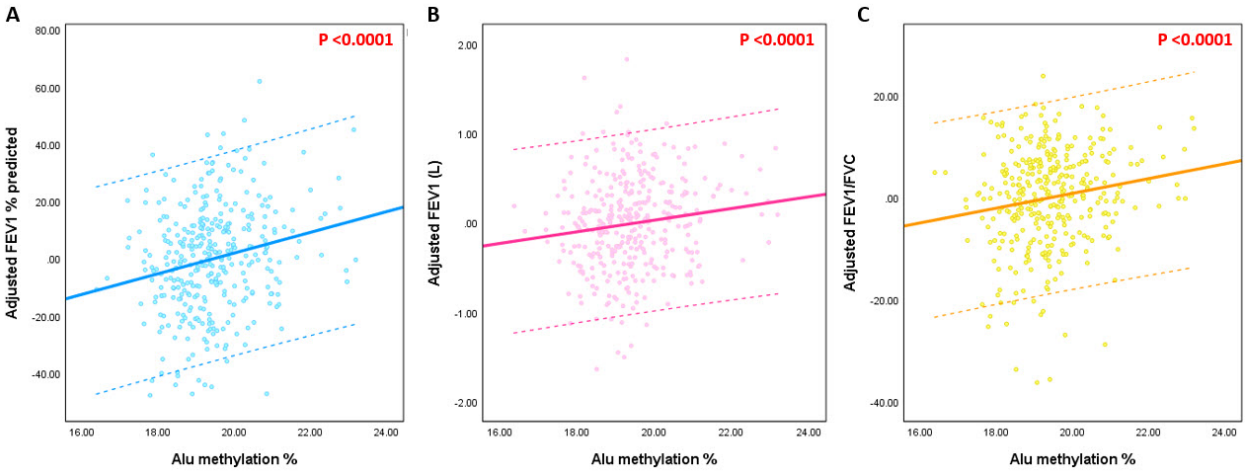
	Mean (SD) or N (%)
No. of total participants	373
Male sex, N (%)	211 (56.57)
Age (years) Range Mean (SD)	 43-92 69.72 (9.99)
White, N(%)	349 (93.6)
Smoking Status, N (%) Current Former Never	 54 (14.48) 176 (47.18) 143 (38.34)
Pack-years (SD)	14.04 (19.86)
BMI (kg/m^2) (SD)	27.15 (5.04)
FEV_1 in L (SD)	2.48 (0.75)
FEV_1%predicted (SD)	91.37 (19.76)
FEV_1/FVC% (SD)	68.03 (10.30)
SGRQ Score (SD) Symptom Activity Impact Total	 18.95 (18.52) 15.63 (20.00) 6.08 (10.61) 11.11 (13.22)
Spirometric COPD, N (%)	202 (54.2)
Coronary Artery Disease, N (%)	25 (6.70)
Myocardial Infarction, N (%)	20 (5.36)
% Alu methylation (SD)	19.49 (1.49)
% LINE-1 methylation (SD)	76.14 (3.12)

Mean (SD) values are given for continuous variables whereas dichotomous data are given as counts (% of total).

2.3.1.1 Percentage (%) Methylation and Lung Function Measures

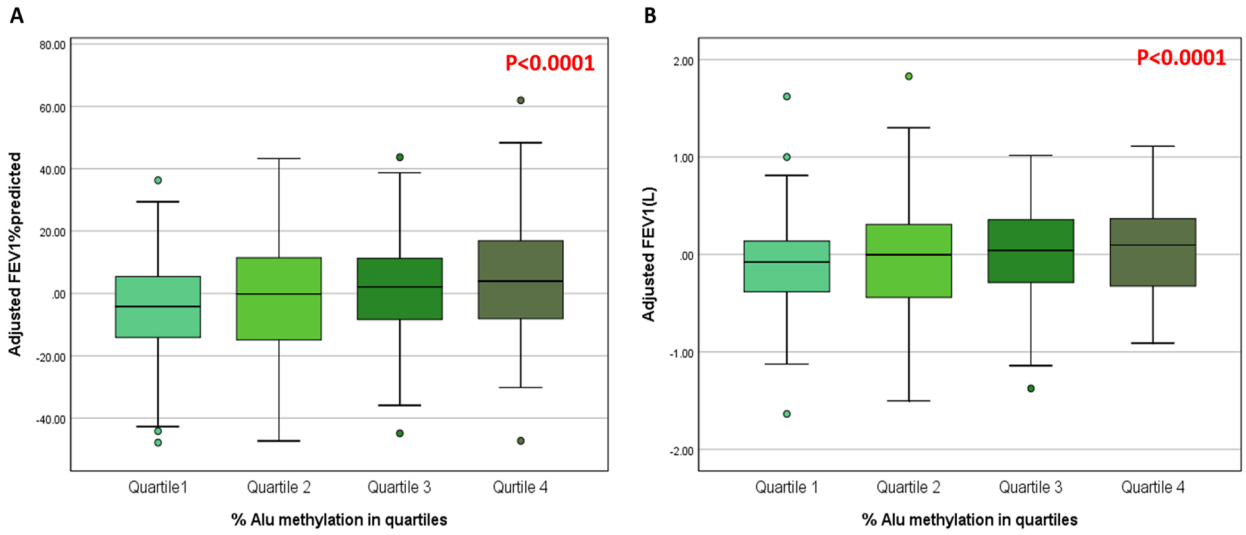
% Alu methylation was significantly related to FEV_1 % predicted, FEV_1 in liters (L) and FEV_1/FVC ratio (β = 0.205, 0.098 & 0.158, respectively, P<0.0001) (**Figure 6A and B**). With the multivariable linear regression with a backward stepwise selection process, age, sex, BMI, current smoking status and pack-years were considered as co-variables. Co-variables that were not significantly related to % Alu methylation were not retained from the final analysis. Current smoking status was not retained in the final analysis with FEV_1 % predicted and BMI and current smoking status were not retained in the final analyses with FEV_1 (L) and FEV_1/FVC ratio. % Alu methylation was analyzed based on quartiles with Quartile 4 indicating higher methylation, and as a continuous variable. The higher the % Alu methylation quartile, the better the lung function measures (P < 0.0001, β = 2.20 for FEV_1 % predicted & β = 0.085 for FEV_1 in liters (L)) (**Figure 7A and B**). % Alu methylation was significantly related to decline of FEV_1 % predicted (%/year) (β = 2.20, P = 0.004), and FEV_1/FVC ratio (%/year) (β = 1.01, P = 0.007) (**Figure 8A, B and C**) based on linear regression (slope) model. LME analysis demonstrated that % Alu methylation was significantly related to FEV_1 % predicted decline (%/year) (β = 0.531, P = 0.0073). L1 methylation status was not significantly related to cross-sectional lung function measures or lung function decline (Appendix G 1, 2, 3 and 4).

Figure 6A, B and C Results of multivariable linear regression analyses using a stepwise selection between percentage Alu methylation and lung function measures: (A) FEV_1 % predicted, (B) FEV_1 in liters and (C) FEV1/FVC.



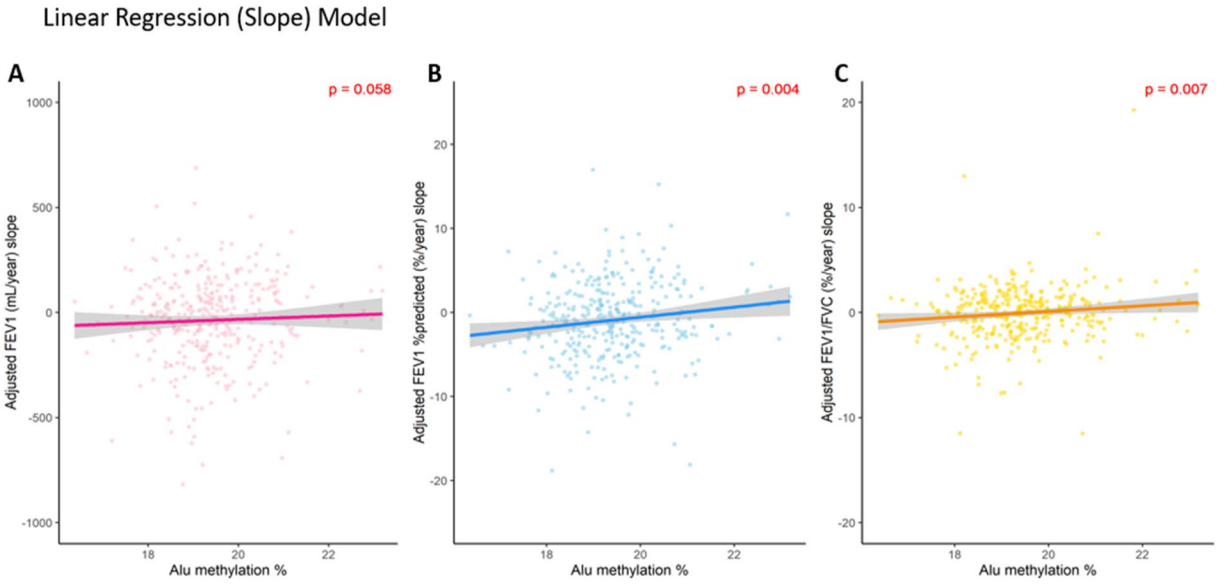
Stepwise regression was analyzed with age, sex, race, BMI, smoking-status and pack-years. The final model included age, sex, BMI, race and pack-years for FEV_1% predicted and FEV_1 in liters; and age, sex and pack-years for FEV_1/FVC. Broken lines represent 95% confidence interval.

Figure 7A and B Results of multivariable linear regression analyses using a stepwise selection between percentage Alu methylation in quartiles and (A) FEV1 % predicted and (B) FEV1 in liters.



% Alu methylation was divided into quartiles with Quartile 4 indicating higher methylation and a stepwise regression was analyzed with age, sex, race, BMI, smoking-status and pack-years. The final model for Alu methylation included age, sex, BMI, race and pack-years for FEV_1% predicted and FEV_1 in liters and compared with FEV_1 % predicted and FEV_1(L) with the same manner as above, the higher the % Alu methylation quartile, the better the lung function measures.

Figure 8A, B and C The relationship between percentage Alu methylation and lung function decline in FEV_1 (mL/year), FEV_1 % predicted (%/year) and FEV_1 /FVC (%/year) using multivariable linear regression analyses (A, B and C) and linear mixed effect analyses (D).



D Linear Mixed Effect (LME) Model

Methylation %	Lung function decline	Estimated beta	Standard error	P-value
Alu	FEV_1 (mL/year)	0.00422	0.00634	0.506
Alu	**FEV_1 % predicted (%/year)**	0.531	0.1977	**0.0073**
Alu	FEV_1/FVC (%/year)	0.182	0.1011	0.072

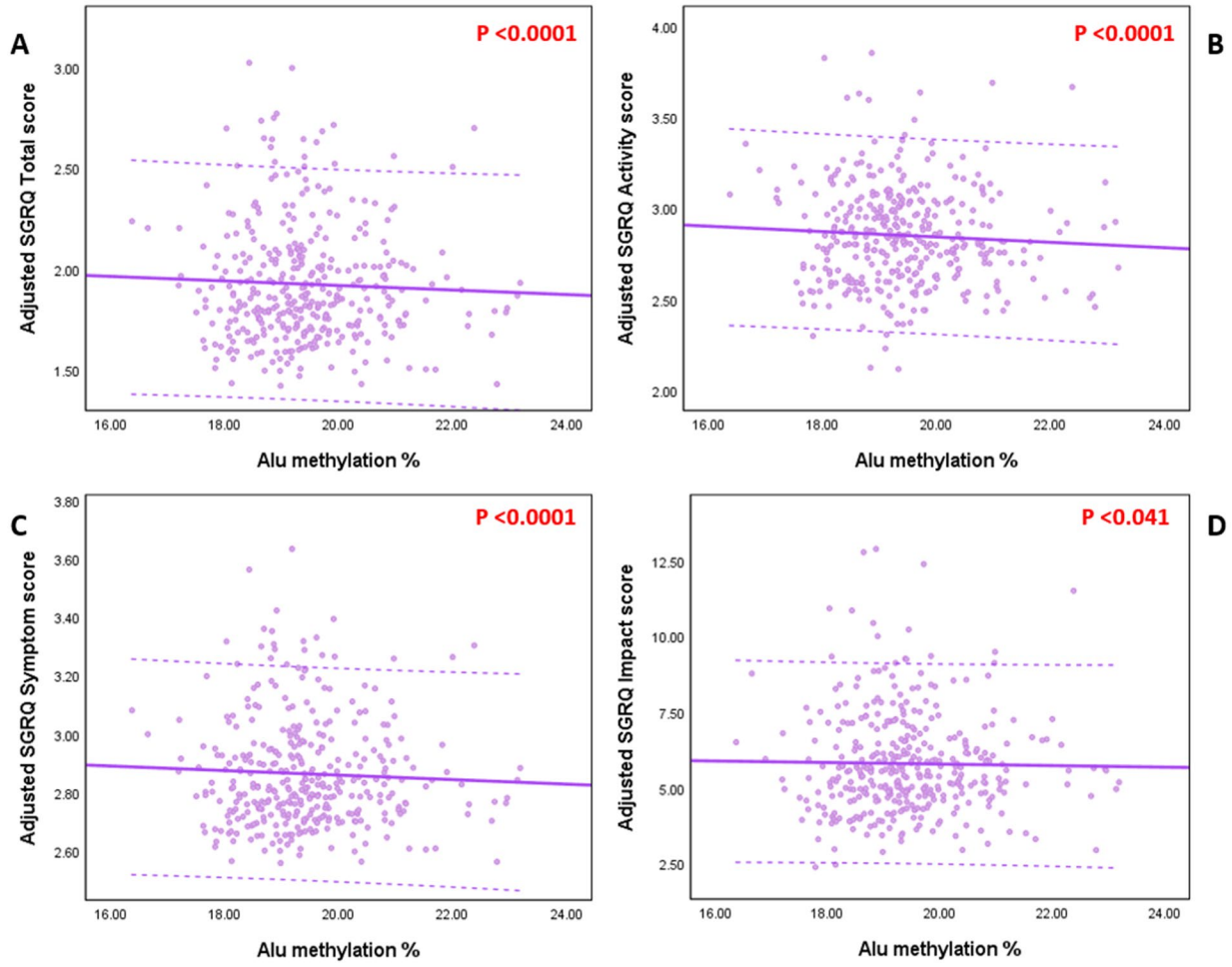
A multivariable linear regression was analyzed with co-variables: age, sex, race, BMI, smoking-status, pack-years and baseline lung function. The slope of each lung function measures was estimated using a linear regression of four lung function measures over 8 years. A linear mixed effect was analyzed with co-variables: age, sex, race, BMI, smoking-status, pack-years and baseline lung function.

2.3.1.2 % Methylation and Quality of Life

SGRQ Total as well as the sub-domain Activity, Symptom and Impact scores were significantly associated with % Alu methylation (β =-0.116, -0.138, -0.152, & -0.135, respectively, *P* <0.0001 except SGRQ Impact score *P*=0.041). (**Figure 9A, B, C and D**) In the multivariable linear regression that used a backward stepwise selection process retained co-variables: age, BMI and

current smoking status in the final analysis that evaluated the relationship with SGRQ Activity score. Similarly, BMI and pack-years were retained in the final analyses that evaluated the relationship with SGRQ Symptom and Total scores, and BMI was retained in the final analysis that evaluated the relationship with SGRQ Impact score. L1 % methylation was not significantly related to SGRQ scores.

Figure 9A, B, C and D Results of multivariable linear regression analyses using stepwise selection between percentage Alu methylation and (A) SGRQ Total scores, (B) SGRQ Activity scores, (C) SGRQ Symptom scores and (D) SGRQ Impact scores.



The final model included BMI and pack-years for Total and Symptom scores, age, BMI and smoking status for Activity scores, and BMI for Impact scores.

2.3.2 RTP

The mean age of participants from RTP was 66.75±11.82 years, and 63.37% of the study were men. Table 4 provides clinical demographic data of participant including % methylation of Alu and L1. The cohort was divided into two groups based on using the median value, 19.31 and 73.45, respectively and there was no significant relationship between those with higher and lower % methylation (Appendix H 1 and 2).

Table 4 Clinical and demographic characteristics of COPD Rapid Transition program participants.

	Mean (SD) or N (%)
No. of total participants	495
No. of stable COPD (control) participants, N (%)	72 (14.5)
Male sex, N (%)	310 (62.37)
Age (years) Range Mean (SD)	 20-97 66.75 (11.82)
White, N (%)	409 (82.29)
Smoking Status, N (%) Current Former Never	 279 (56.36) 181 (36.57) 35 (7.07)
Cancer, N (%)	79 (15.90)
Myocardial Infarction, N (%)	75 (15.10)
Coronary Artery Disease, N (%)	98 (19.72)
% Alu methylation (SD)	19.17 (1.67)
% LINE-1 methylation (SD)	73.10 (3.90)

Mean (SD) values are given for continuous variables whereas dichotomous data are given as counts (% of total).

2.3.3 CanCOLD and RTP

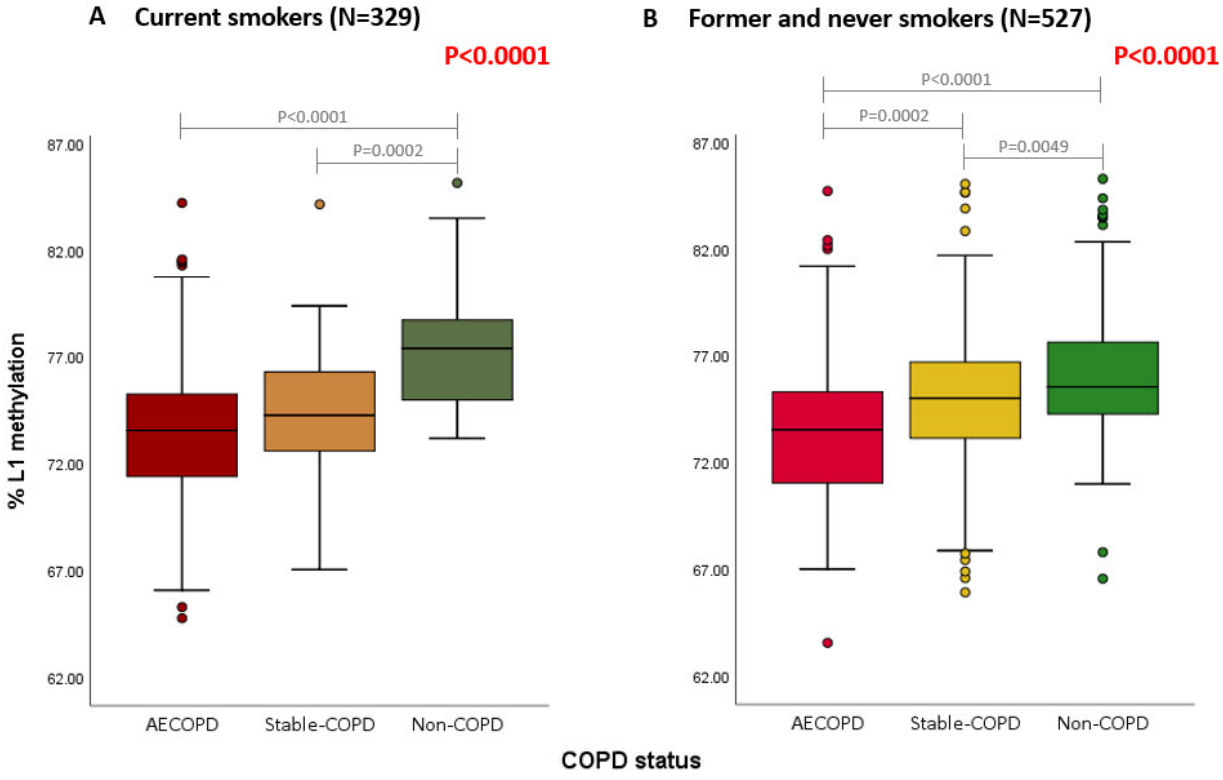
When the data from CanCOLD and RTP were pooled, the total number of participants was 857. **Table 5** provides demographics and clinical characteristics of all and stratified participants by different states of COPD. The mean age of all participants was 68.03±11.16 years, and 60% of them were men. In a univariate correlation analysis, smoking-status was significantly related to % methylation (P<0.0001), therefore, participants were divided into current-smokers vs. former- and never-smokers. % L1 methylation was significantly different between COPD states in both current-smokers and former-/never-smokers where non-COPD participants had the highest, and AECOPD participants had the lowest % methylation (P<0.0001) (**Figure 10**). Within current-smokers, participants with non-COPD had significantly higher % L1 methylation compared to those with AECOPD (P<0.0001) and stable-COPD (P=0.0002). Within former- and never-smokers, all COPD states were significantly related to one another: P<0.0001 between AECOPD and non-COPD, P=0.0002 between AECOPD and stable-COPD, and P=0.0049 between stable-COPD and non-COPD. % Alu methylation was not significantly different between different COPD states.

Table 5 Clinical and demographic characteristics of all and stratified participants by different states of COPD from Canadian Cohort of Obstructive Lung disease and COPD Rapid Transition program participants.

COPD Status	**All**	**AECOPD**	**Stable-COPD**	**Non-COPD**
No. of participants	857	425 (49.6)	198 (23.1)	234 (27.3)
Age (years) Range Mean (SD)	 20-97 68.03 (11.16)	 31-97 66.92 (11.75)	 20-92 69.97 (11.23)	 45-91 68.20 (9.61)
Male sex, N (%)	521 (60.8)	265 (62.4)	122 (61.6)	130 (55.6)
White, N (%)	758 (88.4)	343 (80.7)	186 (93.9)	218 (93.2)
Smoking Status, N (%) Current Former Never	 331 (38.6) 349 (40.7) 177 (20.7)	 250 (58.8) 145 (34.1) 30 (7.1)	 48 (24.2) 99 (50) 51 (25.8)	 33 (14.1) 105 (44.9) 96 (41)
Coronary Artery Disease, N (%)	123 (14.4)	91 (21.4)	16 (8.01)	16 (6.8)
% Alu methylation (SD)	19.31 (1.33)	19.21 (1.48)	19.26 (1.1)	19.47 (1.1)
% LINE-1 methylation (SD)	74.53 (3.45)	73.38 (3.26)	74.94 (3.5)	76.21 (2.9)

Mean (SD) values are given for continuous variables whereas dichotomous data are given as counts (% of total).

Figure 10A and B Comparisons of percentage LINE-1 methylation between non-, stable- and AE-COPD in (A) current smokers and (B) former and never smokers pooled from Canadian Cohort of Obstructive Lung Disease (CanCOLD) study and COPD Rapid Transition program (RTP).



Non-COPD participants from CanCOLD study were categorized as 'Non-COPD', AECOPD participants from RTP were categorized as 'AECOPD' and stable-COPD control from RTP and physician-diagnosed-COPD participants from CanCOLD study were categorized as 'Stable-COPD'. The difference of % methylation between COPD statuses was analyzed using Kruskal-Wallis H test and each status was compared with one another with Dunn's multiple comparison tests.

2.4 Discussion

The most important finding from this study was that hypomethylation of Alu elements in peripheral leukocytes was associated with lower lung function, faster lung function decline and the reduced quality of life in patients with COPD. We also found that L1 was the most de-

methylated in individuals during AECOPD compared to those with stable COPD. Healthy control subjects demonstrated the highest methylation of L1.

One of the strengths in this study was that we used real-time DNA sequencing technology called bisulfite-pyrosequencing which is known as a "gold standard" for the measurement of DNA methylation at a single base resolution [68, 91]. Compared to gene-specific CpG analysis such as epigenome-wide association studies (EWAS), pyrosequencing analyzes the level of 'global' methylation throughout the genome at a lower cost with a faster turnaround time and greater accuracy [91, 92]. Pyrosequencing is especially useful in assessing *hypo*methylation of highly methylated regions of the genome such as Alu and L1. Yang et al. have previously validated pyrosequencing for Alu and L1 by examining the Alu and L1 methylation in cell lines before and after treatment with a methylation inhibitor [21]. We found a significant relationship of Alu hypomethylation with lung function, which is consistent with data from a previous study that used pyrosequencing [69]. Although Lange et al. found a significant relationship of L1 hypomethylation with more rapid lung function decline, we found a significant relationship only with Alu hypomethylation. We did not find a significant relationship between % L1 methylation and lung function. However, when we evaluated participants from both CanCOLD and RTP cohorts, we found that L1 methylation status but not Alu methylation was significantly related to AECOPD. Together, these data suggest that Alu methylation may be a biomarker for lung function and lung function decline while L1 methylation may be more useful for identifying AECOPD.

The other strength in this study was that we related % Alu methylation to patients' health status as measured by the SGRQ score. Participants with lower % Alu methylation had significantly worse

health status than those with higher %, for all SGRQ sub-domains including: *Activity*, *Symptom* and *Impact*.

Global DNA methylation changes over time [93, 94]. Although the mechanism by which methylation patterns are altered is not fully known, it is widely accepted that hypomethylation of retrotransposons is associated with increased genomic instability that can lead to a higher rate of mutations by disrupting transcription factor binding sites or altering regulatory sequences [77, 87, 95, 96]. Genetic influence is one of the factors that can alter global DNA methylation pattern [93, 94]. Previous studies have found that DNA methylation pattern is more similar in monozygotic twins compared with dizygotic twins [93] and there is also a familial pattern of changes in DNA methylation with aging [94]. Furthermore, previous EWAS have identified a large number of single nucleotide polymorphism (SNP)s that are associated with specific DNA methylation patterns, which suggests that DNA methylation (and de-methylation) is in part controlled by genetics [97, 98].

Environmental exposure such as cigarette smoke can also influence DNA methylation pattern. It has been shown that smoking is significantly related to global hypomethylation of repetitive elements [77, 99], and EWAS studies have found specific sites of hypomethylation in smokers, [100, 101] which persisted even after smoking cessation [101]. Cigarette smoke can contribute to DNA hypomethylation by modifying DNA-binding factors [102]. Transcription factor specificity protein 1 (Sp1) is a general transcription factor, which is involved in many cellular processes including cell proliferation and death [103, 104] and its expression and DNA-binding capacity are modified by cigarette smoke [105, 106]. Once activated, Sp1 binds to CpG-rich motifs within gene promoters and prevents CpG sites from being methylated [102, 103, 107].

Furthermore, cigarette smoke may cause DNA hypomethylation through hypoxia which is caused when carbon monoxide in cigarette smoke competitively binds to hemoglobin and thus displacing oxygen [102]. *S*-adenosylmethionine (SAM) is a major methyl donor and its level is considered a marker of methylation status [108-110]. Hypoxia induces DNA hypomethylation by significantly decreasing intracellular levels of SAM level *in vitro* and *in vivo*, although the exact mechanism of SAM downregulation is not completely understood. It should be noted, however, that these results were derived from hepatoma cell lines and liver tissue retrieved from mice or rats. Whether these data apply to human lung tissue requires further studies. Other environmental exposures such as air or organic pollutants that can insult the lung directly and are known to reactivate transposons by causing cellular stress [111, 112] have also been associated with global DNA hypomethylation [113-116].

Alu and L1 sequences are CpG rich. In basal physiological states, retrotransposons are heavily methylated and kept silenced, however, oxidative stress/damage, which is implicated in COPD pathogenesis [117], is known to interfere with the binding of methyl-CpG binding proteins leading to hypomethylation [118]. Aberrant methylation in small airways in COPD could modify gene expression and pathways that are important to COPD pathogenesis, such as the nuclear factor erythroid 2-related factor 2(Nrf2) oxidative response pathway [119].

A limitation to this study is that we measured % methylation in peripheral leukocytes rather than that in the lung tissue. Previous studies have found that DNA methylation in the blood is associated with that in tissue [120] including the lung tissue [121] and may be used as a surrogate marker. However, leukocyte DNA methylation status tends to underestimate the level of methylation in

tissue [85]. Future studies are needed to confirm strength of the relationship between the lung specific cells and tissue with the degree of DNA methylation.

Conclusion

In summary, we have shown that moderate to severe COPD participants who had shorter leukocyte telomere lengths were more likely to have poor quality of life with a higher risk of exacerbation and death. Moreover, we have shown that hypomethylation of Alu elements in peripheral leukocytes was associated with lower lung function, faster lung function decline and reduced quality of life in patients with COPD. We also showed that AECOPD participants had the least methylated L1 elements. Telomeres and Alu and L1 elements in peripheral leukocyte are readily accessible and easy to measure; thus, may represent a clinically translatable biomarker for patient risk-stratification and identifying individuals at increased risk of poor patient-centered outcomes in COPD. Additional work will be needed to develop these measurements as clinically available blood tests for clinical application

Bibliography

1. Lee, J., et al., *The relationship between telomere length and mortality in chronic obstructive pulmonary disease (COPD).* PLoS One, 2012. **7**(4): p. e35567.
2. Mapel, D.W., et al., *Health care utilization in chronic obstructive pulmonary disease. A case-control study in a health maintenance organization.* Archives of Internal Medicine, 2000. **160**(17): p. 2653-8.
3. Soriano, J.B., et al., *Patterns of comorbidities in newly diagnosed COPD and asthma in primary care.* Chest, 2005. **128**(4): p. 2099-107.
4. Valdes, A.M., et al., *Telomere length in leukocytes correlates with bone mineral density and is shorter in women with osteoporosis.* Osteoporosis International, 2007. **18**(9): p. 1203-10.
5. Hornsby, P.J., *Telomerase and the aging process.* Experimental Gerontology, 2007. **42**(7): p. 575-81.
6. Shammas, M.A., *Telomeres, lifestyle, cancer, and aging.* Curr Opin Clin Nutr Metab Care, 2011. **14**(1): p. 28-34.
7. Greider, C.W., *Telomere length regulation.* Annual Review of Biochemistry, 1996. **65**: p. 337-65.
8. Bodnar, A.G., et al., *Extension of life-span by introduction of telomerase into normal human cells.* Science, 1998. **279**(5349): p. 349-52.
9. Shin, J.S., et al., *The role of telomeres and telomerase in the pathology of human cancer and aging.* Pathology, 2006. **38**(2): p. 103-13.
10. Mena, S., A. Ortega, and J.M. Estrela, *Oxidative stress in environmental-induced carcinogenesis.* Mutation Research, 2009. **674**(1-2): p. 36-44.
11. Yanbaeva, D.G., et al., *Systemic effects of smoking.* Chest, 2007. **131**(5): p. 1557-66.
12. Lopez-Otin, C., et al., *The hallmarks of aging.* Cell, 2013. **153**(6): p. 1194-217.
13. Franceschi, C., et al., *The Continuum of Aging and Age-Related Diseases: Common Mechanisms but Different Rates.* Frontiers in Medicine, 2018. **5**(61).
14. Jin, M., et al., *Relationship of Absolute Telomere Length With Quality of Life, Exacerbations, and Mortality in COPD.* Chest, 2018. **154**(2): p. 266-273.
15. Moore, L.D., T. Le, and G. Fan, *DNA methylation and its basic function.* Neuropsychopharmacology, 2013. **38**(1): p. 23-38.
16. Holliday, R. and J.E. Pugh, *DNA modification mechanisms and gene activity during development.* Science, 1975. **187**(4173): p. 226-32.
17. Compere, S.J. and R.D. Palmiter, *DNA methylation controls the inducibility of the mouse metallothionein-I gene lymphoid cells.* Cell, 1981. **25**(1): p. 233-40.
18. Field, A.E., et al., *DNA Methylation Clocks in Aging: Categories, Causes, and Consequences.* Molecular Cell, 2018. **71**(6): p. 882-895.
19. Fraga, M.F., et al., *Epigenetic differences arise during the lifetime of monozygotic twins.* Proc Natl Acad Sci U S A, 2005. **102**(30): p. 10604-9.
20. Wilson, A.S., B.E. Power, and P.L. Molloy, *DNA hypomethylation and human diseases.* Biochim Biophys Acta, 2007. **1775**(1): p. 138-62.
21. Yang, A.S., et al., *A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA elements.* Nucleic Acids Res, 2004. **32**(3): p. e38.

22. Levine, M.E., et al., *DNA methylation age of blood predicts future onset of lung cancer in the women's health initiative.* Aging (Albany NY), 2015. **7**(9): p. 690-700.
23. Kim, M., et al., *DNA Methylation as a Biomarker for Cardiovascular Disease Risk.* PLOS ONE, 2010. **5**(3): p. e9692.
24. Baccarelli, A., et al., *Ischemic heart disease and stroke in relation to blood DNA methylation.* Epidemiology, 2010. **21**(6): p. 819-28.
25. Levine, A.J., et al., *Accelerated epigenetic aging in brain is associated with pre-mortem HIV-associated neurocognitive disorders.* J Neurovirol, 2016. **22**(3): p. 366-75.
26. Weng, Y.L., et al., *DNA modifications and neurological disorders.* Neurotherapeutics, 2013. **10**(4): p. 556-67.
27. Lu, H., et al., *DNA methylation, a hand behind neurodegenerative diseases.* Front Aging Neurosci, 2013. **5**: p. 85.
28. Perna, L., et al., *Epigenetic age acceleration predicts cancer, cardiovascular, and all-cause mortality in a German case cohort.* Clin Epigenetics, 2016. **8**: p. 64.
29. Marioni, R.E., et al., *DNA methylation age of blood predicts all-cause mortality in later life.* Genome Biol, 2015. **16**: p. 25.
30. Saferali, A., et al., *Longer telomere length in COPD patients with alpha1-antitrypsin deficiency independent of lung function.* PLoS One, 2014. **9**(4): p. e95600.
31. Houben, J.M., et al., *Telomere shortening in chronic obstructive pulmonary disease.* Respiratory Medicine, 2009. **103**(2): p. 230-6.
32. Mui, T.S., et al., *Telomere length and chronic obstructive pulmonary disease: evidence of accelerated aging.* Journal of the American Geriatrics Society, 2009. **57**(12): p. 2372-4.
33. Savale, L., et al., *Shortened telomeres in circulating leukocytes of patients with chronic obstructive pulmonary disease.* Am J Respir Crit Care Med, 2009. **179**(7): p. 566-71.
34. Brouilette, S.W., et al., *Telomere length, risk of coronary heart disease, and statin treatment in the West of Scotland Primary Prevention Study: a nested case-control study.* Lancet, 2007. **369**(9556): p. 107-14.
35. Fitzpatrick, A.L., et al., *Leukocyte telomere length and cardiovascular disease in the cardiovascular health study.* American Journal of Epidemiology, 2007. **165**(1): p. 14-21.
36. van der Harst, P., et al., *Telomere length of circulating leukocytes is decreased in patients with chronic heart failure.* Journal of the American College of Cardiology, 2007. **49**(13): p. 1459-64.
37. McGrath, M., et al., *Telomere length, cigarette smoking, and bladder cancer risk in men and women.* Cancer Epidemiology, Biomarkers and Prevention, 2007. **16**(4): p. 815-9.
38. Wu, X., et al., *Telomere dysfunction: a potential cancer predisposition factor.* Journal of the National Cancer Institute, 2003. **95**(16): p. 1211-8.
39. Albert, R.K., et al., *Azithromycin for Prevention of Exacerbations of COPD.* The New England journal of medicine, 2011. **365**(8): p. 689-698.
40. Jones, P.W., F.H. Quirk, and C.M. Baveystock, *The St George's Respiratory Questionnaire.* Respir Med, 1991. **85 Suppl B**: p. 25-31; discussion 33-7.
41. Niewoehner, D.E., et al., *Prevention of exacerbations of chronic obstructive pulmonary disease with tiotropium, a once-daily inhaled anticholinergic bronchodilator: a randomized trial.* Annals of Internal Medicine, 2005. **143**(5): p. 317-26.
42. O'Callaghan, N.J. and M. Fenech, *A quantitative PCR method for measuring absolute telomere length.* Biol Proced Online, 2011. **13**: p. 3.

43. Liu, J.C., et al., *Absolute leukocyte telomere length in HIV-infected and uninfected individuals: evidence of accelerated cell senescence in HIV-associated chronic obstructive pulmonary disease.* PLoS One, 2015. **10**(4): p. e0124426.
44. Leung, J.M., et al., *The determinants of poor respiratory health status in adults living with human immunodeficiency virus infection.* Aids Patient Care and STDS, 2014. **28**(5): p. 240-7.
45. Gardner, M., et al., *Gender and telomere length: systematic review and meta-analysis.* Exp Gerontol, 2014. **51**: p. 15-27.
46. Motulsky, H.J. and R.E. Brown, *Detecting outliers when fitting data with nonlinear regression – a new method based on robust nonlinear regression and the false discovery rate.* BMC Bioinformatics, 2006. **7**(1): p. 123.
47. Hyatt, R.E., et al., *Conditions associated with an abnormal nonspecific pattern of pulmonary function tests.* Chest, 2009. **135**(2): p. 419-424.
48. Washko, G.R., et al., *Lung volumes and emphysema in smokers with interstitial lung abnormalities.* N Engl J Med, 2011. **364**(10): p. 897-906.
49. Rode, L., et al., *Short telomere length, lung function and chronic obstructive pulmonary disease in 46,396 individuals.* Thorax, 2013. **68**(5): p. 429-35.
50. Albrecht, E., et al., *Telomere length in circulating leukocytes is associated with lung function and disease.* Eur Respir J, 2014. **43**(4): p. 983-92.
51. Amsellem, V., et al., *Telomere dysfunction causes sustained inflammation in chronic obstructive pulmonary disease.* Am J Respir Crit Care Med, 2011. **184**(12): p. 1358-66.
52. Tsuji, T., K. Aoshiba, and A. Nagai, *Alveolar cell senescence in patients with pulmonary emphysema.* Am J Respir Crit Care Med, 2006. **174**(8): p. 886-93.
53. Alder, J.K., et al., *Telomere length is a determinant of emphysema susceptibility.* Am J Respir Crit Care Med, 2011. **184**(8): p. 904-12.
54. O'Donovan, A., et al., *Cumulative inflammatory load is associated with short leukocyte telomere length in the Health, Aging and Body Composition Study.* PLoS One, 2011. **6**(5): p. e19687.
55. Gan, W.Q., et al., *Association between chronic obstructive pulmonary disease and systemic inflammation: a systematic review and a meta-analysis.* Thorax, 2004. **59**(7): p. 574-80.
56. von Zglinicki, T., *Oxidative stress shortens telomeres.* Trends Biochem Sci, 2002. **27**(7): p. 339-44.
57. Agusti, A. and D.D. Sin, *Biomarkers in COPD.* Clin Chest Med, 2014. **35**(1): p. 131-41.
58. Vogelmeier, C.F., et al., *Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease 2017 Report: GOLD Executive Summary.* Eur Respir J, 2017. **49**(3).
59. Daniali, L., et al., *Telomeres shorten at equivalent rates in somatic tissues of adults.* Nat Commun, 2013. **4**: p. 1597.
60. Robertson, J.D., et al., *Dynamics of telomere shortening in neutrophils and T lymphocytes during ageing and the relationship to skewed X chromosome inactivation patterns.* Br J Haematol, 2000. **109**(2): p. 272-9.
61. Kimura, M., et al., *Synchrony of telomere length among hematopoietic cells.* Exp Hematol, 2010. **38**(10): p. 854-9.

62. Qiu, W., et al., *Variable DNA methylation is associated with chronic obstructive pulmonary disease and lung function.* Am J Respir Crit Care Med, 2012. **185**(4): p. 373-81.
63. Wielscher, M., et al., *Diagnostic Performance of Plasma DNA Methylation Profiles in Lung Cancer, Pulmonary Fibrosis and COPD.* EBioMedicine, 2015. **2**(8): p. 929-36.
64. Lepeule, J., et al., *Gene promoter methylation is associated with lung function in the elderly: the Normative Aging Study.* Epigenetics, 2012. **7**(3): p. 261-9.
65. Bell, J.T., et al., *Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population.* PLoS Genet, 2012. **8**(4): p. e1002629.
66. Machin, M., et al., *Systematic review of lung function and COPD with peripheral blood DNA methylation in population based studies.* BMC Pulmonary Medicine, 2017. **17**(1): p. 54.
67. Marioni, R.E., et al., *The epigenetic clock is correlated with physical and cognitive fitness in the Lothian Birth Cohort 1936.* Int J Epidemiol, 2015. **44**(4): p. 1388-96.
68. Matthew Poulin, et al., *Pyrosequencing Methylation Analysis*, in *Cancer Epigenetics for Precision Medicine: Methods and Protocols, Methods in Molecular Biology*, Dumitrescu R and V. M., Editors. 2018, Springer Protocols. p. 283-296.
69. Lange, N.E., et al., *Alu and LINE-1 methylation and lung function in the normative ageing study.* BMJ Open, 2012. **2**(5).
70. Ayarpadikannan, S. and H.-S. Kim, *The impact of transposable elements in genome evolution and genetic instability and their implications in various diseases.* Genomics & informatics, 2014. **12**(3): p. 98-104.
71. Schulz, W.A., C. Steinhoff, and A.R. Florl, *Methylation of endogenous human retroelements in health and disease.* Curr Top Microbiol Immunol, 2006. **310**: p. 211-50.
72. Wu, M., et al., *Inherited somatic mosaicism caused by an intracisternal A particle insertion in the mouse tyrosinase gene.* Proc Natl Acad Sci U S A, 1997. **94**(3): p. 890-4.
73. Deininger, P., *Alu elements: know the SINEs.* Genome Biology, 2011. **12**(12): p. 236.
74. Qin, S., et al. *The Role of Transposable Elements in the Origin and Evolution of MicroRNAs in Human*. PloS one, 2015. **10**, e0131365 DOI: 10.1371/journal.pone.0131365.
75. Carnell, A.N. and J.I. Goodman, *The Long (LINEs) and the Short (SINEs) of It: Altered Methylation as a Precursor to Toxicity.* Toxicological Sciences, 2003. **75**(2): p. 229-235.
76. Kazazian, H.H. and J.V. Moran, *The impact of L1 retrotransposons on the human genome.* Nature Genetics, 1998. **19**(1): p. 19-24.
77. Ade, C., A.M. Roy-Engel, and P.L. Deininger, *Alu elements: an intrinsic source of human genome instability.* Curr Opin Virol, 2013. **3**(6): p. 639-45.
78. Sargurupremraj, M. and M. Wjst, *Transposable elements and their potential role in complex lung disorder.* Respiratory Research, 2013. **14**(1): p. 99.
79. Pray, L., *Functions and Utility of Alu Jumping Genes.* Nature Education, 2008. **1**(1): p. 93.
80. Sela, N., et al., *Comparative analysis of transposed element insertion within human and mouse genomes reveals Alu's unique role in shaping the human transcriptome.* Genome Biology, 2007. **8**(6): p. R127.

81. Zhang, X.H.F. and L.A. Chasin, *Comparison of multiple vertebrate genomes reveals the birth and evolution of human exons.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(36): p. 13427-13432.
82. Park, S.Y., et al., *Alu and LINE-1 Hypomethylation Is Associated with HER2 Enriched Subtype of Breast Cancer.* PLOS ONE, 2014. **9**(6): p. e100429.
83. Wong, J.Y., et al., *The association between global DNA methylation and telomere length in a longitudinal study of boilermakers.* Genet Epidemiol, 2014. **38**(3): p. 254-64.
84. Weisenberger, D.J., et al., *Analysis of repetitive element DNA methylation by MethyLight.* Nucleic Acids Res, 2005. **33**(21): p. 6823-36.
85. Kankava, K., et al., *LINE-1 METHYLATION IN BLOOD AND TISSUES OF PATIENTS WITH BREAST CANCER.* Georgian Med News, 2018(276): p. 107-112.
86. Roman-Gomez, J., et al., *Promoter hypomethylation of the LINE-1 retrotransposable elements activates sense/antisense transcription and marks the progression of chronic myeloid leukemia.* Oncogene, 2005. **24**(48): p. 7213-23.
87. Bollati, V., et al., *Changes in DNA methylation patterns in subjects exposed to low-dose benzene.* Cancer Res, 2007. **67**(3): p. 876-80.
88. Chen, Y.-W.R., et al., *C-reactive protein and N-terminal prohormone brain natriuretic peptide as biomarkers in acute exacerbations of COPD leading to hospitalizations.* PloS one, 2017. **12**(3): p. e0174063-e0174063.
89. Bourbeau, J., et al., *Canadian Cohort Obstructive Lung Disease (CanCOLD): Fulfilling the need for longitudinal observational studies in COPD.* Copd, 2014. **11**(2): p. 125-32.
90. CanCOLD. *CanCOLD Study*. 2013 [cited 2019 January 28]; Available from: http://cancold.ca/en/study.html.
91. England, R. and M. Pettersson, *Pyro Q-CpG™: quantitative analysis of methylation in multiple CpG sites by Pyrosequencing®.* Nature Methods, 2005. **2**: p. 798.
92. Royo, J.L., M. Hidalgo, and A. Ruiz, *Pyrosequencing protocol using a universal biotinylated primer for mutation detection and SNP genotyping.* Nature Protocols, 2007. **2**: p. 1734.
93. Kaminsky, Z.A., et al., *DNA methylation profiles in monozygotic and dizygotic twins.* Nat Genet, 2009. **41**(2): p. 240-5.
94. Bjornsson, H.T., et al., *Intra-individual change over time in DNA methylation with familial clustering.* Jama, 2008. **299**(24): p. 2877-83.
95. Norris, J., et al., *Identification of a new subclass of Alu DNA repeats which can function as estrogen receptor-dependent transcriptional enhancers.* J Biol Chem, 1995. **270**(39): p. 22777-82.
96. Asada, K., et al., *LINE-1 hypomethylation in a choline-deficiency-induced liver cancer in rats: dependence on feeding period.* J Biomed Biotechnol, 2006. **2006**(1): p. 17142.
97. Bell, J.T., et al., *DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines.* Genome biology, 2011. **12**(1): p. R10-R10.
98. Numata, S., et al., *DNA methylation signatures in development and aging of the human prefrontal cortex.* American journal of human genetics, 2012. **90**(2): p. 260-272.
99. Furniss, C.S., et al., *Line region hypomethylation is associated with lifestyle and differs by human papillomavirus status in head and neck squamous cell carcinomas.* Cancer epidemiology, biomarkers & prevention : a publication of the American Association for

Cancer Research, cosponsored by the American Society of Preventive Oncology, 2008. **17**(4): p. 966-971.
100. Breitling, L.P., et al., *Tobacco-smoking-related differential DNA methylation: 27K discovery and replication.* American journal of human genetics, 2011. **88**(4): p. 450-457.
101. Wan, E.S., et al., *Cigarette smoking behaviors and time since quitting are associated with differential DNA methylation across the human genome.* Human molecular genetics, 2012. **21**(13): p. 3073-3082.
102. Lee, K.W.K. and Z. Pausova, *Cigarette smoking and DNA methylation.* Frontiers in genetics, 2013. **4**: p. 132-132.
103. Beishline, K. and J. Azizkhan-Clifford, *Sp1 and the 'hallmarks of cancer'.* Febs j, 2015. **282**(2): p. 224-58.
104. Black, A.R., *Sp1 and kruppel-like factor family of transcription factors in cell growth regulation and cancer.* Journal of cellular physiology, 2001. **188**: p. 143-160.
105. Di, Y.P., J. Zhao, and R. Harper, *Cigarette smoke induces MUC5AC protein expression through the activation of Sp1.* The Journal of biological chemistry, 2012. **287**(33): p. 27948-27958.
106. Mercer, B.A., et al., *Identification of a cigarette smoke-responsive region in the distal MMP-1 promoter.* Am J Respir Cell Mol Biol, 2009. **40**(1): p. 4-12.
107. Holler, M., et al., *Sp1 transcription factor binds DNA and activates transcription even when the binding site is CpG methylated.* Genes Dev, 1988. **2**(9): p. 1127-35.
108. Lu, S.C. and J.M. Mato, *S-Adenosylmethionine in cell growth, apoptosis and liver cancer.* J Gastroenterol Hepatol, 2008. **23 Suppl 1**: p. S73-7.
109. Lu, S.C. and J.M. Mato, *Role of methionine adenosyltransferase and S-adenosylmethionine in alcohol-associated liver cancer.* Alcohol, 2005. **35**(3): p. 227-34.
110. Pogribny, I.P., et al., *Elevation in S-Adenosylhomocysteine and DNA Hypomethylation: Potential Epigenetic Mechanism for Homocysteine-Related Pathology.* The Journal of Nutrition, 2002. **132**(8): p. 2361S-2366S.
111. Nätt, D. and A. Thorsell, *Stress-induced transposon reactivation: a mediator or an estimator of allostatic load?* Environmental epigenetics, 2016. **2**(3): p. dvw015-dvw015.
112. Paulsen, M., S. Tierling, and J. Walter, *DNA methylation and the mammalian genome*, in *Epigenetics*, J. Tost, Editor. 2008, Caister Academic Press. p. 1-21.
113. Rusiecki, J.A., et al., *Global DNA hypomethylation is associated with high serum-persistent organic pollutants in Greenlandic Inuit.* Environ Health Perspect, 2008. **116**(11): p. 1547-52.
114. Baccarelli, A. and V. Bollati, *Epigenetics and environmental chemicals.* Curr Opin Pediatr, 2009. **21**(2): p. 243-51.
115. Baccarelli, A., et al., *Rapid DNA Methylation Changes after Exposure to Traffic Particles.* American Journal of Respiratory and Critical Care Medicine, 2009. **179**(7): p. 572-578.
116. Sood, A., et al., *Wood smoke exposure and gene promoter methylation are associated with increased risk for COPD in smokers.* American journal of respiratory and critical care medicine, 2010. **182**(9): p. 1098-1104.
117. Kirkham, P.A. and P.J. Barnes, *Oxidative Stress in COPD.* Chest, 2013. **144**(1): p. 266-273.

118. Valinluck, V., et al., *Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2).* Nucleic Acids Research, 2004. **32**(14): p. 4100-4108.
119. Vucic, E.A., et al., *DNA methylation is globally disrupted and associated with expression changes in chronic obstructive pulmonary disease small airways.* Am J Respir Cell Mol Biol, 2014. **50**(5): p. 912-22.
120. Barault, L., et al., *Leukocyte DNA as Surrogate for the Evaluation of Imprinted Loci Methylation in Mammary Tissue DNA.* PLOS ONE, 2013. **8**(2): p. e55896.
121. de Vries, M., et al., *From blood to lung tissue: effect of cigarette smoke on DNA methylation and lung function.* Respiratory Research, 2018. **19**(1): p. 212.

Appendices

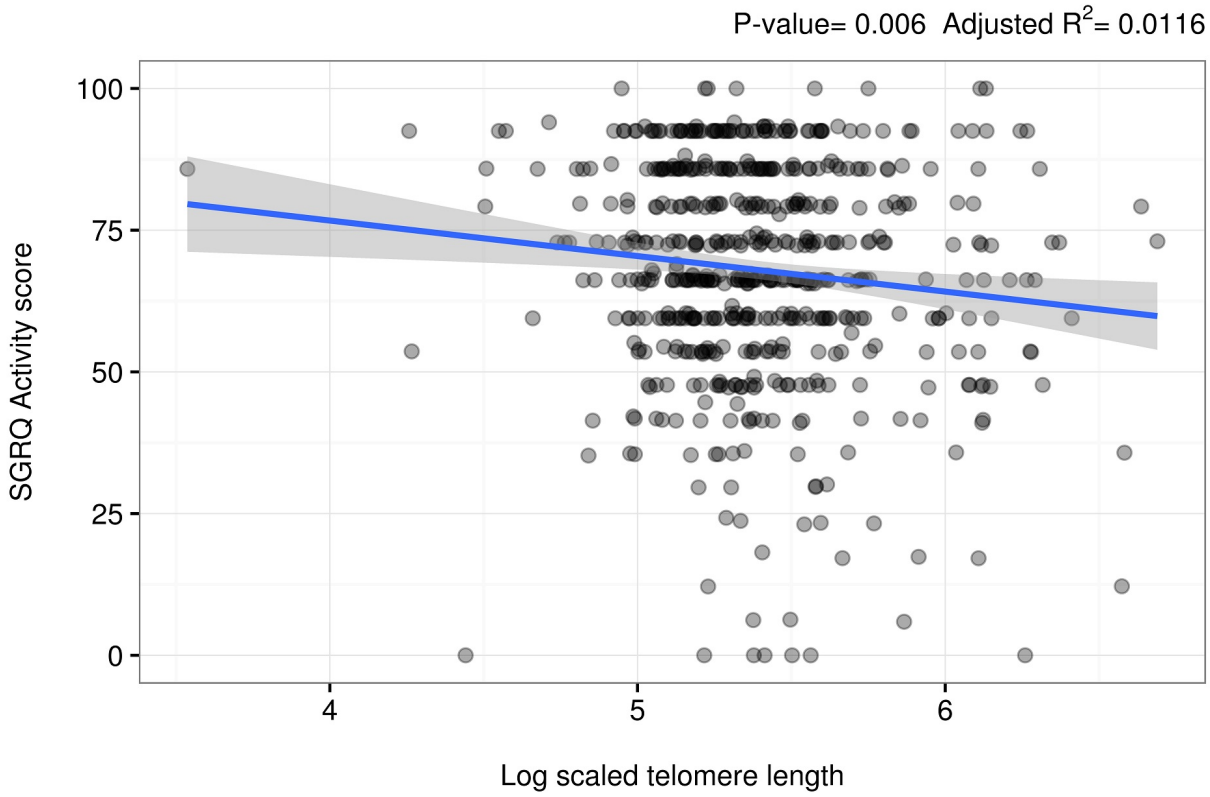
Appendix A **A comparison of the main demographic and clinical characteristics of MACRO participants who did and did not have telomeres measured from their peripheral circulation.**

Study participants were divided into two groups based on whether their telomere lengths were measured (aTL Measured) or not (aTL Non-Measured). Mean (SD) values are given for normally distributed variables, while dichotomous data are given as counts (% of total). Jonckheere-Terpstra trend test was used for continuous variables and the Cochran-Armitage test was used for dichotomous variables for *p*-trend.

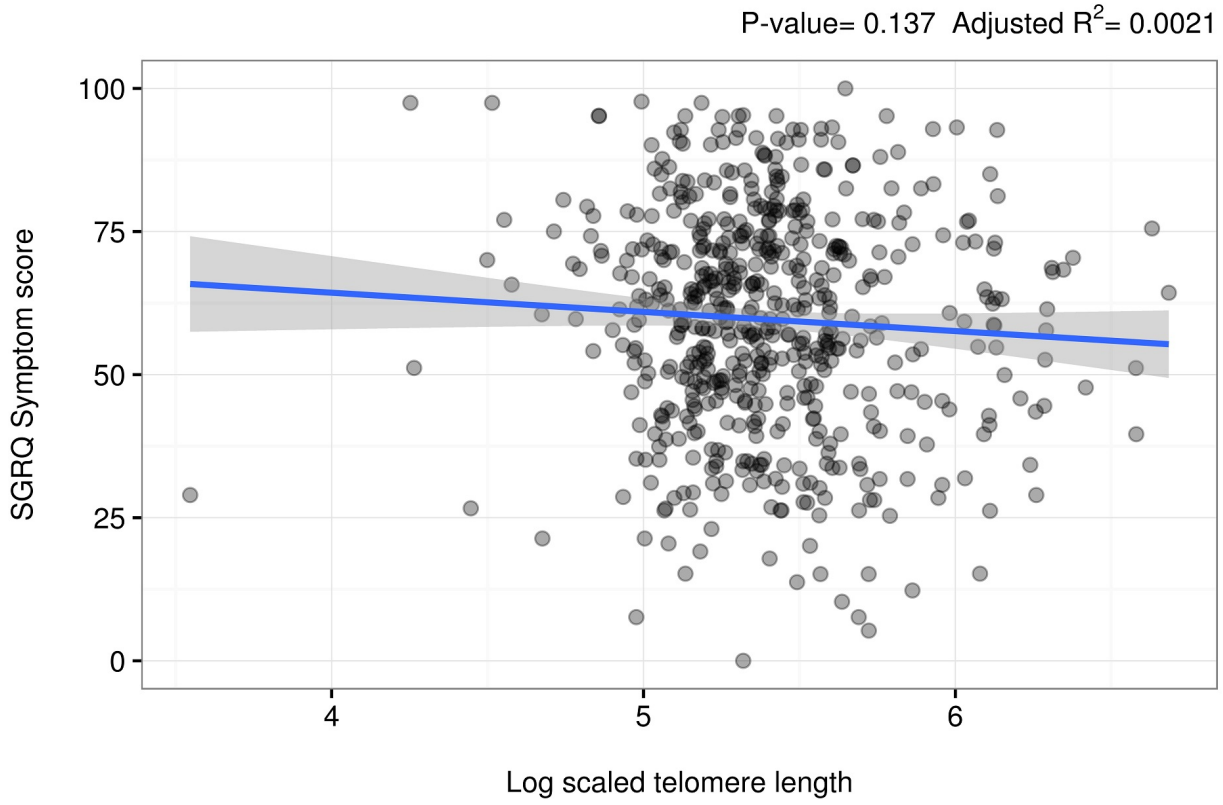
	Measured	**Non-Measured**	**P-value**
No. of participants	576	566	
Male sex	319 (55.4)	360 (63.6)	**0.005**
Azithromycin group	302 (52.4)	268 (47.3)	0.110
Age	66.0 (8.5)	64.4 (8.6)	**0.001**
White	488 (84.7)	440 (77.7)	**0.002**
Smoking (pack-years)	57.3 (31.1)	60.0 (33.4)	0.156
Current smoking status	107 (18.6)	145 (25.6)	**0.004**
Chronic bronchitis	255 (44.6)	280 (49.9)	0.072
Any cancer history	91 (15.8)	76 (13.4)	0.257
BMI (kg/m^2)	27.5 (6.0)	28.0 (6.6)	0.199
SGRQ, total score	47.9 (15.8)	53.3 (16.7)	**< 0.001**
Symptom score	59.6 (19.4)	63.7 (19.6)	**< 0.001**
Activity score	67.8 (19.6)	71.9 (19.1)	**< 0.001**
Impact score	32.9 (17.4)	39.3 (19.6)	**< 0.001**
Oxygen use	357 (62.0)	319 (56.4)	0.053
CRP (mg/L)	5.1 (4.0)	4.6 (3.7)	**0.040**
FEV_1/FVC	42.3 (12.4)	42.9 (13.1)	0.443
FEV_1, % predicted	39.6 (15.7)	39.5 (15.7)	0.933
FVC, % predicted	70.5 (18.7)	69.6 (17.7)	0.407

Appendix B **1, 2 and 3 The relationship of telomere length with SGRQ (B.1) Activity scores, (B.2) Symptom scores and (B.3) Impact scores.**

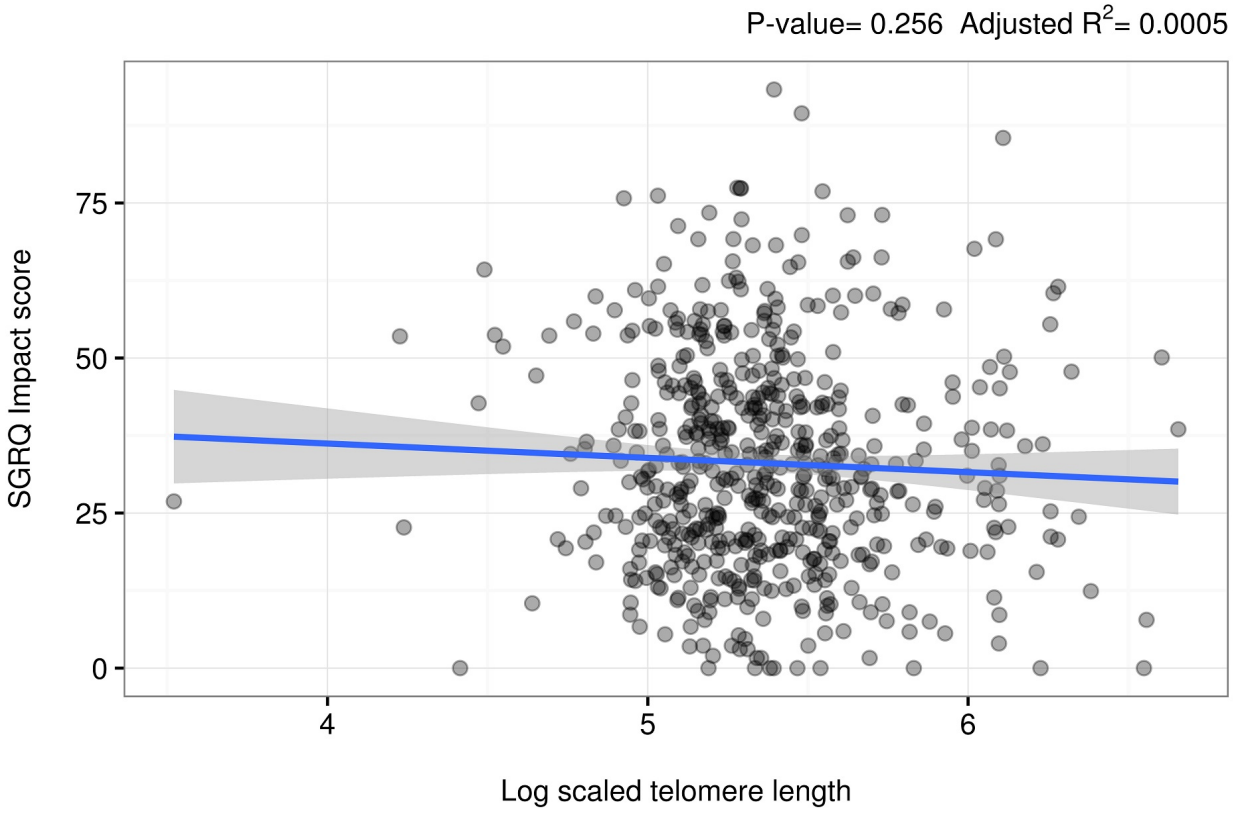
B.1 The relationship of telomere length with the SGRQ Activity scores



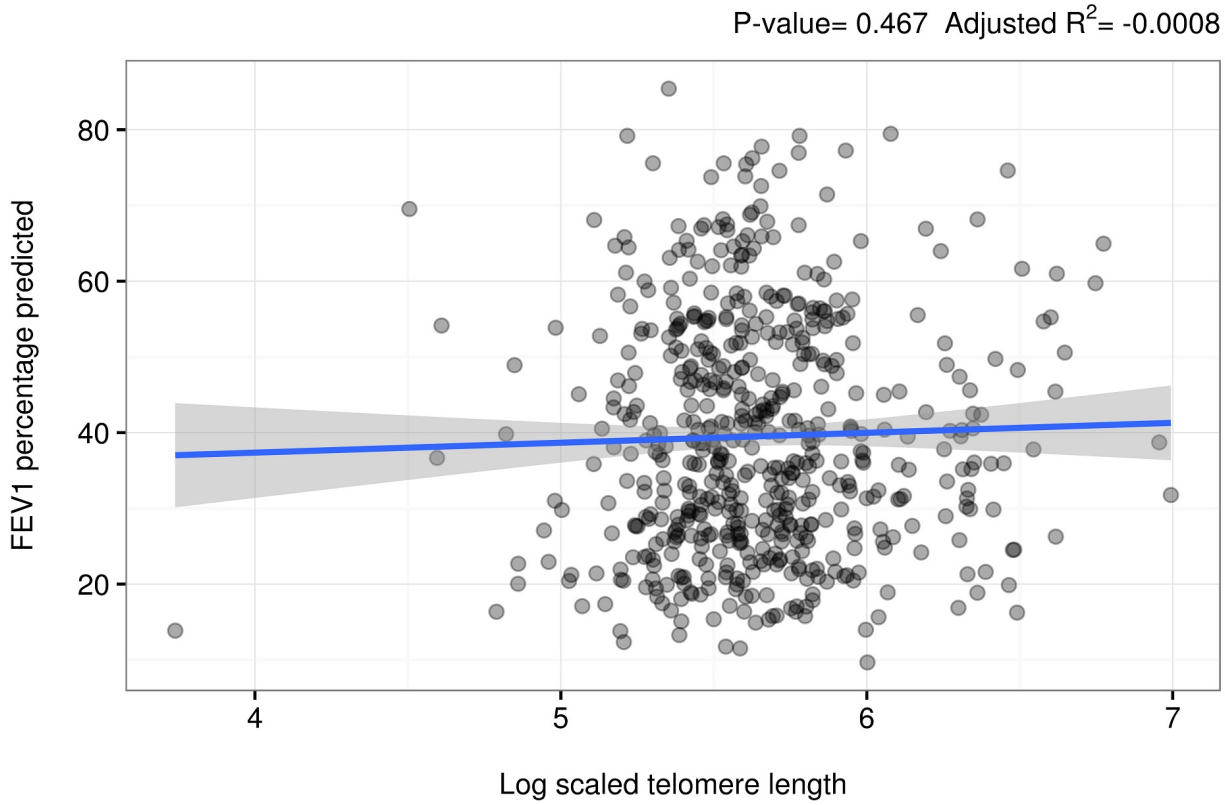
B.2 The relationship of telomere length with the SGRQ Symptom scores



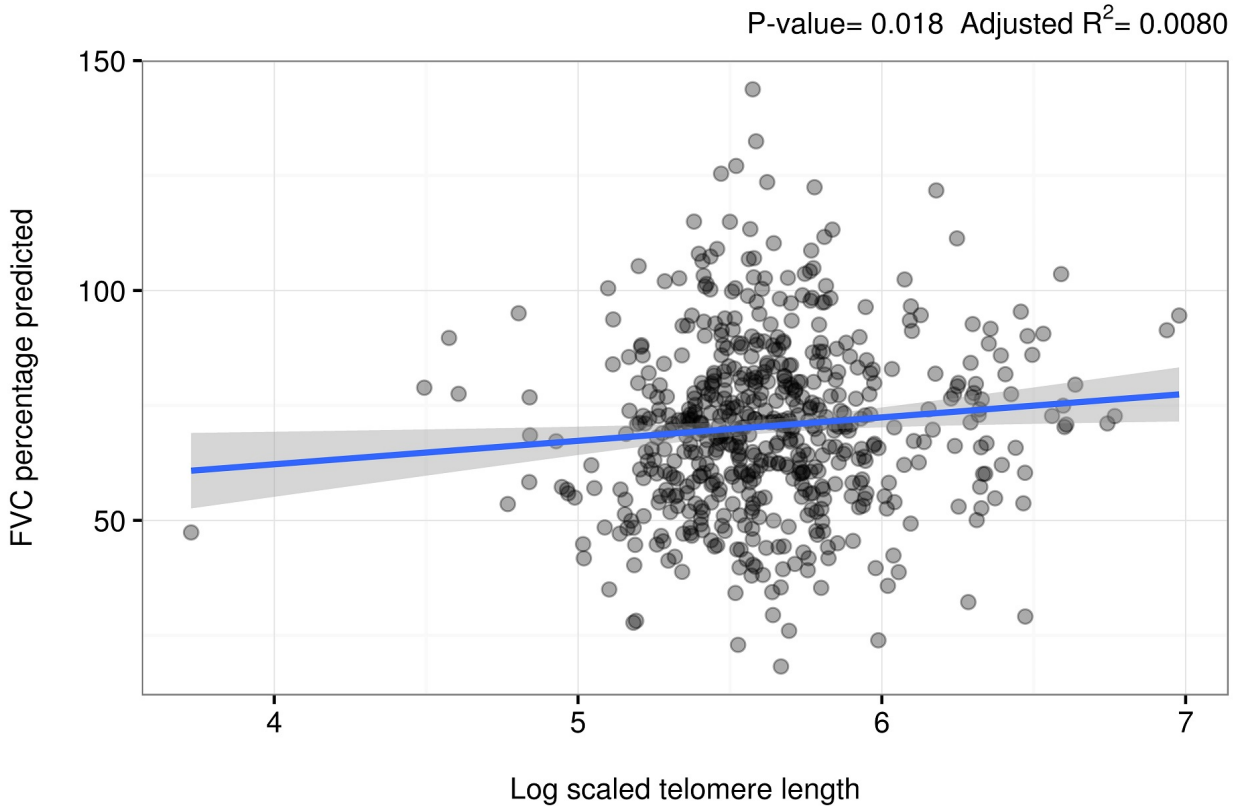
B.3 The relationship of telomere length with the SGRQ Impact scores



Appendix C **The relationship of telomere length with FEV_1 percentage predicted**



Appendix D **The relationship of telomere length with FVC percentage predicted**



Appendix E **A comparison of the main comorbidities of MACRO participants according to leukocyte telomere length in quartiles.**

All variables were in dichotomous data, and Cochran-Armitage test was used for *p*-trend.

	All	Telomere Quartiles (kbp/genome)				
		Q1 (<112.4)	Q2 (112.4-137.5)	Q3 (137.5-171.8)	Q4 (>171.8)	P-trend
No. of participants	576	144	144	144	144	–
Chronic bronchitis	255 (44.6)	66 (46.5)	67 (47.2)	65 (45.1)	57 (39.6)	0.221
Cancer history	91 (15.8)	23 (16.0)	27 (18.8)	23 (16.0)	18 (12.5)	0.332
High blood pressure	278 (48.3)	70 (48.6)	71 (49.3)	66 (45.8)	71 (49.3)	0.941
Previous heart attack	73 (12.7)	18 (12.5)	25 (17.4)	17 (11.8)	13 (9.0)	0.198
Cardio Vascular diseases	104 (18.1)	29 (20.1)	37 (25.7)	15 (10.4)	23 (16.0)	0.053
Diabetes Mellitus	86 (15.0)	23 (16.0)	22 (15.3)	22 (15.3)	19 (13.2)	0.531
Sleep apnea	126 (21.9)	32 (22.2)	29 (20.1)	29 (20.1)	36 (25.0)	0.589
Thyroid diseases	93 (16.1)	24 (16.7)	20 (13.9)	18 (12.5)	31 (21.5)	0.337
Clotting disorder	48 (8.3)	13 (9.0)	8 (5.6)	17 (11.8)	10 (6.9)	1.000
GERD	237 (41.1)	61 (42.4)	69 (48.0)	48 (33.3)	59 (41.0)	0.307
Depression	173 (30.0)	45 (31.3)	38 (26.4)	35 (24.3)	55 (38.2)	0.273

Appendix F **A comparison of the leukocyte cell proportions of MACRO participants according to telomere length in quartiles.**

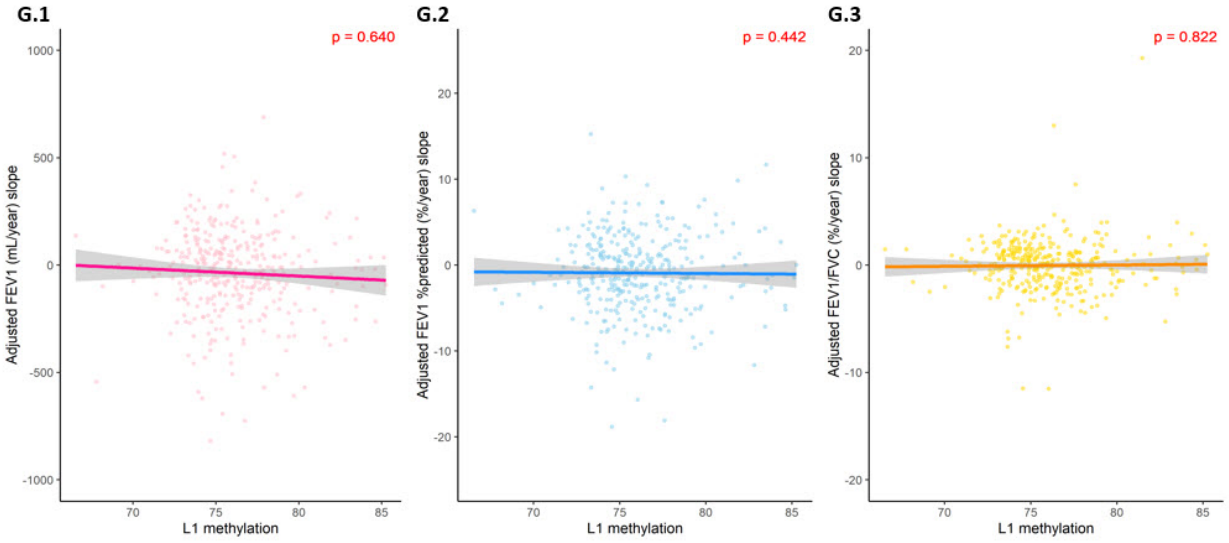
CBC analysis was performed on the samples of 276 participants. All variables were in continuous data, and Jonchkeere-Terpstra trend test was used for *p*-trend

	All	Telomere Quartiles (kbp/genome)				
		Q1 (<112.6)	Q2 (113.1-139.5)	Q3 (140.1-170.4)	Q4 (>170.8)	P-trend
No. of participants	276	69	69	69	69	–
CD8T	0.039 (0.035)	0.028 (0.032)	0.031 (0.034)	0.031 (0.031)	0.031 (0.038)	0.979
CD4T	0.112 (0.067)	0.096 (0.062)	0.107 (0.062)	0.123 (0.061)	0.121 (0.079)	**0.029**
NK	0.053 (0.038)	0.059 (0.037)	0.050 (0.036)	0.051 (0.041)	0.052 (0.038)	0.292
B cell	0.024 (0.030)	0.020 (0.021)	0.021 (0.028)	0.028 (0.031)	0.029 (0.037)	0.257
Mono	0.092 (0.031)	0.092 (0.032)	0.089 (0.030)	0.097 (0.033)	0.092 (0.027)	0.394
Gran	0.686 (0.115)	0.703 (0.101)	0.699 (0.108)	0.669 (0.115)	0.671 (0.133)	0.133

Appendix G **1, 2, 3 and 4 The relationship between percentage L1 methylation and the lung function decline in FEV_1 (mL/year), FEV_1% predicted (%/year) and FEV_1/FVC (%/year) using multivariable linear regression analyses (G.1, G.2 and G.3) and linear mixed effect analyses (G.4).**

A multivariable linear regression was analyzed with age, sex, race, BMI, smoking-status, pack-years and baseline lung function. The slope of each lung function measures was estimated using a linear regression of four lung function measures over 8 years. A linear mixed effect was analyzed with age, sex, race, BMI, smoking-status, pack-years and baseline lung function.

Linear Regression (Slope) Model



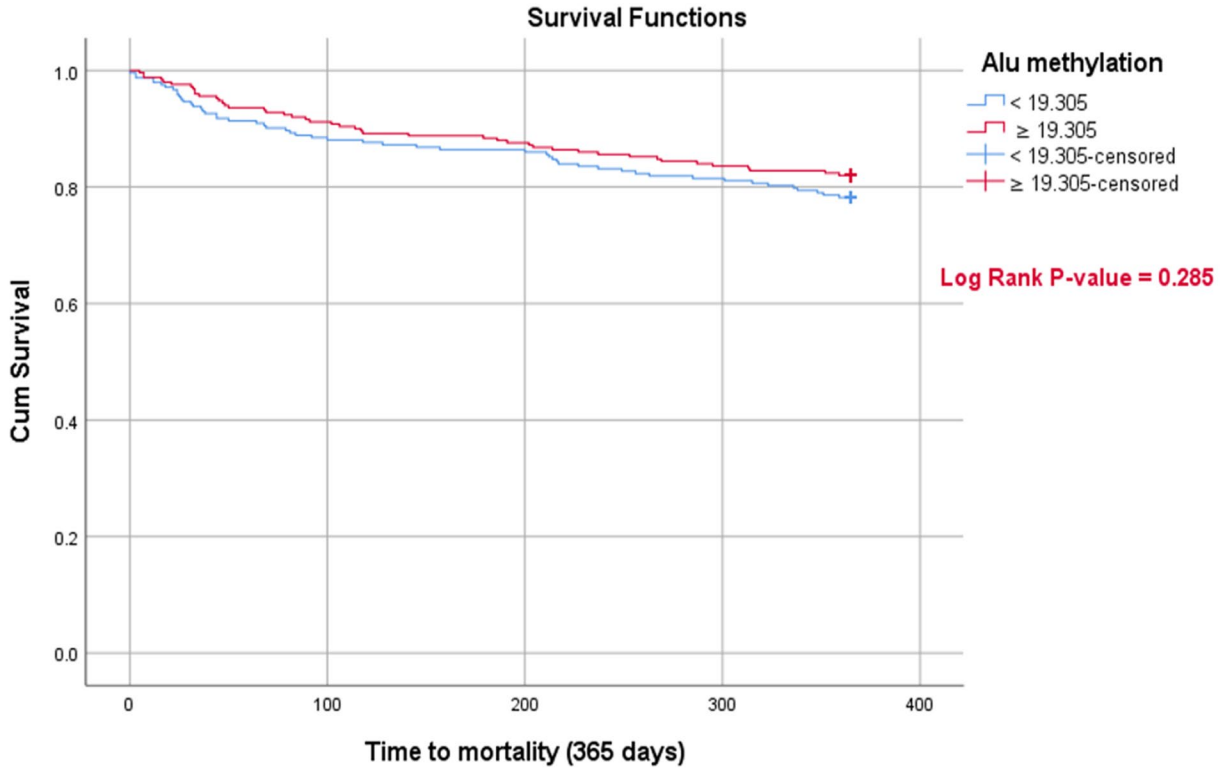
G.4 Linear Mixed Effect (LME) Model

Methylation %	Lung function decline	Estimated beta	Standard error	P-value
LINE-1	FEV_1 (mL/year)	-0.00141	0.00231	0.542
LINE-1	FEV_1 % predicted (%/year)	-0.0504	0.0728	0.488
LINE-1	FEV_1/FVC (%/year)	-0.010	0.0374	0.786

Appendix H **1 and 2 Kaplan-Meier (KM) survival curves for mortality according to (H.1) % Alu and (H.2) % L1 methylation in the COPD Rapid Transition program (RTP).**

Participants were divided into lower vs higher % methylation using the median value (19.31% for Alu and 73.45% for L1). P values were generated using the log-rank test.

H.1 A KM survival curve for mortality according to lower vs higher % Alu methylation in RTP



H.2 KM survival curve for mortality according to lower vs higher % L1 methylation in RTP