Supplement

Functional aging in fibrotic interstitial lung disease: The impact of frailty on adverse health outcomes

Sabina A Guler, MD [a,b,c]
Joanne M Kwan, MD [a,b]
Janice M Leung, MD [a,b]
Nasreen Khalil, MD [d]
Pearce G Wilcox, MD [a]
Christopher J Ryerson, MD [a,b]

[a] Department of Medicine, University of British Columbia, Vancouver, Canada
[b] Centre for Heart Lung Innovation, University of British Columbia, Vancouver, Canada
[c] Department of Pulmonary Medicine, University Hospital and University of Bern, Switzerland
[d] The Lung Centre, Vancouver General Hospital, Vancouver, Canada

METHODS

Measurement of blood leukocyte telomere length

Absolute telomere length as a marker of biological and cellular aging was measured in peripheral blood leukocytes from a sub-cohort of patients with fibrotic ILD. Studies comparing telomere attrition in different tissues have shown that telomere length correlates well between blood leukocytes and lung parenchyma, and that the rate of telomere attrition is consistent across organ systems within individual patients, suggesting that there is an intra-individual synchrony of telomere length in somatic tissues.[1,2]

A modified version of the Cawthon method for relative measurement of telomere length using quantitative real-time polymerase chain reaction and introduction of an oligomer standard was applied.[3,4] Genomic DNA was isolated from peripheral blood buffy coat using the QIAamp DNA blood mini kit (Qiagen, Toronto, Canada). Samples underwent only one freeze-thaw cycle before DNA extraction. Standard curves were generated from known quantities of synthesized oligomers of telomere DNA [TTAGGG repeated 14 times] and single copy gene (36B4) DNA [CAGCAAGTGGGAAGGTGTAATCCGTCTCCACAGACAAGGCCAGGACTCGTTTGTACCCG-TTGATGATAGAATGGG] (Sigma-Aldrich, St. Louis, MO). The standard curves allow the assessment of the sample telomere DNA length based on the ratio of telomere DNA length to 36B4 DNA length. DNA from a short telomere cell line (HEK293) and a long telomere cell line (K562, ATCC, Manassas, VA) were used as inter-experimental plate controls.[5] The ABI ViiA 7 Real Time PCR System (Applied Biosystems, Foster City, CA) was used to run samples in triplicate. The telomere lengths measured reflect an average length across the population of leukocyte cells included in the sample.

Expanded statistical methods

Data structure

The frailty index (FI) as the main predictor variable, as well as age and other demographic variables were collected at every visit; visits within time frames shorter than 6 months were excluded in order to avoid overlapping observation periods. Absolute telomere length (aTL) as the secondary predictor variable was collected only once in a subset of patients who consented to donation of blood for research purposes.

The primary mortality endpoint was time to the composite of death or lung transplantation based on previous observations of comparable disease severity in patients that are about to decease and patients undergoing lung transplantation. We performed a pre-specified sensitivity analysis with death and lung transplantation as competing risks (i.e., once a patient was transplanted, he or she was unable to contribute a subsequent mortality event to the analysis).

We divided the observation time in intervals defined by the time points of FI and covariate assessment in order to account for repeated FI measurement per patient and for time-dependent covariates (e.g. FI, age, pulmonary function).[6] Other outcomes were assessed within the 6-month time periods after every FI assessment: 1) rate of all-cause hospitalisations, 2) rate of respiratory related hospitalisations, 3) time to hospital discharge for the patients with hospitalisations, 4) occurrence of medication adverse reaction (MAR) for the patients treated with antifibrotic or immunosuppressive medications.

Data analysis

Descriptive statistics are reported as mean (standard deviation [SD]) or median (interquartile range [IQR]). Between group differences were analysed for statistical significance by chi-square or Fisher's exact test for categorical variables and by two-sample t-test or Wilcoxon rank sum test for continuous variables. Data were analyzed using R version 3.5.1 (R Foundation for Statistical Computing, Vienna, Austria).[7]

Mixed effects models

We applied (generalized) linear mixed models with random intercepts for every patient since multiple FI measurements from the same patient cannot be regarded as independent from each other. The unadjusted models included FI as a fixed effect. These models were adjusted for potential confounders of the effect of FI on adverse health outcomes. We considered confounders with either conceptual importance (age, sex) or a statistically relevant relationship to the outcome of interest ($p<0.1$). Consequently, the adjusted models included age, sex, forced vital capacity %-predicted (%FVC), diffusion capacity of the lung for carbon monoxide %-predicted (%DLCO), and a diagnosis of idiopathic pulmonary fibrosis (IPF) as fixed effects. We used the *R package lme4* for these analyses with functions lmer and glmer for linear and generalized linear mixed models, respectively.[8]

The same data analysis strategy was applied for different outcomes. The rates of all-cause and respiratory-related hospitalisations within 6 months were modeled by generalized linear mixed models with a Poisson distribution family and a log link function. The probability of MAR within 6 months from FI assessment was modeled using a generalized linear mixed model with a binomial family distribution and a logit link function (i.e. a logistic mixed model). SGRQ was modeled using linear mixed models fitted by restricted maximum likelihood (REML) with Satterthwaite's approximations for the degrees of freedom.

Model specification and standardized residuals were examined, including assessment for normality and homoscedasticity, over-dispersion, zero-inflation, and auto-correlation. The *R package DHARMa* was used for these analyses.[9]

Survival analysis

Time to death or lung transplantation and time to hospital discharge were each modeled with Cox proportional hazards regression models with intervals of time accounting for time-dependent covariates. Unadjusted models for FI and adjusted models including the above covariates were used to test the independent association of frailty with mortality. Model performance was measured using the Harrell's C-statistic. The independence between residuals and time (proportional hazards assumption) was tested using Schoenfeld residual tests.[10]

A prespecified sensitivity analysis with death (without lung transplantation) and lung transplantation as competing risks was performed by subdistribution hazard models according to Fine and Gray.[11] The *R packages survival* and *cmprsk* were used for these analyses.[12,13]

Causal mediation analysis

A causal mediation analysis (CMA) was performed with the goal to estimate average direct effects of chronological age on adverse health outcomes and indirect effects of chronological age mediated by

either biological age (aTL) or functional age (FI) (Figure S1). We performed a three-step procedure: First the *mediator models* were created by modelling the mediators separately (aTL and FI) as a function of the exposure (chronological age), second the *outcome models* for 2-year survival (logistic regression), rate of all-cause and respiratory-related hospitalisations (Poisson regression) were built, and third the two models were integrated into the *mediation model*, which estimates the strength of direct and indirect effects for an increase in chronological age by one year.[14]

To keep the models parsimonious, no additional covariates were included in the models. Assumptions for causal mediation analysis include the absence unmeasured confounding between the mediator and the outcome, which is typically untestable. We performed sensitivity analyses in order to estimate how strong a confounder would have to be to change the conclusion of the model: Unmeasured confounding between the mediator and the outcome leads to correlation between the residuals in the *mediator* and the *outcome regression models*. We tested the potential strength of the correlation between model residuals that would cause the estimated indirect effect to change direction.[14] CMA was performed with the *R mediation* package.[15]

Table S1. Baseline characteristics of patients with CTD-ILD, unclassifiable ILD, IPF and hypersensitivity pneumonitis.

	CTD-ILD (n=227)	Unclassifiable (n=127)	IPF (n=100)	HP (n=47)
DEMOGRAPHICS				
Sex, men	53 (23%)	72 (57%)	75 (75%)	18 (38%)
Age, years	60.1 (12.7)	68.9 (10.6)	70.9 (8.0)	63.9 (10.0)
Body mass index, kg/m^2	26.3 (5.6)	29.5 (5.3)	27.9 (4.9)	30.8 (5.8)
Ever-smoker	101 (43%)	65 (51%)	62 (62%)	25 (53%)
Smoked pack-years*	10 (2.25-26.3)	19.5 (9.5-36)	72 (58.8-82.3)	18.1 (7.8-38.2)
ILD SEVERITY				
FVC, %-predicted	76.1 (21.5)	73.8 (21.4)	72.0 (17.3)	70.4 (20.6)
FEV_1, %-predicted	76.5 (22.1)	76.9 (22.1)	76.0 (17.2)	74.1 (21.2)
DLCO, %-predicted	54.0 (19.2)	56.6 (19.5)	45.1 (13.7)	53.1 (15.9)

Data are presented as mean (standard deviation), median (interquartile range), or frequency (percentage).
*in ever-smokers

Abbreviations: CTD, connective tissue disease; DLCO%, diffusion capacity of the lung for carbon monoxide %-predicted; FEV_1, forced vital capacity in one second; FVC%, forced vital capacity %-predicted; HP, hypersensitivity pneumonitis; ILD, interstitial lung disease; IPF, idiopathic pulmonary fibrosis

Table S2. Baseline characteristics of patients with IPF and non-IPF ILDs in the full cohort and the sub-cohort with blood samples for absolute telomere length measurement available.

	Full cohort		Telomere cohort	
	IPF (n=100)	Non-IPF ILD (n=440)	IPF (n=41)	Non-IPF ILD (n=148)
Age	69.7 (8.5)	62.9 (12.6)	68.1 (8.2)	61.6 (11.9)
Sex, male	75 (74%)	157 (36%)	30 (73%)	57 (39%)
Body mass index, kg/m^2	27.9 (4.9)	27.8 (5.8)	29.0 (4.7)	27.3 (5.4)
Ever smoker	62 (62%)	210 (48%)	33 (80%)	85 (57%)
Smoked pack-years	23.5 (13-37)	15 (4-31)	20.0 (12.8-35)	12 (4-25.3)
ILD SEVERITY				
FVC, %-predicted	72.0 (17.3)	74.7 (21.5)	75.0 (17.0)	76.0 (20.3)
FEV$_1$, %-predicted	76.0 (17.2)	76.1 (22.1)	79.6 (17.5)	78.0 (20.3)
DLCO, %-predicted	45.1 (13.7)	54.6 (19.2)	45.3 (13.9)	54.4 (17.4)
QUALITY OF LIFE				
SGRQ, total	46.9 (23.0)	42.9 (21.8)	50.9 (21.0)	46.4 (18.6)
SGRQ, activity	63.3 (28.1)	57.2 (25.8)	57.8 (24.4)	55.3 (25.9)
SGRQ, symptom	48.5 (24.7)	48.5 (23.4)	54.5 (23.9)	46.1 (24.0)
SGRQ, impact	36.7 (23.1)	32.6 (23.1)	31.3 (17.3)	30.1 (23.1)
FRAILTY				
Frailty Index	0.167 (0.092-0.288)	0.214 (0.095-0.333)	0.146 (0.071-0.262)	0.181 (0.043-0.348)
Co-FI	0.111 (0.097-0.211)	0.158 (0.055-0.214)	0.105 (0.105-0.158)	0.105 (0.053-0.211)
I&SC-FI	0.174 (0.077-0.391)	0.261 (0.087-0.435)	0.174 (0.043-0.348)	0.217 (0.045-0.400)
Frail (FI >0.21)	39 (39%)	233 (53%)	15 (37%)	70 (47%)
Prefrail (FI 0.1-0.21)	28 (28%)	91 (21%)	10 (24%)	25 (17%)

Data are presented as mean (standard deviation), median (interquartile range), or frequency (percentage).
*in ever-smokers

Abbreviations: DLCO%, diffusion capacity of the lung for carbon monoxide %-predicted; FEV$_1$, forced vital capacity in one second; FVC%, forced vital capacity %-predicted; ILD, interstitial lung disease; SGRQ, St. George's Respiratory Questionnaire.

Figure S1. Directed acyclic graph

The directed acyclic graph illustrates the hypothesized mediation of the effect of chronological age on adverse health outcomes by either functional age or biological age.

Abbreviations: FI, frailty index; aTL, absolute telomere length

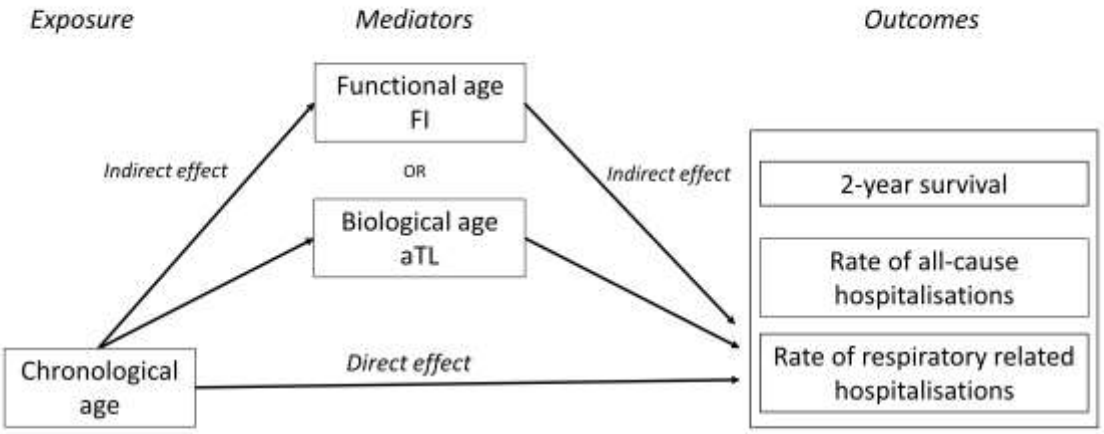
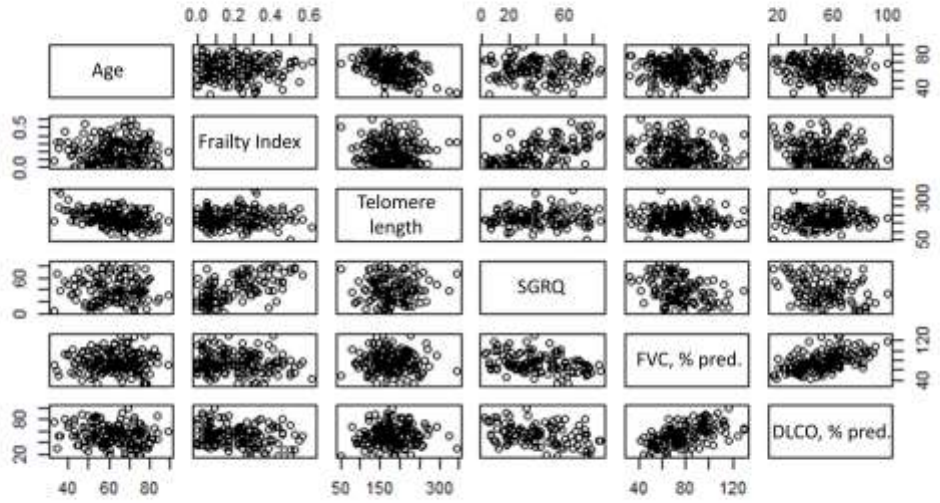


Figure S2. Pairwise scatterplots.

Scatterplots for age, Frailty Index, absolute telomere length, quality of life, and pulmonary function tests.

Abbreviations: DLCO%, diffusion capacity of the lung for carbon monoxide %-predicted; FVC%, forced vital capacity %-predicted; SGRQ, St. George's Respiratory Questionnaire.



REFERENCES

1. Daniali L, Benetos A, Susser E, et al. Telomeres shorten at equivalent rates in somatic tissues of adults. *Nature communications.* 2013;4:1597.
2. Friedrich U, Griese E, Schwab M, Fritz P, Thon K, Klotz U. Telomere length in different tissues of elderly patients. *Mechanisms of ageing and development.* 2000;119(3):89-99.
3. Cawthon RM. Telomere measurement by quantitative PCR. *Nucleic Acids Res.* 2002;30(10):e47.
4. O'Callaghan NJ, Fenech M. A quantitative PCR method for measuring absolute telomere length. *Biological procedures online.* 2011;13:3.
5. Fehrer C, Voglauer R, Wieser M, et al. Techniques in gerontology: cell lines as standards for telomere length and telomerase activity assessment. *Experimental gerontology.* 2006;41(6):648-651.
6. Therneau T CC, Atkinson C. Using Time Dependent Covariates and Time Dependent Coefficients in the Cox Model. 2017.
7. Team RC. R: A language and environment for statistical computing. . *R Foundation for Statistical Computing, Vienna, Austria.* 2016.
8. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. *2015.* 2015;67(1):48.
9. (2018). FH. DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models. R package version 0.2.0. *https://CRAN.R-project.org/package=DHARMa.* 2018.
10. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika.* 1994;81(3):515-526.
11. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *Journal of the American Statistical Association.* 1999;94(446):496-509.
12. T T. A Package for Survival Analysis in S. Version 2.38. 2015.
13. Gray B. cmprsk: Subdistribution Analysis of Competing Risks. R package version 2.2-8. . 2019.
14. Imai K, Keele L, Yamamoto T. Identification, Inference and Sensitivity Analysis for Causal Mediation Effects. *Statist. Sci.* 2010;25(1):51-71.
15. Dustin Tingley TY, Kentaro Hirose, Luke Keele, Kosuke Imai. mediation: R Package for Causal Mediation Analysis. *Journal of Statistical Software.* 2014;59(5):1-38.