Controlling for baseline telomere length biases estimates of the effect of smoking on leukocyte telomere attrition in longitudinal studies

Melissa Bateson and Daniel Nettle

Centre for Behaviour and Evolution and Institute of Neuroscience, Newcastle University, UK.

Email: melissa.bateson@ncl.ac.uk

In analyses of longitudinal leukocyte telomere length (LTL) data it is common practice to adjust for baseline LTL when modelling the factors that potentially affect LTL attrition. However, the apparent dependence of LTL attrition on baseline LTL is largely brought about by regression to the mean as a result of measurement error. We used simulated LTL data to explore whether adjusting for baseline LTL results in biased estimates of the true effect of smoking on attrition. We show that when baseline LTL is shorter in smokers than non-smokers and LTL measurement error is non-zero, adjusting for baseline LTL results in overestimating the true effect of smoking on telomere attrition. The size of this latter bias increases with increasing LTL measurement error. Since it is a robust finding that smokers have shorter LTL than non-smokers and LTL measurement error is substantial, we conclude that the type 1 error rate for reports of effects of smoking on telomere attrition is likely to be above 5%. We therefore recommend that models of LTL attrition should not be adjusted for baseline LTL. Although we have couched our analysis in terms of the effects of smoking, our findings are likely to have general relevance to other lifestyle factors and exposures studied in relation to telomere attrition.

Introduction

Leukocyte telomere length (LTL)—the length of the repeated TTAGGG sequence at the end of leukocyte chromosomes—is emerging as a widely studied biomarker of human health. Many recent studies have attempted to test whether lifestyle factors such as diet, exercise and smoking affect the rate of LTL attrition (e.g. Bendix et al. 2014; Ehrlenbach et al. 2009; Huzen et al. 2014; Müezzinler et al. 2015; Révész et al. 2016; Weischer et al. 2014). LTL attrition rates are estimated from longitudinal datasets in which LTL is measured twice, at baseline (LTL_b) and follow up (LTL_{fu}). Attrition rate is calculated via the following formula:

LTL attrition (bp/year) = (LTL_b – LTL_{fu})/follow-up years (Equation 1)

In studies of LTL attrition, it is typical to find a strong positive correlation between LTL attrition rate and baseline LTL: individuals with relatively longer telomeres at baseline have faster attrition than individuals with shorter telomeres (e.g. Aviv et al. 2009). This phenomenon is largely explained as a statistical artefact arising from measurement error and regression to the mean (Verhulst et al. 2013). Given the strong statistical dependence of LTL attrition on baseline LTL it is common practice to adjust for baseline LTL in multiple regression models of LTL attrition. For example, we have found nine studies that report the effect of smoking on LTL attrition and all of these adjust for baseline LTL in their models (Bendix et al. 2014; Ehrlenbach et al. 2009; Huzen et al. 2014; Müezzinler et al. 2015; Révész et al. 2016; Weischer et al. 2014; Aviv et al. 2009; Farzaneh-Far et al. 2010; Toupance et al. 2017).

It has long been recognised that statistical adjustment for a third variable can alter the size of the effect between two variables of interest. Whether adjustment for a variable removes or introduces bias depends on the causal relations among the variables being studied (Greenland 2003). In the situation described above, variation in measured baseline LTL is potentially caused by both measurement error and smoking, since there is robust evidence that smokers have shorter LTL than non-smokers in cross-sectional analyses (Astuti et al. 2017). Measured baseline LTL is therefore what is termed a 'collider' variable (Figure 1). Controlling for collider variables can, in some circumstances, exaggerate effect sizes between variables of interest, a phenomenon known as collider bias (Greenland 2003). Since most of the longitudinal studies of the effects of smoking on telomere attrition also report an effect of smoking on baseline LTL, this raises the question of whether the results reported for the effects of smoking on LTL attrition are biased. Furthermore, since measurement error affects baseline LTL, and baseline LTL affects apparent telomere attrition (via regression to the mean), how does measurement error affect the size of this bias?

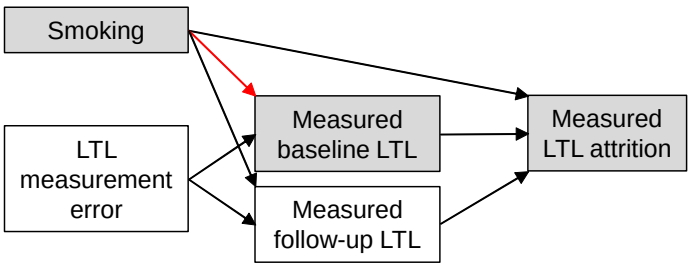

Figure 1. Assumed causal relations between smoking and LTL. The variables in the shaded boxes are those typically included in published multiple regression models of LTL attrition. The red arrow indicates the relationship that turns baseline LTL into a collider variable and potentially makes adjusting for this variable problematic.

Methods

To address the above questions, we simulated and then modelled LTL data in which we knew the true effects of smoking and measurement error and hence could calculate any biases in the parameter estimates from models of the data adjusting for baseline LTL. We simulated four different scenarios to describe the true differences in LTL data between smokers and non-smokers: (A) No difference in attrition and no difference in baseline LTL; (B) A true difference in attrition, but no difference in baseline LTL; (C) No difference in attrition, but a true difference in baseline LTL; and (D) A true difference in attrition and a true difference in baseline LTL. Since scenarios C and D both assumed a true difference in baseline LTL between smokers and non-smokers (shorter LTL in smokers), we predicted that estimates of the effect of smoking on telomere attrition would be biased in these scenarios if baseline LTL was included in the models.

The parameter values used in each scenario were taken from Aviv et al. (2009), who found a significant baseline difference in LTL between smokers and non-smokers of 141 bp and a non-significant difference in attrition between smokers and non-smokers of 2 bp/year (Table 1). (Note that the standard deviations of baseline LTL and annual attrition in Table 1 are likely to be overestimates of the standard deviations of the true variables, since both true variation and measurement error contribute to the measured values. However, in the absence of error-free measurements we used these published standard deviations as the best estimates available.)

Table 1: Parameter values used in the simulations.

		Scenario A	Scenario B	Scenario C	Scenario D
Non-smokers	Baseline LTL (bp; mean±sd)	7451±777	7451±777	7481±777	7481±777
	Attrition (bp/year; mean±sd)	40.7±46	40±46	40.7±46	40±46
Smokers	Baseline LTL (bp; mean±sd)	7451±777	7451±777	7392±777	7481±777
	Attrition (bp/year; mean±sd)	40.7±46	42±46	40.7±46	42±46

The simulation of LTL values was based on one previously described by Bateson and Nettle (2016) and was implemented in the statistical computing language R. The script for the simulation is available as supplementary information: "R_script_Bias_Simuation.R"; two R functions required by this script are also included as supplementary information. In each replicate simulation, true baseline LTLs were generated for 2000 individuals (1000 non-smokers and 1000 smokers) by drawing independent random samples from normal distributions for baseline LTL with means and standard deviations for non-smokers and smokers as given in Table 1. Each individual was then assigned an annual attrition by again drawing an independent random sample from normal distributions for attrition with means and standard deviations given in Table 1. This rate of attrition was applied for 10 years starting with the true baseline LTL to yield a true follow-up LTL for each individual. (Note that we assumed that each individual experienced a consistent rate of attrition over the follow-up interval, equivalent to setting r = 1 in Bateson and Nettle's original simulation). Measurement error was introduced into both baseline and follow-up LTL values by assuming that measured LTL was an independent random sample from a normal distribution with the mean equal to the true LTL and the standard deviation equal to the true TL * CV/100. Measured LTL attrition for each individual was calculated according to Equation 1 using the 'measured' values for LTL_b and LTL_{fu}. We assumed values of CV of 0, 2, 4, 8, 10 and 12% and generated 1000 replicate data sets for each value of CV in each of the four scenarios (A, B, C and D).

We modelled the data from each replicate with the three different general linear models:

Attrition ~ Smoking (Model 1)

Attrition ~ Baseline LTL + Smoking (Model 2)

RTM-corrected attrition ~ Smoking (Model 3)

where attrition and baseline LTL were continuous variables and smoking was categorical (smokers versus non-smokers). Thus model 2 includes adjustment for baseline LTL, as is typical in the literature. In model 3 we used the equation suggested by Verhulst et al. (2013) to correct the raw measured attrition values for regression to the mean caused by measurement error. For completeness, we initially considered a fourth alternative for how the data could be modelled that has occasionally been used in the literature:

Follow-up LTL ~ Baseline LTL + Smoking (Model 4)

However, since model 4 produces identical results to model 2 we did not consider it further.

For each model we recorded the β coefficient for the estimated effect of smoking and whether the parameter estimate for smoking was significantly different from 0 at $p < 0.05$. Data on the results of significance tests were used to calculate the probability of a type 1 error being committed in scenarios A and C (where there is no true difference between smokers and non-smokers in attrition) and the power to reject the null hypothesis of no effect of smoking on attrition in scenarios B and D (where smokers had a true attrition rate 2 bp/year higher than non-smokers).

Results

Summarised output data from the simulation are available as "Supplementary_dataset.CSV". These data were used to create Figures 2 and 3.

Figure 2 shows the estimated effects of smoking on LTL attrition derived from models 1, 2 and 3 in the four different scenarios that we simulated. Scenarios A and C simulate datasets in which the null hypothesis of no effect of smoking on telomere attrition is true. In scenario A, in which there is also no effect of smoking on baseline LTL, models 1, 2 and 3 all correctly estimate the effect of smoking on telomere attrition as zero. However, in scenario C, in which there is a true effect of smoking on baseline LTL, while model 1 correctly estimates the effect of smoking on attrition as zero, models 2 and 3 overestimate the effect of smoking on attrition at non-zero values of measurement error and this overestimation increases as LTL measurement error increases. Model 3 produces biased estimates even when the measurement error is zero and the biases are greater than those from model 2 up to CVs of 12%.

Scenarios B and D simulate datasets in which there is a true effect of smoking on telomere attrition of an additional 2 bp/year in smokers. In scenario B, in which there is no effect of smoking on baseline LTL, models 1, 2 and 3 correctly estimate the effect of smoking on attrition as 2 bp/year. However, in scenario D, in which there is a true effect of smoking on baseline LTL, while model 1 correctly estimates the effect of smoking on attrition, models 2 and 3 overestimate the effect at non-zero values of measurement error and this overestimation increases as measurement error increases. As in scenario C, model 3 produces biased estimates even when the measurement error is zero and the biases are greater than those from model 2 up to CVs of 12%.

In summary, it appears that when there is a true effect of smoking on baseline LTL (scenarios C and D), adjusting for baseline LTL, either via including baseline LTL as a covariate (model 2) or via using a measure of attrition that is corrected for regression to the mean (model 3), results in estimating an exaggerated effect of smoking on attrition when measurement error is non-zero.

Figure 3 shows the probability of type 1 errors in scenarios A and C (where there is no true effect of smoking on attrition) and the power to reject the null hypothesis in scenarios B and D (where smokers have an additional 2 bp/year attrition compared to non-smokers). In scenario A, the probability of type 1 errors with all three models is around 0.05, as would be expected. However, in scenario C (where there is a difference in baseline LTL between smokers and non-smokers) the type 1 error rates with models 2 and 3 reflect the exaggerated effect sizes seen in Figure 2C, rising as measurement error increases.

In scenario B the power to detect the small true difference in attrition between smokers and non-smokers is low (around 0.15 at a CV of 0%) for all models and decreases as measurement error increases, as would be expected. However, in scenario D (where there is a difference in baseline LTL between smokers and non-smokers) the relationship between power and measurement error is different for each of the models: with models 1 and 3 power decreases as measurement error increases and the power is overall higher in model 3 than in model 2 (around 0.3 compared to 0.15 at a CV of 12%), but with model 2 power increases as measurement error increases and the power is intermediate between that of models 1 and 3 for much of the range of values of CV explored. Note that the rather low absolute levels of power in Figure 3 reflect the very small true difference in attrition between smokers and non-smokers (only 2 bp/year) assumed in the simulations.

Discussion

Our results show that as long as there is no true difference in baseline LTL between smokers and non-smokers (scenarios A and B), then all of the modelling approaches that we have considered accurately estimate the effect of smoking on LTL attrition and do this with equivalent power. However, if there is a true difference between smokers and non-smokers in baseline LTL (scenarios C

and D) and LTL measurement error is non-zero, then adjusting for baseline LTL in multiple regression models biases estimates of the effect of smoking on attrition. Estimates of the effect of smoking on attrition are overestimated and the size of this overestimation increases as LTL measurement error increases for realistic values of measurement error. This overestimation of the effect sizes of smoking on attrition has two inter-related consequences. First, in scenarios in which there is no true effect of smoking on attrition (A and C), it translates into a type 1 error rate of above the usually-accepted 5% level. Second, in scenarios in which there is a true effect of smoking on attrition (B and D), it increases the power of the model to reject the null hypothesis.

Figure 4 provides an intuitive explanation for why the above overestimation occurs in models with adjustment for baseline LTL (such as model 2). The three panels depict cartoon data from a scenario in which there is no true effect of smoking on attrition (mean attrition is the same in smokers and non-smokers), but there is a baseline LTL difference, with smokers having shorter LTL than non-smokers (i.e. scenario C in this manuscript). Adding progressively higher measurement error introduces a positive relationship between baseline LTL and apparent attrition as indicated by the overall regression line that rotates about the mean values of baseline LTL and attrition, increasing in slope as the CV increases (panels A-C). The effect of this rotation is to create mean positive residuals of the data from the regression line for smokers and mean negative residuals for non-smokers (indicated by the vertical arrows in Figure 4). This bias occurs because the smokers have a mean baseline LTL that is lower than that of non-smokers and would not occur if there was no true difference in baseline LTL (as is the case in scenarios A and B). The result is that the estimated effect of smoking on attrition grows as measurement error increases.

Longitudinal studies of LTL often report significant effects of smoking on baseline LTL (Aviv et al. 2009; Huzen et al. 2014; Müezzinler et al. 2015; Révész et al. 2016; Weischer et al. 2014). Measurement error is also a substantial problem in telomere biology (Aviv et al. 2011). Therefore, estimates of the effects of smoking on LTL attrition are likely to be exaggerated in the literature given the common practice of adjusting for baseline LTL in regression models. Type 1 error rates for effects of smoking on attrition are consequently likely to be above 5% in the published literature and reported significant effects should be interpreted with this in mind (Bendix et al. 2014; Huzen et al. 2014). Furthermore, using the method recommend by Verhulst et al. (2013) for correcting for regression to the mean does not solve this problem, and actually makes it worse for realistic values of measurement error.

We have couched our analysis in terms of the effects of smoking on LTL. However, our general findings are relevant to estimating the effect on telomere attrition of any lifestyle choice or exposure that is associated with a true difference in LTL at the time of baseline measurement. On the basis of our analyses we recommend that models of LTL attrition should not be adjusted for baseline LTL.

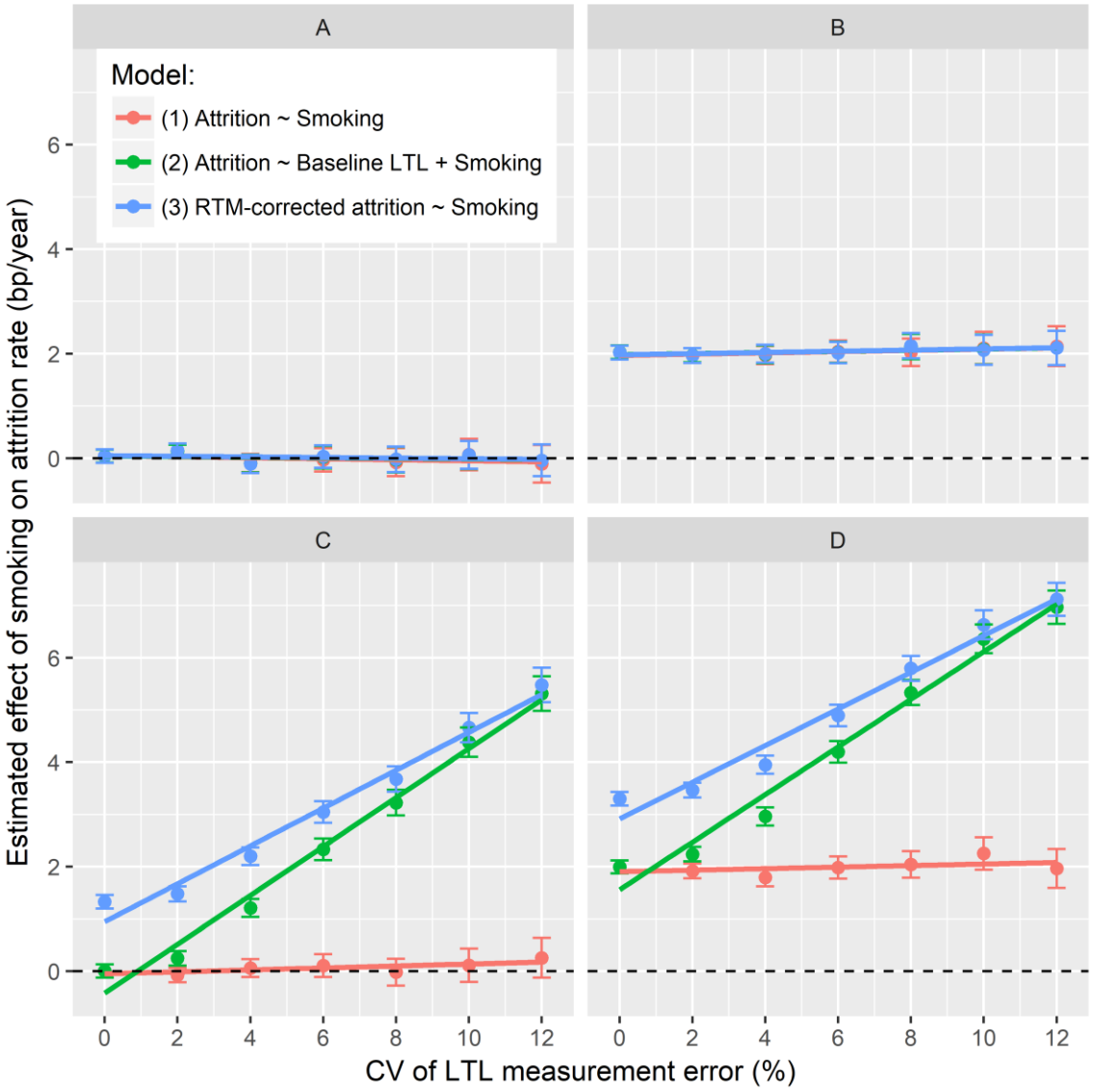

Figure 2. The estimated effects of smoking on LTL attrition rate obtained from three general linear models (models 1, 2 and 3) fitted to data simulated given four sets of assumptions regarding the true effects of smoking (scenarios A-D). The dashed lines indicate no effect of smoking on attrition. Data points are the mean ± 95% confidence intervals obtained from modelling the data from 1000 replicate simulations and lines are simple linear fits. The four scenarios are as follows: (A) No difference in attrition and no difference in baseline LTL; (B) A true difference in attrition, but no difference in baseline LTL; (C) No difference in attrition, but a true difference in baseline LTL; and (D) A true difference in attrition and a true difference in baseline LTL. The true effect of smoking on LTL attrition rate in scenarios B and D was an additional 2 bp/year in smokers. The true effect of smoking on baseline LTL in scenarios C and D was that smoker's baseline LTL were 141 bp shorter. Note that in order to express the results from model 3 in the same units as those from models 1 and 2 we followed the advice of Verhulst et al. (2013) of adding the difference between the mean baseline LTL and the mean follow-up LTL to the RTM-corrected attrition values.

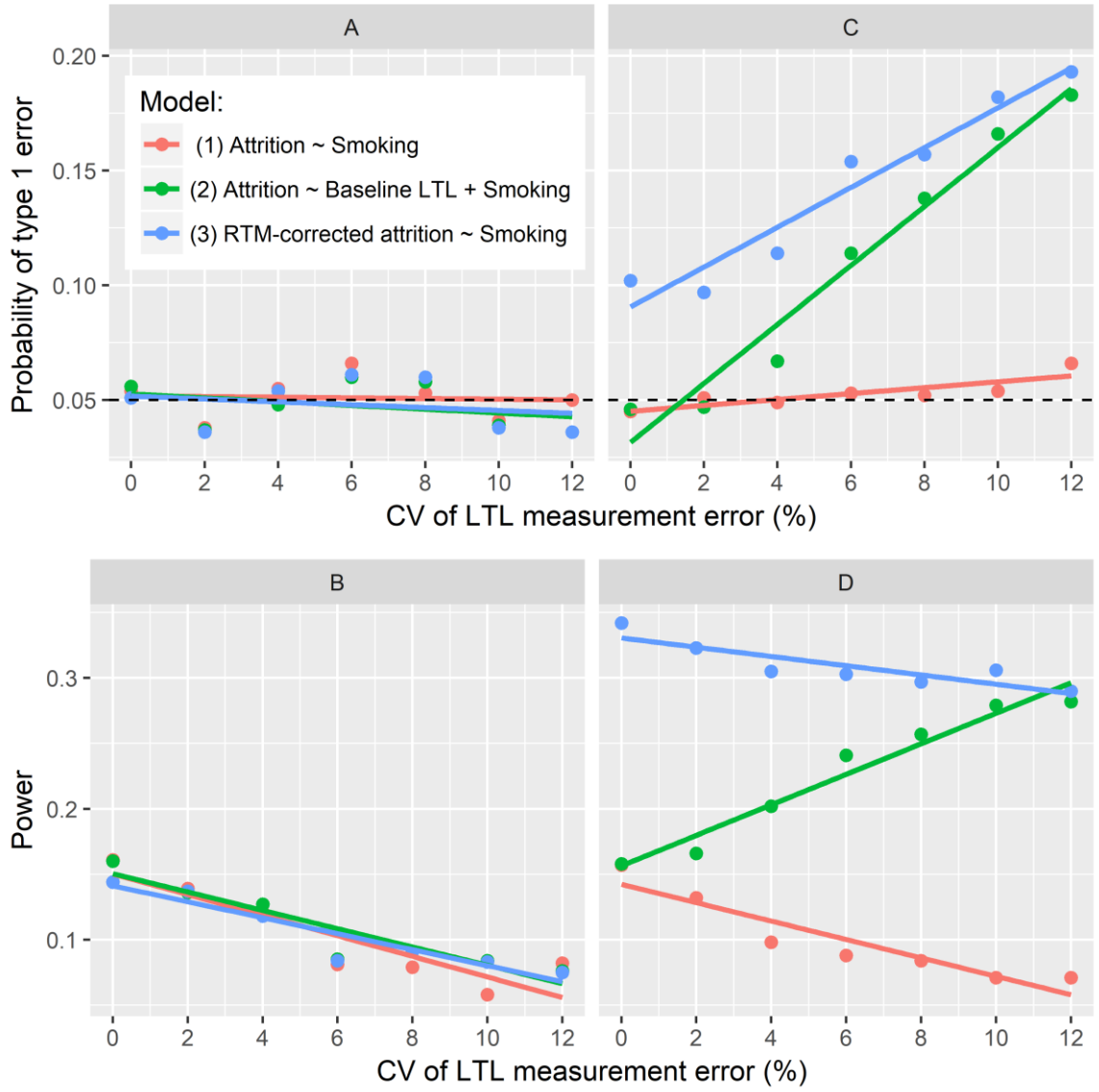

Figure 3. Type 1 error rate and power for the three different models under consideration. Data points represent the proportion of simulations yielding a p-value below 0.05 in 1000 replicate simulations, and lines are simple linear fits. Panels A and C show the probability of Type 1 errors occurring in scenarios A and C where there is no true effect of smoking on the LTL attrition. Panels B and D show the power to reject the null hypothesis of no effect of smoking on attrition in scenarios B and D where smokers have higher attrition than non-smokers by 2 bp/year.

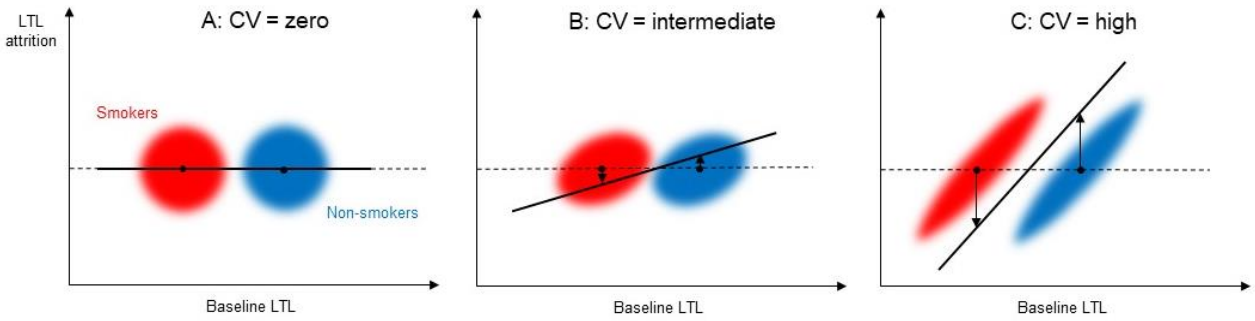

Figure 4. Cartoons illustrating how a true baseline difference in LTL between smokers and non-smokers results in a bias in the estimated effect of smoking on attrition in models that adjust for baseline LTL. The red and blue clouds indicate distributions of LTL and attrition measurements for smokers and non-smokers respectively. Note that in the scenario modelled here, there is no difference in attrition between smokers and non-smokers (the centres of all the clouds lie on the horizontal dotted line indicating an identical level of attrition). Panel A assumes no measurement error, panel B assumes an intermediate level of measurement error and panel C assumes high measurement error. The solid black line shows the relationship between baseline LTL and attrition; the slope of this line increases as measurement error increases due to regression to the mean, as shown by Verhulst et al. (2013). The arrows show the mean residual values for smokers and non-smokers after controlling for baseline LTL. These residuals increase (becoming more positive for smokers and more negative for non-smokers) as measurement error increases, thereby generating a spurious effect of smoking on attrition for non-zero values of measurement error.

Acknowledgements

We thank Dan Eisenberg for drawing our attention to the fact that the multiple regression models reported in the telomere attrition literature might lead to biased estimates of the effect of smoking on telomere attrition. Intriguingly for us all, Dan was right about the bias, but his original intuition about its direction appears to be the opposite to what is actually the case. DN is funded by an Advanced Grant from the European Research Council (AdG 666669, COMSTAR).

Supplementary information

"R_script_Bias_Simulation.R": Script for the simulations presented in the current manuscript.
"R_script_BatesonNettle_AgingCell_function.R": Function required by the above script.
"R_script_Verhulst_function.R": Function required by the above script.
"Supplementary_dataset.CSV": Dataset used to produce the figures in the current manuscript.

References

Astuti, Y. et al., 2017. Cigarette smoking and telomere length : A systematic review of 84 studies and meta-analysis. *Environmental Research*, 158, pp.480–489.

Aviv, A. et al., 2011. Impartial comparative analysis of measurement of leukocyte telomere length/DNA content by Southern blots and qPCR. *Nucleic Acids Research*, 39(20), pp.3–7.

Aviv, A. et al., 2009. Leukocyte telomere dynamics: Longitudinal findings among young adults in the Bogalusa Heart Study. *American Journal of Epidemiology*, 169(3), pp.323–329.

Bateson, M. & Nettle, D., 2016. The telomere lengthening conundrum - it could be biology. *Aging Cell*, 1474–9726.

Bendix, L. et al., 2014. Longitudinal changes in leukocyte telomere length and mortality in humans. *Journals of Gerontology - Series A Biological Sciences and Medical Sciences*, 69 A(2), pp.231–239.

Ehrlenbach, S. et al., 2009. Influences on the reduction of relative telomere length over 10 years in the population-based bruneck study: Introduction of a well-controlled high-throughput assay. *International Journal of Epidemiology*, 38(6), pp.1725–1734.

Farzaneh-Far, R. et al., 2010. Telomere length trajectory and its determinants in persons with coronary artery disease: Longitudinal findings from the heart and soul study. *PLoS ONE*, 5(1), p.e8612.

Greenland, S., 2003. Quantifying biases in causal models: classical confounding vs collider-stratification bias. *Epidemiology*, 14(3), pp.300–306.

Huzen, J. et al., 2014. Telomere length loss due to smoking and metabolic traits. *Journal of Internal Medicine*, 275, pp.155–163.

Müezzinler, A. et al., 2015. Smoking habits and leukocyte telomere length dynamics among older adults: results from the ESTHER cohort. *Experimental Gerontology*, 70, pp.18–25.

Révész, D. et al., 2016. Baseline biopsychosocial determinants of telomere length and 6-year attrition rate. *Psychoneuroendocrinology*, 67, pp.153–162.

Toupance, S. et al., 2017. Short telomeres, but not telomere attrition rates, are associated with carotid atherosclerosis. *Hypertension*, 70, pp.420–426.

Verhulst, S. et al., 2013. Do leukocyte telomere length dynamics depend on baseline telomere length? An analysis that corrects for “regression to the mean”. *European Journal of Epidemiology*, 28, pp.859–66.

Weischer, M., Bojesen, S.E. & Nordestgaard, B.G., 2014. Telomere shortening unrelated to smoking, body weight, physical activity, and alcohol intake: 4,576 general population individuals with repeat measurements 10 years apart. *PLoS Genetics*, 10(3).