

Supplementary Material

Contents:

Equations S1-S4. Mathematical proof that $m\Delta LTL$ is negatively related to LTL measurement error at baseline.

Equation S5. Correction for regression the mean.

Figure S1. Controlling for LTL_b increases power when there is a difference in LTL_b . Figure S2.

Figure S2. Varying the number of participants in the simulation had no impact on the accuracy of parameter estimates.

Figure S3. Increasing the number of participants increased the probability of type 1 errors when LTL_b was controlled for in scenario C. Figure S4. Increasing the number of participants increased the power in all scenarios and this effect was exaggerated by controlling for LTL_b in scenario D.

Figure S5. Increasing the true difference in $\Delta LTL \cdot year^{-1}$ between smokers and non-smokers had no further impact on the size of the biases compared to Figure 2.

Figure S6. Increasing the true difference in $\Delta LTL \cdot year^{-1}$ between smokers and non-smokers had no further impact on type 1 errors compared to Figure 3.

Figure S7. Increasing the true difference in $\Delta LTL \cdot year^{-1}$ between smokers and non-smokers increased power compared to Figure S1.

Figure S8. Assuming measurement error to be independent of LTL has no impact on the bias resulting from controlling for LTL_b compared to Figure 2.

Figure S9. Assuming measurement error to be independent of LTL has no impact on the increased probability of false-positive errors resulting from controlling for LTL_b compared to Figure 3.

Figure S10. Assuming measurement error to be independent of LTL has no impact on the increased probability of false-positive errors resulting from controlling for LTL_b compared to Figure S1.

Figure S11. Correcting $m\Delta LTL$ for regression to the mean due to measurement error does not eliminate bias when there are systematic differences in LTL_b .

Figure S12. When there are systematic differences in LTL_b controlling for LTL_b can either eliminate or even reverse estimates of the true difference in ΔLTL between smokers and non-smokers (see panel D).

Figure S13. Measurement error predicts biases for sex and BMI.

Table S1. Summary of the datasets analysed for sex and BMI.

Equations S1-S4. Mathematical proof that mΔLTL is negatively related to LTL measurement error at baseline.

ΔLTL is estimated from longitudinal datasets in which LTL is measured twice, at baseline (mLTL_b) and follow up (mLTL_{fu}). The measured ΔLTL (mΔLTL) for the *i*th individual is calculated via the following formula:

$$m\Delta LTL_i = (mLTL_{fu,i} - mLTL_{b,i}) \quad (\text{Equation S1})$$

Thus, a negative value of mΔLTL indicates telomere attrition and a positive value telomere elongation. An individual's measured LTL can be written as the sum of their true LTL and a measurement error:

$$mLTL_{b,i} = LTL_{b,i} + error_{b,i} \quad (\text{Equation S2})$$

$$mLTL_{fu,i} = LTL_{fu,i} + error_{fu,i} \quad (\text{Equation S3})$$

Here, error_{b,i} and error_{fu,i} are the errors introduced by measurement for that individual at baseline and follow-up respectively. We assume that error_{b,i} and error_{fu,i} are drawn from independent distributions. Equation 1 can now be expressed in terms of Equations S2 and S3:

$$\begin{aligned} m\Delta LTL_i &= LTL_{fu,i} + error_{fu,i} - (LTL_{b,i} + error_{b,i}) \\ &= \Delta LTL_i + error_{fu,i} - error_{b,i} \end{aligned} \quad (\text{Equation S4})$$

From Equation S4 it is evident that there is an inverse relationship between mΔLTL_i and error_{b,i}. In other words, a larger positive baseline measurement error for an individual results in a more negative mΔLTL, which implies greater measured telomere attrition, for that individual. This is an example of so-called regression to the mean: baseline values are negatively correlated with measures of change because individuals with high mLTL_b generally have smaller mLTL_{fu} and vice versa.

Equation S5. Correction for regression the mean.

To correct for regression mean we used the equation suggested by Verhulst et al. (1). The change from the baseline measure X_1 (mLTL_b) to follow up measure X_2 (mLTL_{fu}) is adjusted for the regression to the mean effect to yield a correct value D as follows:

$$D = \rho(X_1 - \bar{X}_1) - (X_2 - \bar{X}_2) \quad (\text{Equation S5})$$

where

$$\rho = \frac{2rS_1S_2}{S_1^2 + S_2^2} \quad (\text{Equation S6})$$

in which r is the correlation between X_1 and X_2 .

An R function that implements this correction is included in the R script accompanying this paper.

Reference

1. Verhulst S, Aviv A, Benetos A, Berenson GS, Kark JD. Do leukocyte telomere length dynamics depend on baseline telomere length? An analysis that corrects for “regression to the mean”. Eur J Epidemiol. 2013;28:859–66.

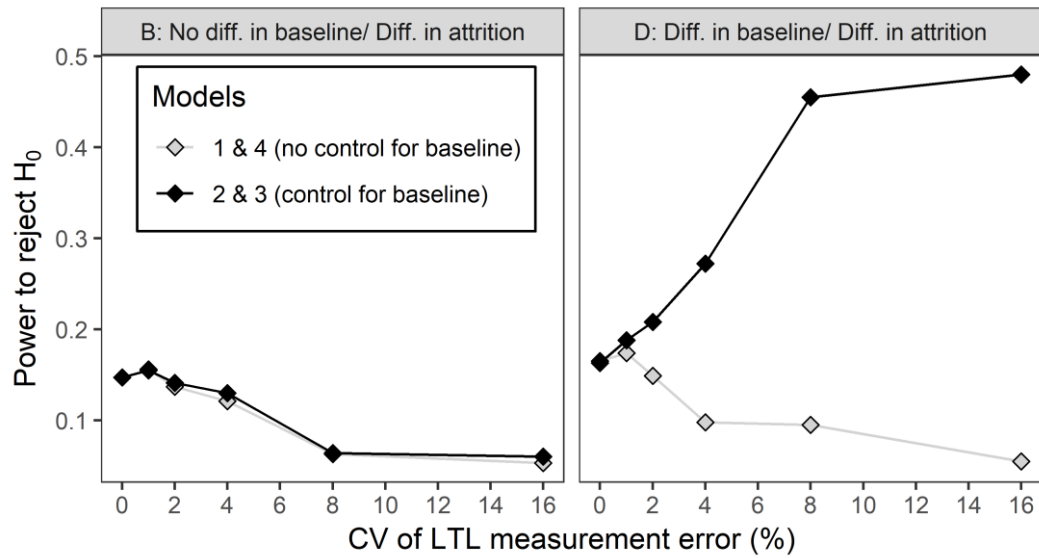


Figure S1. Controlling for LTL_b increases power when there is a difference in LTL_b . Power as a function of measurement error (CV) for the four models under consideration. Data points represent the proportion of simulations yielding a p-value below 0.05 in 1000 replicate simulations. The left and right panels show the power in scenarios B and D respectively. The increase in power with increasing CV in scenario D that occurs with models 2 and 3 reflects the bias in parameter estimates for these models shown in Figure 2D. Power is generally low because of the small true effect size assumed in this simulation of only $-2 \text{ bp} \cdot \text{year}^{-1}$. The difference in LTL_b between smokers and non-smokers in scenario D was LTL_b 141 bp shorter in smokers.

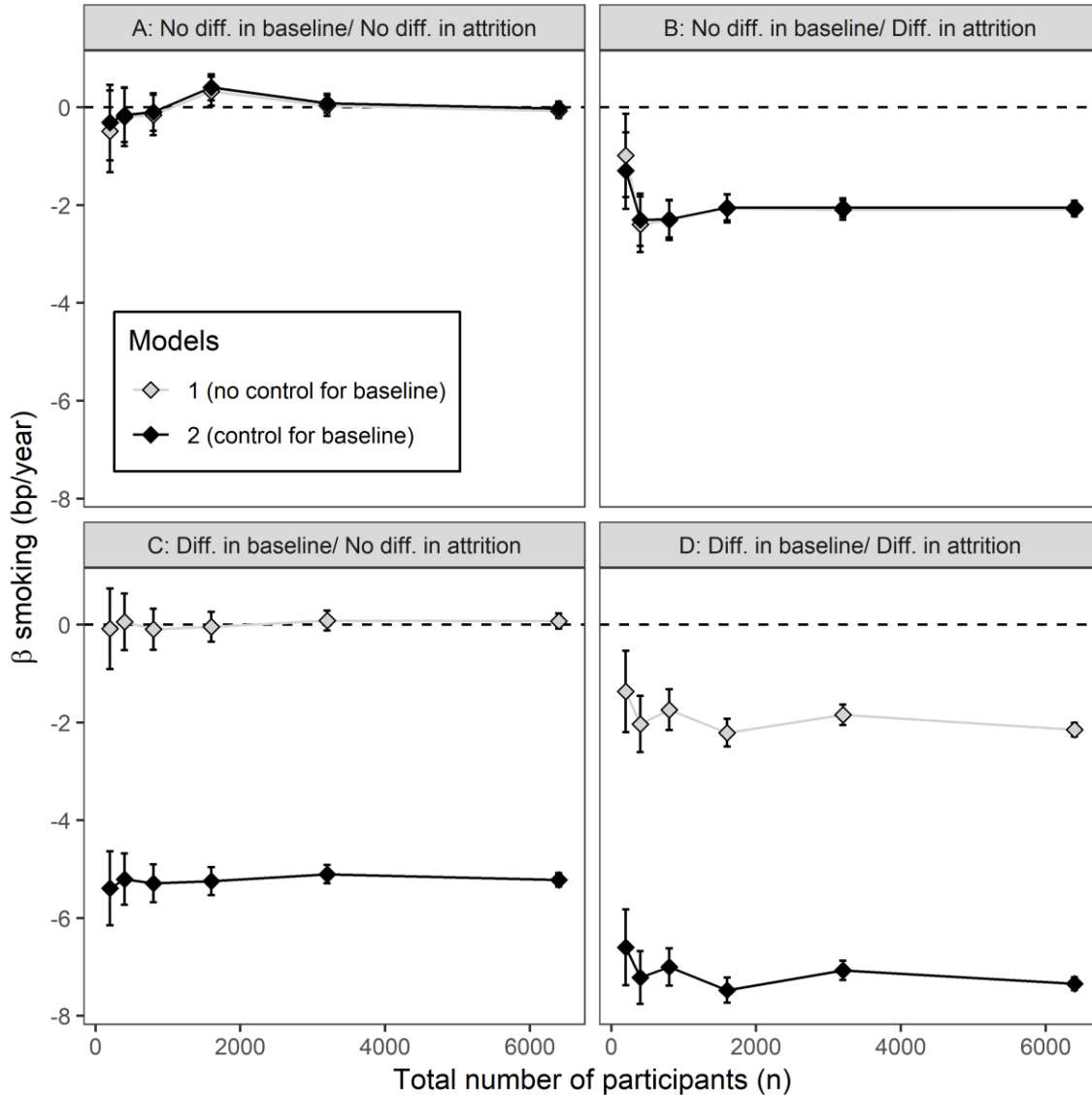


Figure S2. Varying the number of participants in the simulation had no impact on the accuracy of parameter estimates. Panels show the estimated difference in mΔLTL between smokers and non-smokers as a function of the number of participants in the simulations. The β estimates were obtained by fitting two alternative models to data simulated given four sets of assumptions regarding the true differences between smokers and non-smokers (scenarios A-D in Table 2). The dashed lines indicate no difference in mΔLTL between smokers and non-smokers. Data points are the mean \pm 95% confidence intervals obtained from modelling the data from 1000 replicate simulations. The four scenarios were as follows: (A) no difference in LTL_b and no difference in ΔLTL; (B) no difference in LTL_b but a true difference in ΔLTL; (C) a true difference in LTL_b but no difference in ΔLTL; and (D) A true difference in LTL_b and a true difference in ΔLTL. The true difference in LTL_b between smokers and non-smokers in scenarios C and D was LTL_b 141 bp shorter in smokers. The true difference in ΔLTL between smokers and non-smokers in scenarios B and D was ΔLTL -2 bp.year⁻¹ greater in smokers. CV was fixed at 8% for this simulation in order to illustrate the impact of varying participant number.

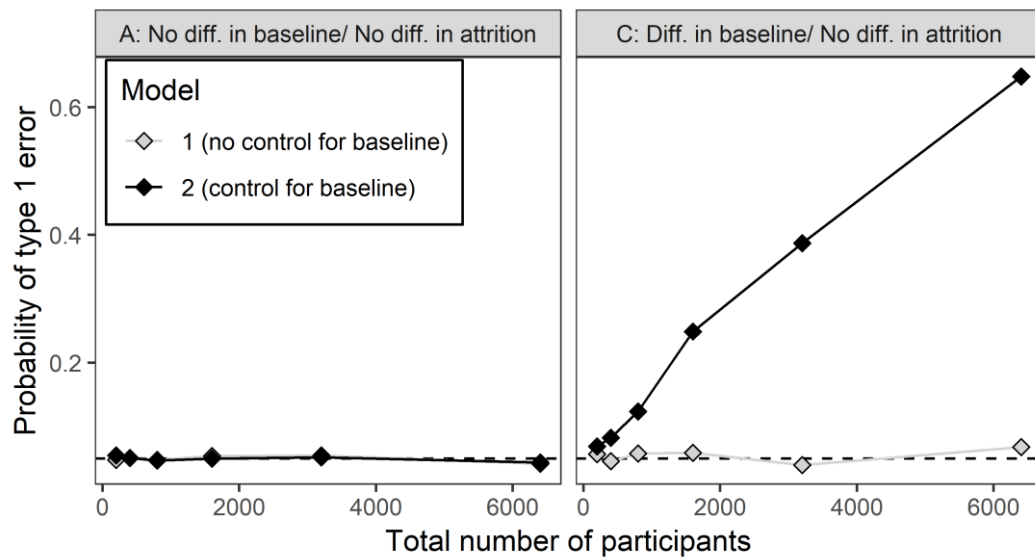


Figure S3. Increasing the number of participants increased the probability of type 1 errors when LTL_b was controlled for in scenario C. Probability of a type 1 error as a function of number of participants for models 1 and 2. Data points represent the proportion of simulations yielding a p-value below 0.05 in 1000 replicate simulations. The left and right panels show the probability of type 1 errors in scenarios A and C respectively. The difference in LTL_b between smokers and non-smokers in scenario C was LTL_b 141 bp shorter in smokers. CV was fixed at 8% for this simulation in order to illustrate the impact of varying participant number.

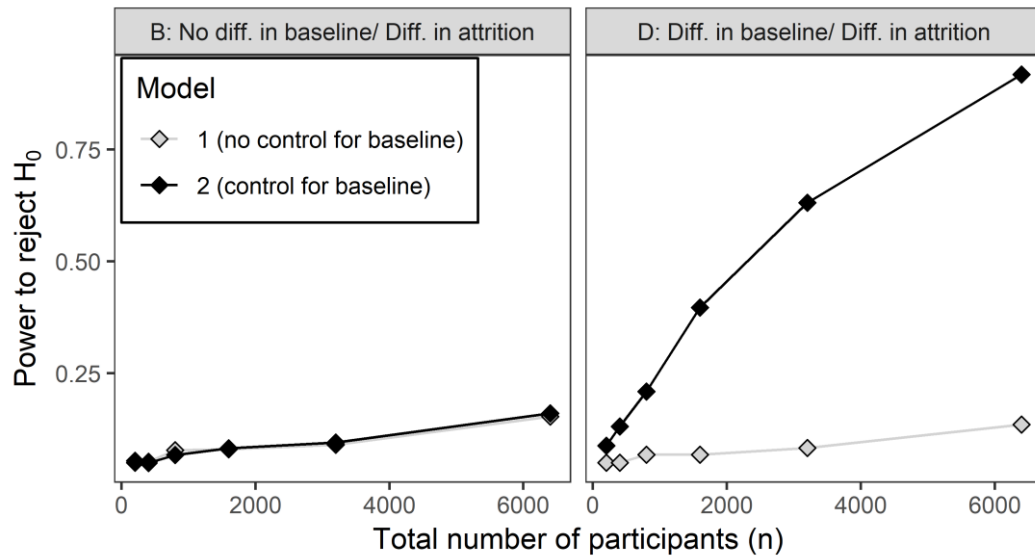


Figure S4. Increasing the number of participants increased the power in all scenarios and this effect was exaggerated by controlling for LTL_b in scenario D. Power as a function of the total number of participants (n) for models 1 and 2. Data points represent the proportion of simulations yielding a p-value below 0.05 in 1000 replicate simulations. The left and right panels show the power in scenarios B and D respectively. The increase in power with increasing CV seen with model 2 in scenario D reflects the bias in parameter estimates shown in Figure S2D. Power is generally low because of the small true effect size assumed in this simulation ($-2 \text{ bp} \cdot \text{year}^{-1}$). The difference in LTL_b between smokers and non-smokers in scenario D was LTL_b 141 bp shorter in smokers. CV was fixed at 8% for this simulation in order to illustrate the impact of varying participant number.

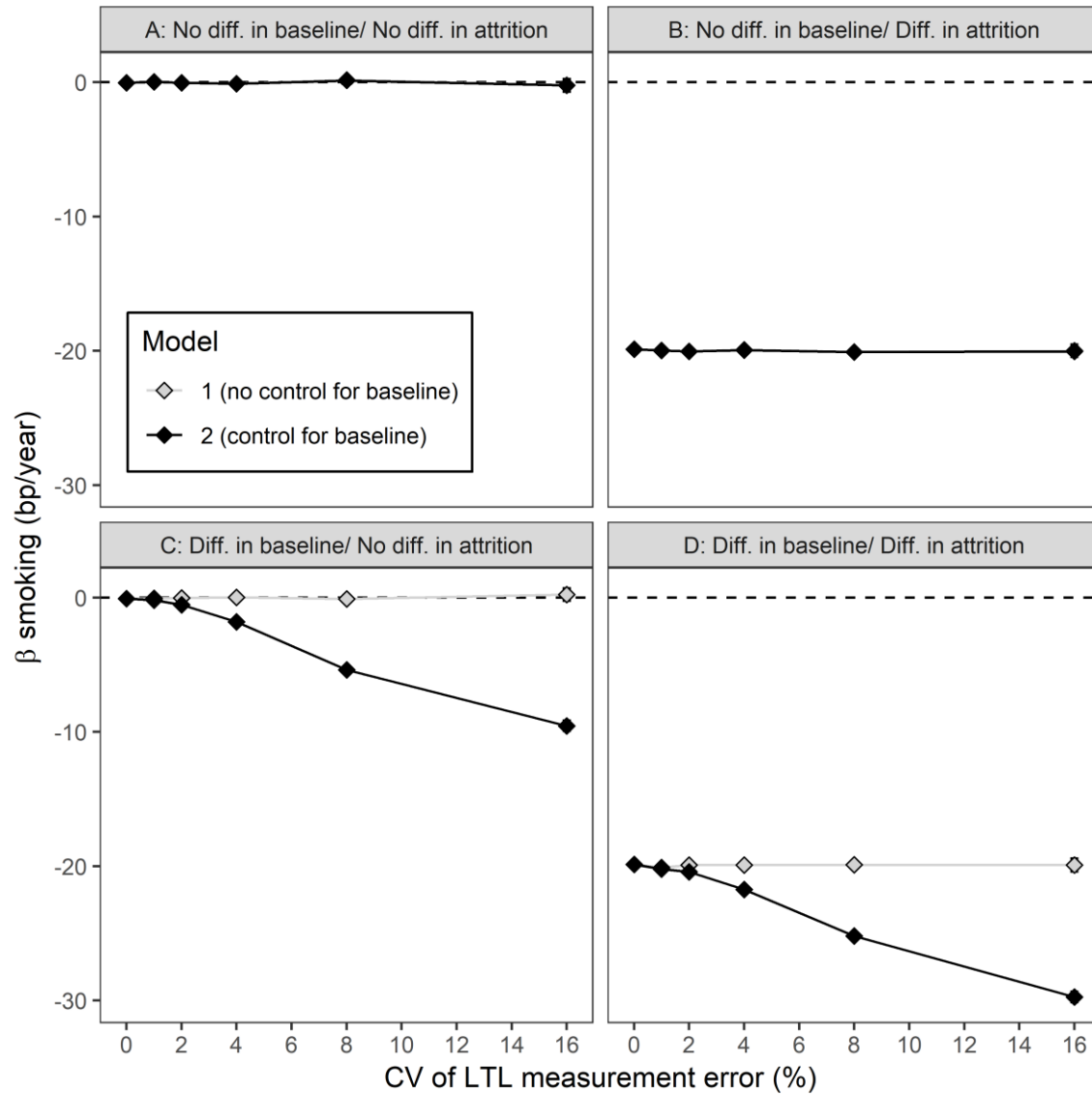


Figure S5. Increasing the true difference in Δ LTL. year^{-1} between smokers and non-smokers had no further impact on the size of the biases compared to Figure 2. Panels show the estimated difference in m Δ LTL between smokers and non-smokers as a function of measurement error. The β estimates were obtained by fitting two alternative models to data simulated given four sets of assumptions regarding the true differences between smokers and non-smokers. The four scenarios are identical to those given in Table 2, other than that the true difference in Δ LTL between smokers and non-smokers in scenarios B and D was Δ LTL -20 bp. year^{-1} (compared with -2 bp. year^{-1} for the simulation shown in Figure 2). The dashed lines indicate no difference in m Δ LTL between smokers and non-smokers. Data points are the mean \pm 95% confidence intervals obtained from modelling the data from 1000 replicate simulations.

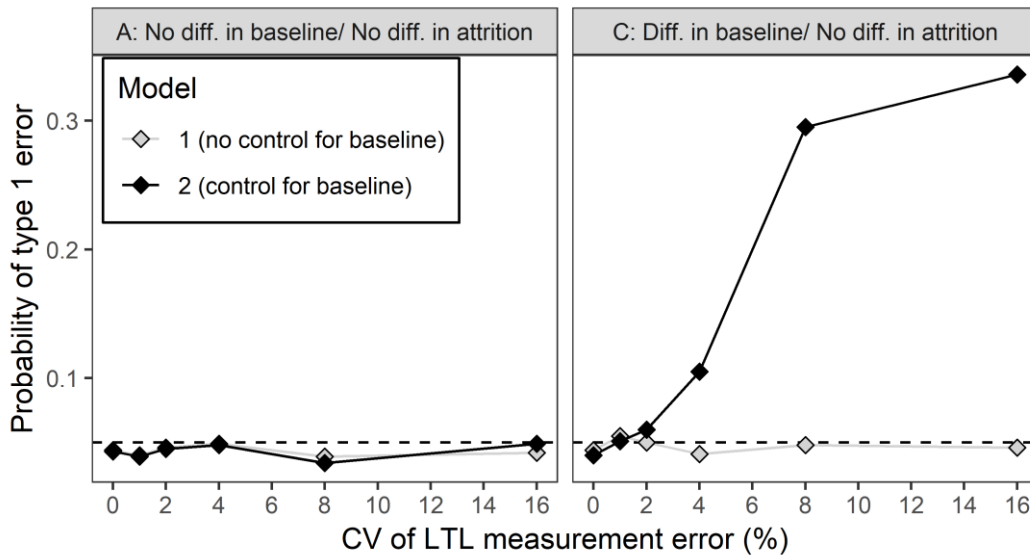


Figure S6. Increasing the true difference in $\Delta\text{LTL}\cdot\text{year}^{-1}$ between smokers and non-smokers had no further impact on type 1 errors compared to Figure 3. Probability of a type 1 error as a function of measurement error (CV) for models 1 and 2. Data points represent the proportion of simulations yielding a p-value below 0.05 in 1000 replicate simulations. The left and right panels show the probability of type 1 errors in scenarios A and C respectively. The difference in LTL_b between smokers and non-smokers in scenario C was LTL_b 141 bp shorter in smokers.

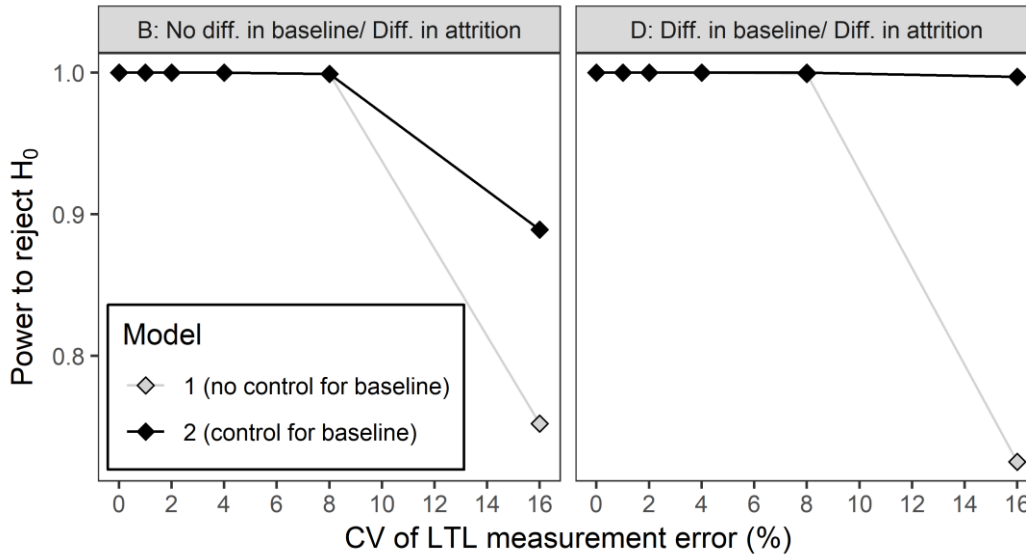


Figure S7. Increasing the true difference in $\Delta\text{LTL} \cdot \text{year}^{-1}$ between smokers and non-smokers increased power compared to Figure S1. Power as a function of measurement error for models 1 and 2. Data points represent the proportion of simulations yielding a p-value below 0.05 in 1000 replicate simulations. The left and right panels show the power in scenarios B and D respectively. The higher power at a CV of 16% seen with model 2 in scenario D reflects the bias in parameter estimates shown in Figure S5D. Power is generally high because of the large true effect size assumed in this simulation ($-20 \text{ bp} \cdot \text{year}^{-1}$). The difference in LTL_b between smokers and non-smokers in scenario D was LTL_b 141 bp shorter in smokers.

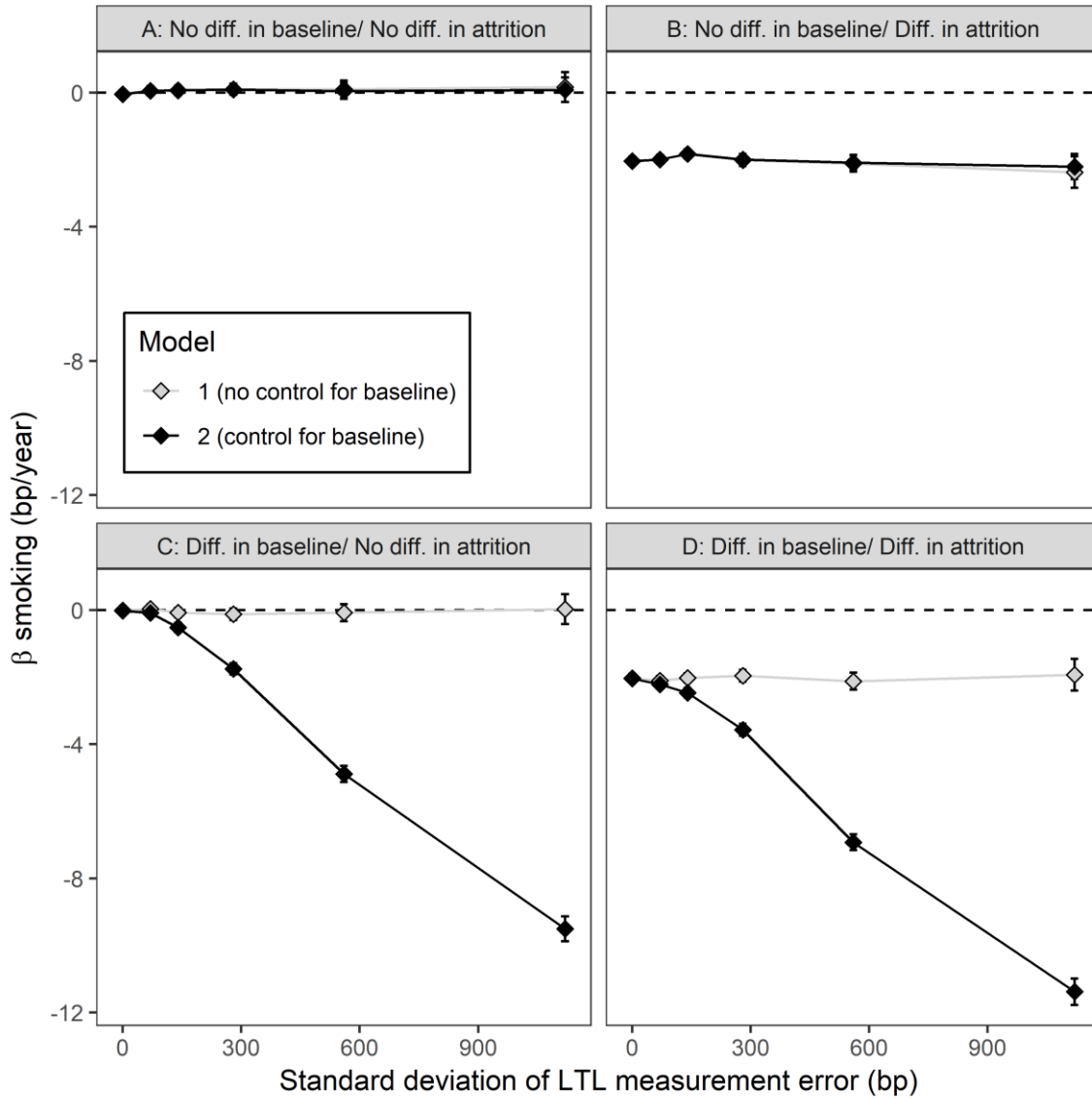


Figure S8. Assuming measurement error to be independent of LTL has no impact on the bias resulting from controlling for LTL_b compared to Figure 2. Panels show the estimated difference in mΔLTL between smokers and non-smokers as a function of measurement error here implemented as a fixed standard deviation (as opposed to as a CV in Figure 2). The β estimates were obtained by fitting two alternative models to data simulated given four sets of assumptions regarding the true differences between smokers and non-smokers. Data points are the mean ± 95% confidence intervals obtained from modelling the data from 1000 replicate simulations. The four scenarios are identical to those given in Table 2 and plotted in Figure 2. The dashed lines indicate no difference in mΔLTL between smokers and non-smokers. The true difference in LTL_b between smokers and non-smokers in scenarios C and D was LTL_b 141 bp shorter in smokers. The true difference in ΔLTL between smokers and non-smokers in scenarios B and D was ΔLTL -2 bp·year⁻¹ greater in smokers.

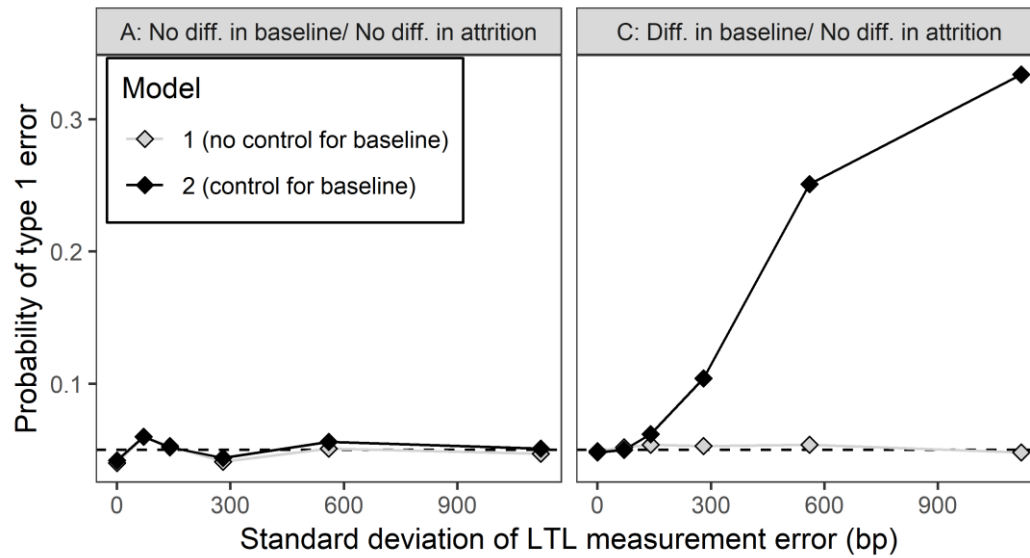


Figure S9. Assuming measurement error to be independent of LTL has no impact on the increased probability of false-positive errors resulting from controlling for LTL_b compared to Figure 3.

Probability of a type 1 error as a function of measurement error here implemented as a fixed standard deviation (as opposed to as a CV) for models 1 and 2. Data points represent the proportion of simulations yielding a p-value below 0.05 in 1000 replicate simulations. The left and right panels show the probability of type 1 errors in scenarios A and C respectively. The difference in LTL_b between smokers and non-smokers in scenario C was LTL_b 141 bp shorter in smokers.

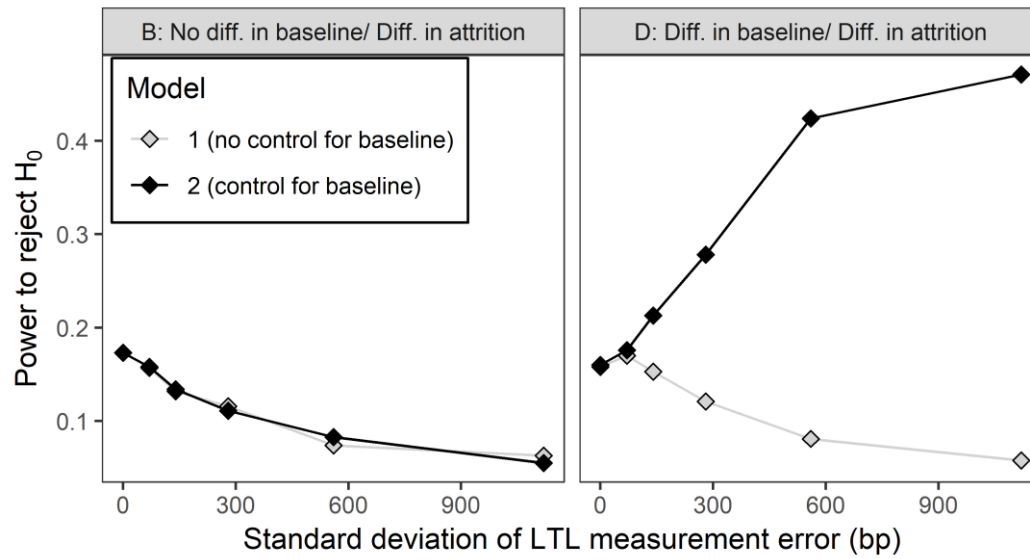


Figure S10. Assuming measurement error to be independent of LTL has no impact on the increased probability of false-positive errors resulting from controlling for LTL_b compared to Figure S1. Power as a function of measurement error here implemented as a fixed standard deviation (as opposed to a CV) for models 1 and 2. Data points represent the proportion of simulations yielding a p-value below 0.05 in 1000 replicate simulations. The left and right panels show the power in scenarios B and D respectively. The increase in power with increasing measurement error in scenario D that occurs with model 2 reflects the bias in parameter estimates for this model shown in Figure S8D. Power is generally low because of the small true effect size assumed in this simulation of only $-2 \text{ bp} \cdot \text{year}^{-1}$. The difference in LTL_b between smokers and non-smokers in scenario D was LTL_b 141 bp shorter in smokers.

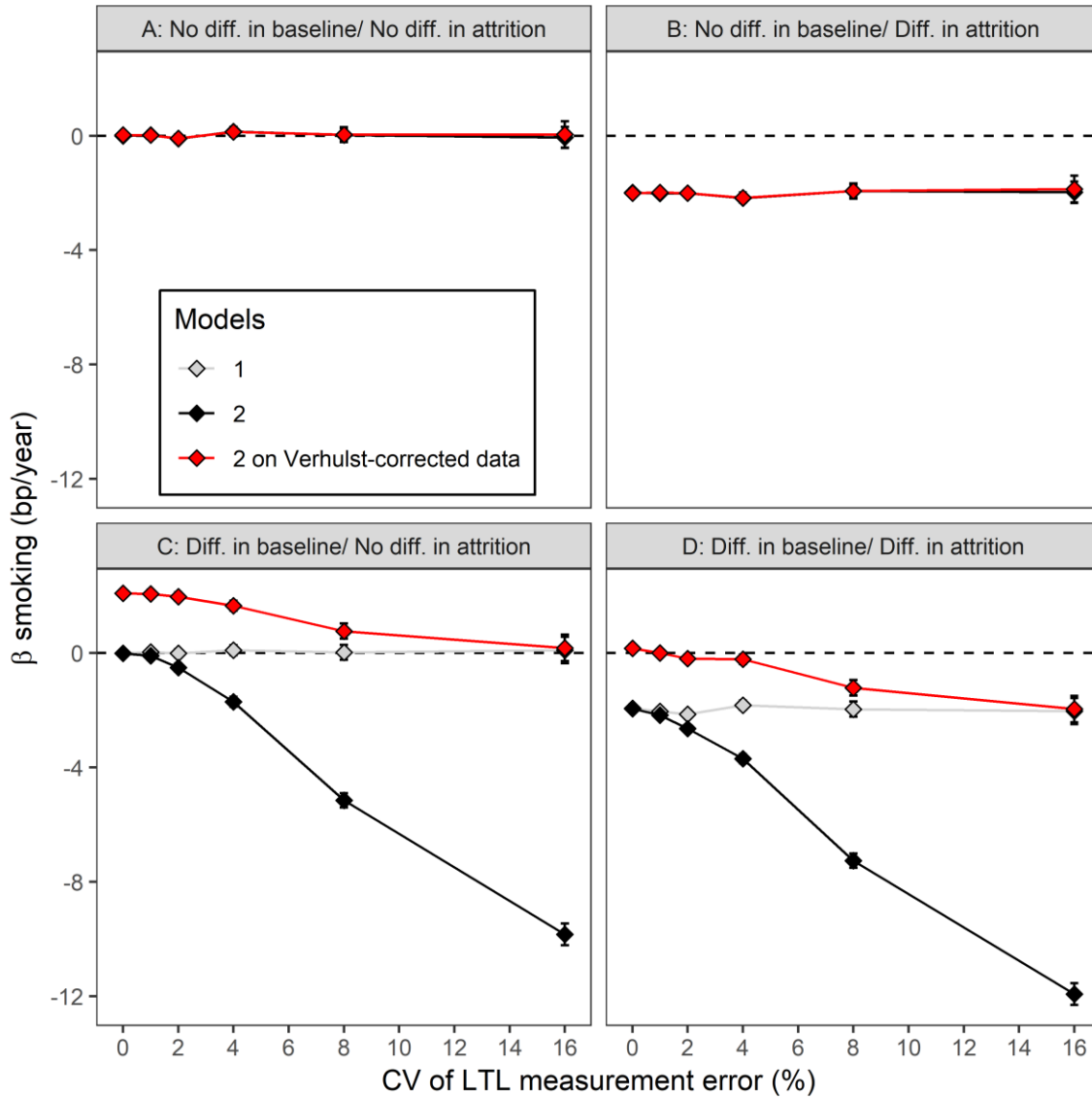


Figure S11. Correcting mΔLTL for regression to the mean due to measurement error does not eliminate bias when there are systematic differences in LTL_b . Panels show the estimated difference in mΔLTL between smokers and non-smokers as a function of measurement error (CV). The β estimates were obtained by fitting two alternative models to data simulated given four sets of assumptions regarding the true differences between smokers and non-smokers (scenarios A-D in Table 2). The dashed lines indicate no difference in mΔLTL between smokers and non-smokers. Data points are the mean \pm 95% confidence intervals obtained from modelling the data from 1000 replicate simulations. The four scenarios were as follows: (A) no difference in LTL_b and no difference in ΔLTL ; (B) no difference in LTL_b but a true difference in ΔLTL ; (C) a true difference in LTL_b but no difference in ΔLTL ; and (D) A true difference in LTL_b and a true difference in ΔLTL . The true difference in LTL_b between smokers and non-smokers in scenarios C and D was LTL_b 141 bp shorter in smokers. The true difference in ΔLTL between smokers and non-smokers in scenarios B and D was ΔLTL -2 bp. $year^{-1}$ greater in smokers. The red data points show the effect of correcting mΔLTL for regression to the mean using Verhulst et al.'s D (see Equation S5) prior to fitting model 2.

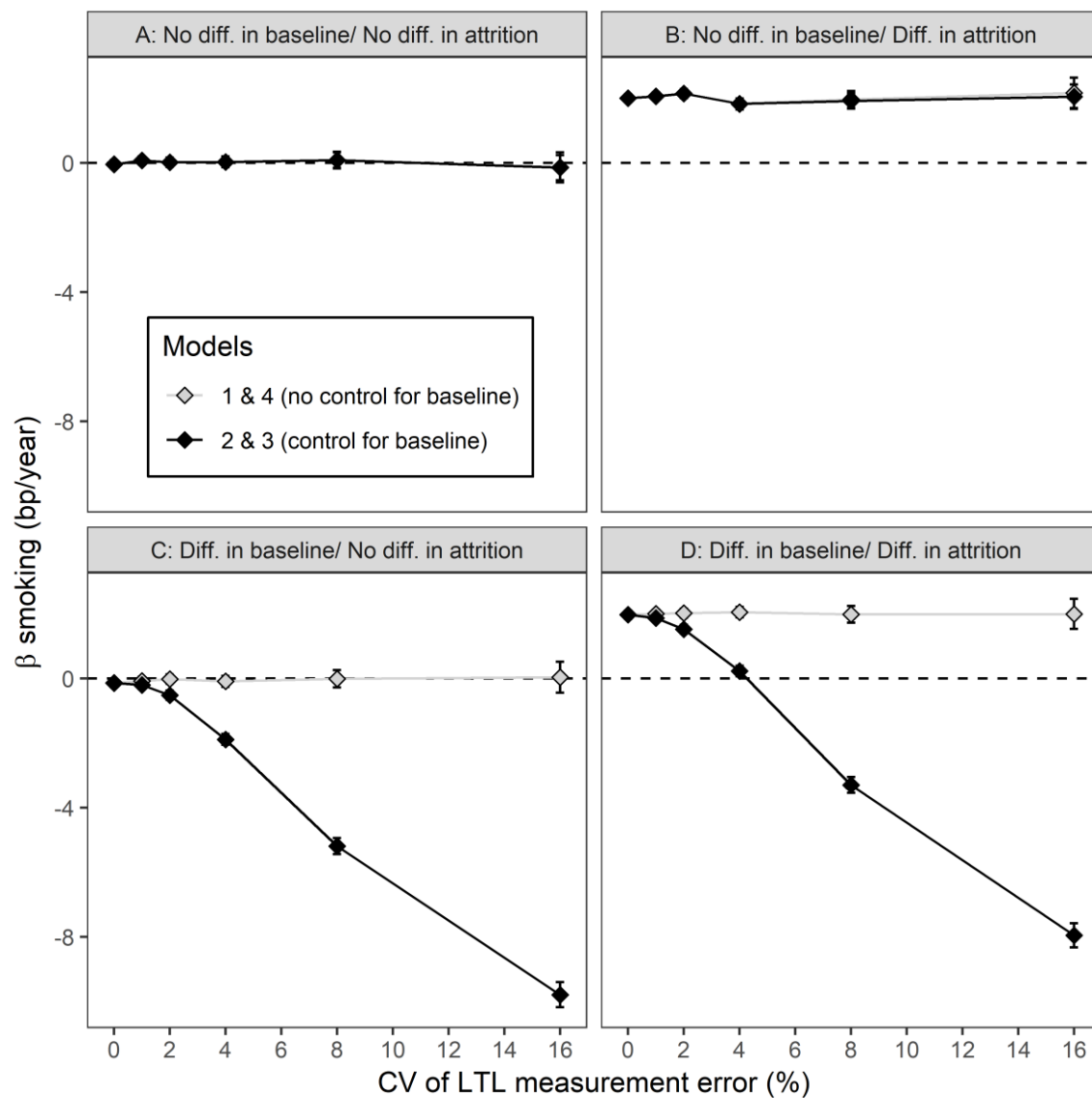


Figure S12. When there are systematic differences in LTL_b controlling for LTL_b can either eliminate or even reverse estimates of the true difference in Δ LTL between smokers and non-smokers (see panel D). Panels show the estimated difference in m Δ LTL between smokers and non-smokers as a function of measurement error (CV). The β estimates were obtained by fitting four alternative models to data simulated given four sets of assumptions regarding the true differences between smokers and non-smokers. The four scenarios are as in Table 2, with the exception that the true difference in Δ LTL in scenarios B and D was reversed, such that Δ LTL was 2 bp·year⁻¹ less in smokers (i.e. β smoking = 2 bp·year⁻¹). The dashed lines indicate no difference in m Δ LTL between smokers and non-smokers. Data points are the mean \pm 95% confidence intervals obtained from modelling the data from 1000 replicate simulations. Panel D shows a scenario in which the true difference in Δ LTL is eliminated (β is estimated as ~ 0) when CV is equal to 4% and reversed when CV is greater than 4% when models 2 and 3 are used.

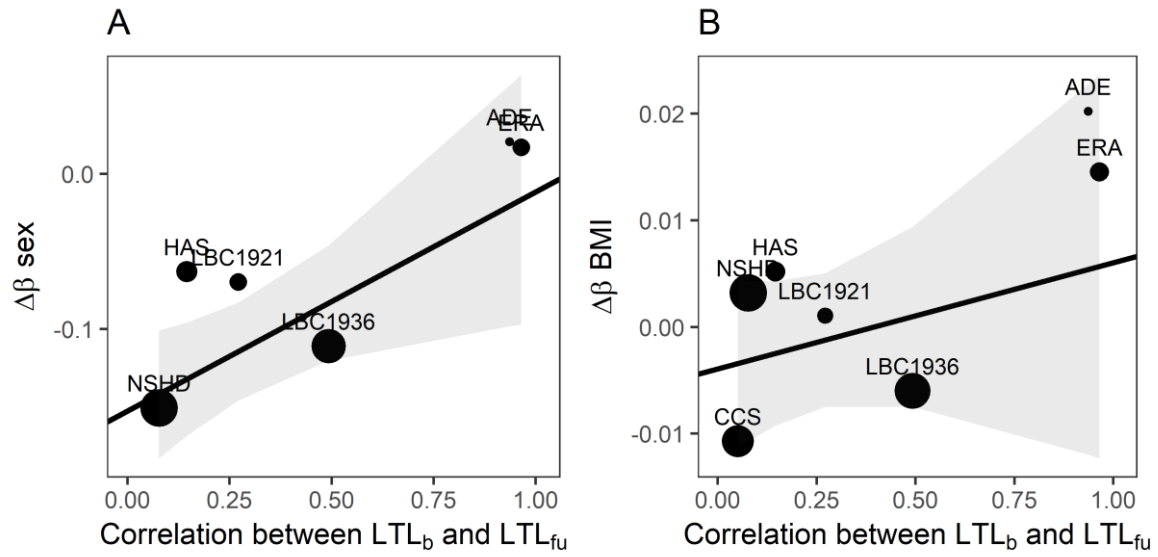


Figure S13. Measurement error predicts biases for sex and BMI. The combined dataset available for analysing effects of sex and BMI included data from 3,313 adults, comprising 2,077 males and 1,236 females (see Table S1). As observed for smoking, there is a positive relationship between the LTL_b-LTL_{fu} correlation coefficient (a proxy for measurement error) and $\Delta\beta$ (the difference between the estimates derived from models 1 and 2) for both sex and BMI. A: Scatterplot showing the relationship between the correlation between LTL_b and LTL_{fu} and the difference between the β coefficients for sex derived from models 1 and 2 (weighted linear regression for sex: $\beta \pm se = 0.14 \pm 0.06$, $t = 2.43$, $p = 0.0722$). B: As panel A, but the β coefficients are for BMI (weighted linear regression for BMI $\beta \pm se = 0.01 \pm 0.01$, $t = 0.82$, $p = 0.4480$). In both panels, the solid black line was derived from a linear regression in which the points were weighted by the number of participants in each cohort and the grey ribbon shows the 95% confidence interval for this line.

Table S1. Summary of the datasets analysed for sex and BMI.

| Cohort ^a | Number of participants | | Correlation between LTL _b and LTL _{fu} | Differences between sexes in telomere length/attrition ^b (standardised β [s.e.]) | | | Associations between BMI and telomere length/attrition ^c (standardised β [s.e.]) | | |
|---------------------|------------------------|--------|--|--|---------------------------------|-----------------|--|---------------------------------|-----------------|
| | | | | | | | | | |
| | Male | Female | | LTL _b | Δ LTL.year ⁻¹ | | LTL _b | Δ LTL.year ⁻¹ | |
| | | | | | Model 1 | Model 2 | | Model 1 | Model 2 |
| ADE | 33 | 35 | 0.94 | -0.55 [0.23] | -0.18 [0.24] | -0.16 [0.26] | -0.22 [0.12] | 0.14 [0.12] | 0.16 [0.13] |
| CCS | 756 | 0 | 0.05 | NA ^d | NA ^d | NA ^d | 0.01 [0.04] | 0.023 [0.04] | 0.012 [0.02] |
| ERA | 108 | 54 | 0.96 | -0.065 [0.17] | -0.01 [0.17] | 0.01 [0.16] | -0.06 [0.08] | -0.05 [0.08] | -0.04 [0.08] |
| HAS | 158 | 95 | 0.15 | 0.09 [0.13] | 0.13 [0.13] | 0.06 [0.09] | -0.01 [0.06] | 0.06 [0.06] | 0.07 [0.04] |
| LBC1921 | 78 | 81 | 0.27 | 0.37 [0.16] | -0.12 [0.16] | -0.19 [0.16] | -0.01 [0.08] | -0.01 [0.10] | -0.01 [0.08] |
| LBC1936 | 444 | 414 | 0.49 | 0.38 [0.07] | 0.21 [0.07] | 0.09 [0.07] | 0.02 [0.03] | 0.00 [0.03] | -0.00 [0.03] |
| NSHD | 500 | 557 | 0.08 | 0.19 [0.06] | 0.25 [0.06] | 0.10 [0.04] | -0.00 [0.03] | -0.01 [0.03] | -0.01 [0.02] |

Notes: ^aAcronyms for cohorts as defined in Table 3 of the main text. ^bFor sex, positive standardised β s indicate that males have longer LTL_b and greater $m\Delta$ LTL.year⁻¹. ^cFor BMI, we used the mean of BMI at baseline and follow-up where this was available; and otherwise either BMI at baseline or BMI at follow-up, whichever was available. For BMI, positive standardised β s indicate that participants with higher BMI have longer LTL_b and greater $m\Delta$ LTL.year⁻¹. ^dWe were unable to analyse the effect of sex for the Caerphilly Cohort Study (CCS), since this cohort was restricted to male participants.