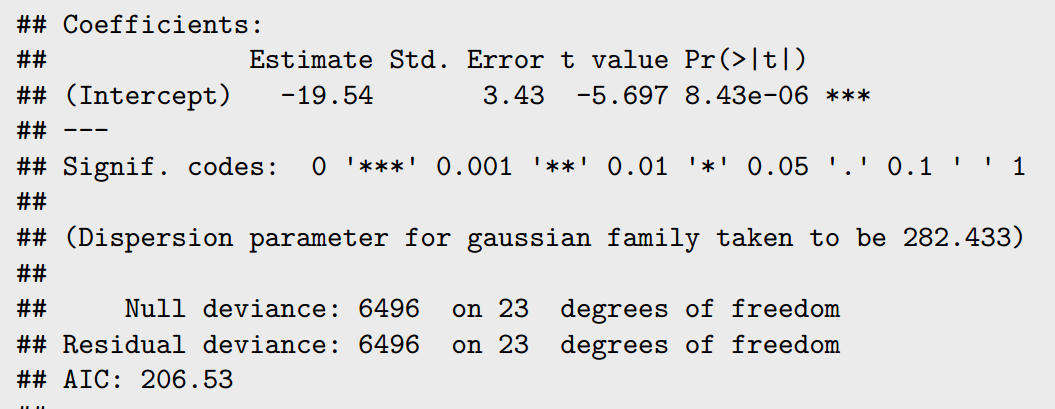
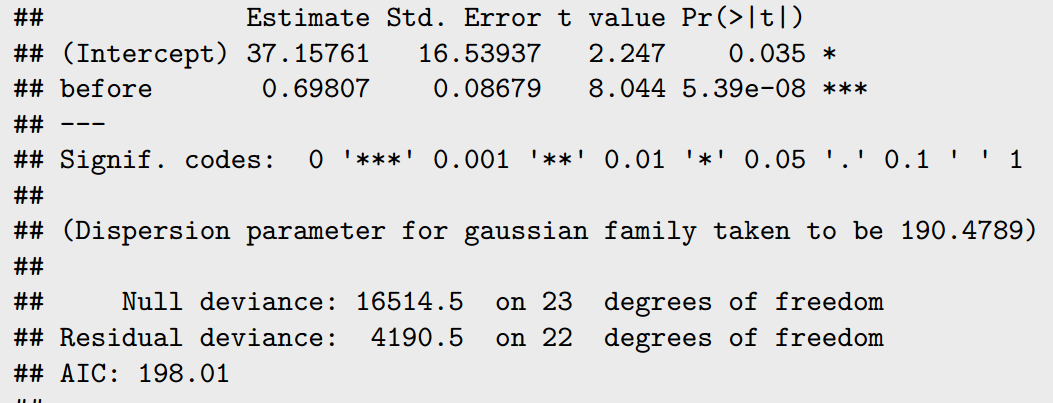
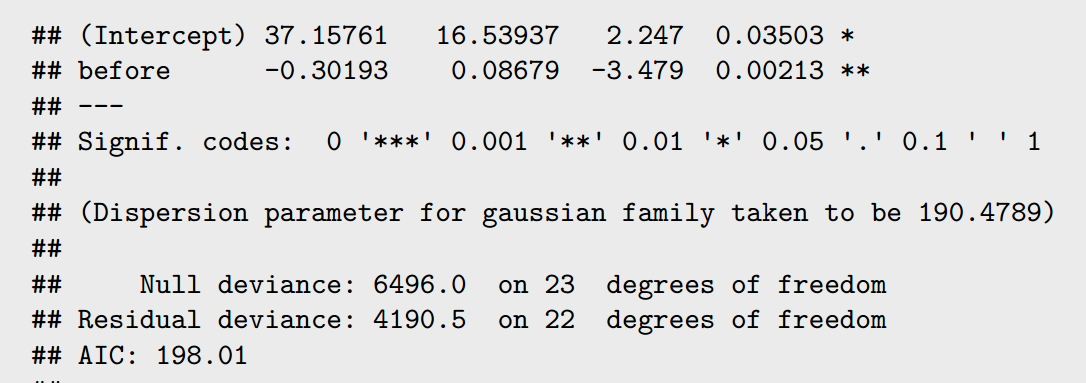
**(2) Comparison of 4 models:**

**(i) change-score model, (ii) baseline-as-covariate model, (iii) hybrid model, (iv) a longitudinal model.**

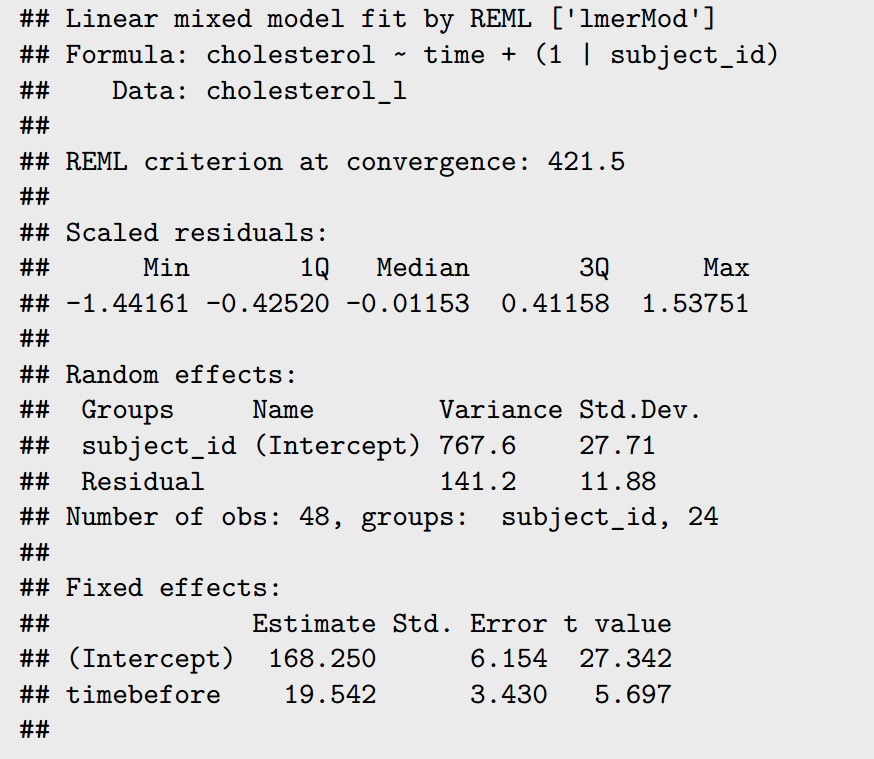


(i) change-score model just accounts for the intercept (mean of the change). Thus, this model not fits the data well and have the highest AIC in all 3 linear models. This model is too simple.

(ii) baseline-as-covariate model is actually the same with the (iii) hybrid model. The difference of the coefficients of “before” is 1, which is just the move of “1 before” from the left side to the right side of the equation. The residual deviance from these two models are the same, and both less than the change-score model.

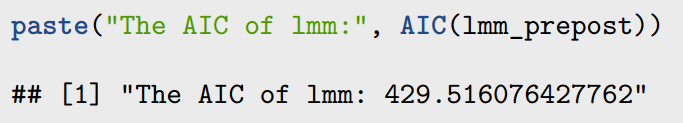


This means these 2 models include more variance. Thus they have the same AIC, lower than the first model. Both of models show the significant association between two time points. These models are reasonable and easy to interpret. But they answer slightly different hypotheses.

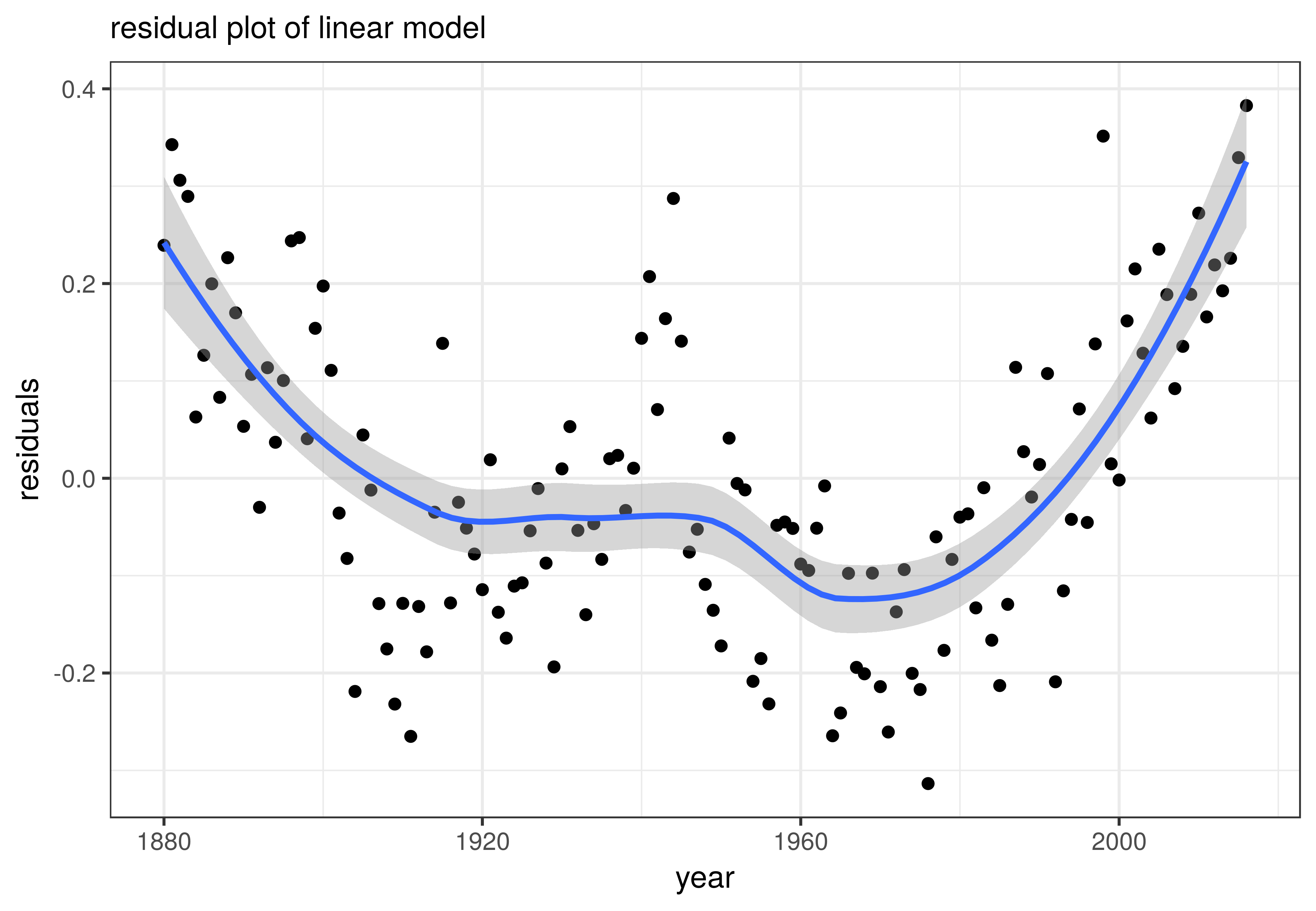


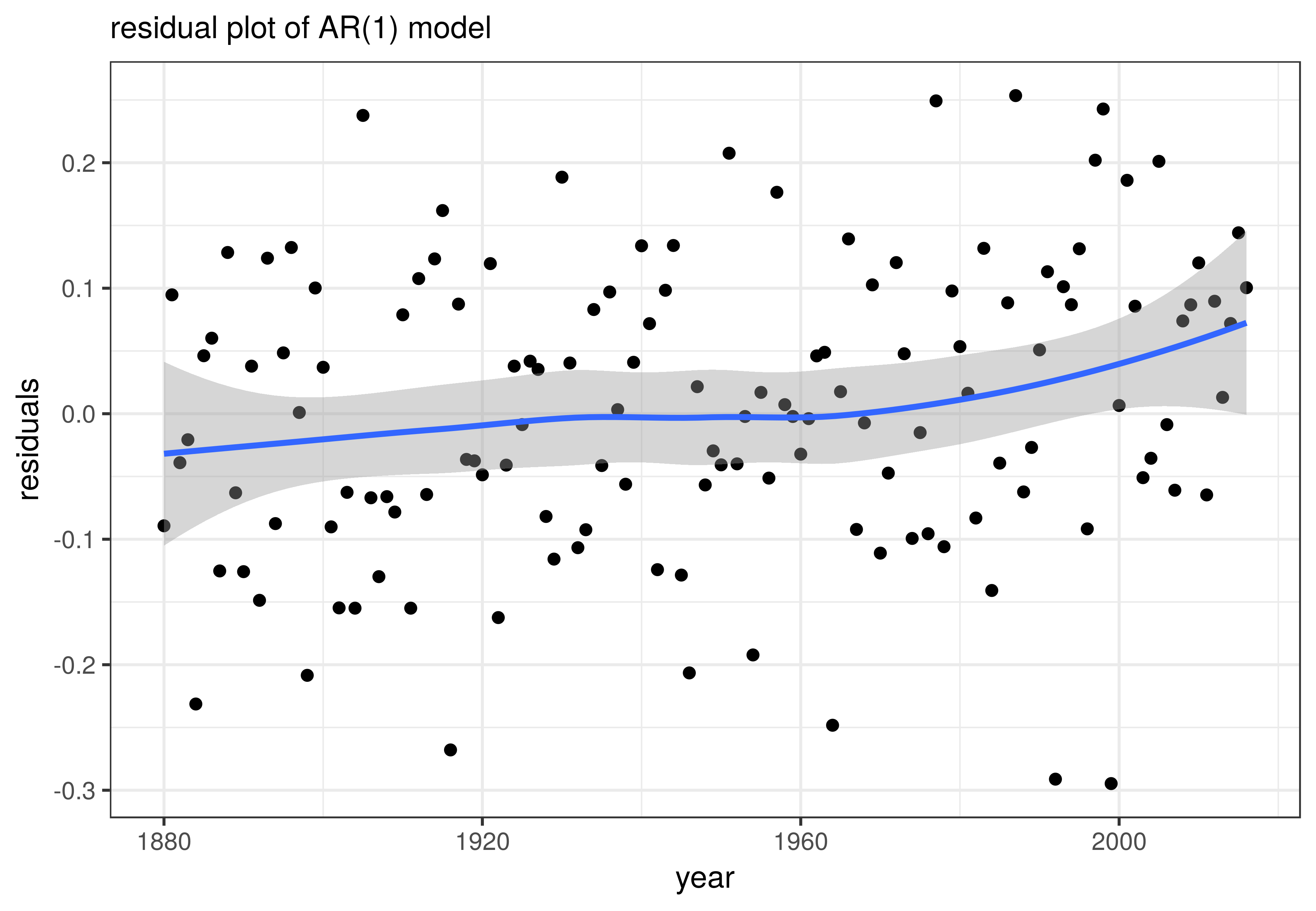
(iv) the linear mixed model with random intercept. The pre-post dataset has 2 repeat measurements, only enough for a compound symmetry covariance structure, which is the random intercept model. This model includes the “time” (before, after) as the binary fixed effect. This model also shows the significant association between cholesterol levels and time points. This model has a better fit to the data points and using the long form of dataset. The hypotheses is to test the association between cholesterol levels and time,

different from the change & baseline. However, since there are only 2 time points, this model is not necessary and too complex.

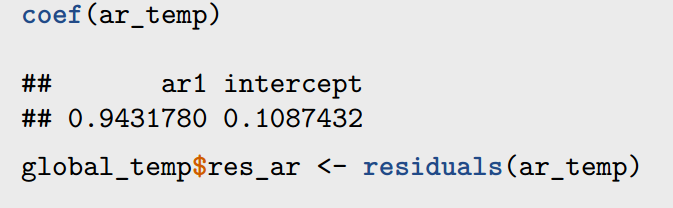


(3) Time series data.

 a. the residual plot shows the simple linear regression not works in this case. The residuals are not normally distributed, with a “W” shape pattern. We need try a different model to get an even and balanced residual cloud.



b. and c. The estimated phi in AR(1) process is 0.9432, there is also the correlation between two neighbors. This model fits the data better compared with the simple linear model. The residual is roughly equally distributed. The complexity of the mean by AR(1) model reduces the error of the model.

 d. the average increase in temperature per decade is 0.06 ºC.

