Supplementary file 2. The Preston-Glei-Wilmoth- method and its application to inequalities in mortality

The method consists of three steps. First, the attributable fraction of lung cancer deaths due to smoking is derived by subtracting the CPS-II rates of lung cancer death among non-smokers from the observed lung cancer death rates. Second, the attributable fraction of non-lung cancer deaths due to smoking is estimated using the observed lung cancer death rate, the expected lung cancer death rate among non-smokers, and a regression parameter given by Preston *et al*. The third and final step combines the attributable fractions and the observed number of deaths from lung cancer and from other causes of death in a weighted average to produce the overall fraction of deaths attributable to smoking.

Step 1 - Estimating mortality from causes other than lung cancer (M_o)

Annual mortality (M_o) is log-linear a function of lung cancer mortality (M_L) and other covariates:

$$\ln M_O = \beta_a X_a + \beta_t X_t + \beta_c X_c + \beta_{ct}(t \times X_c) + \beta_L M_L + \beta_{Lt}(M_L \times t) + \beta_{La}(M_L \times X_a) \qquad (1)$$

Where:

M_O is the death rate from causes other than lung cancer classified by age, sex, year of death, and country (or population);

X_a is a set of dummy variables for each age group;

X_t is a set of dummy variables for each calendar year;

X_c is a set of dummy variables for each country;

$(t \times X_c)$ denotes a set of interactions between calendar year (linear) and each country dummy;

M_L is the observed death rate from lung cancer;

$(M_L \times t)$ is an interaction between M_L and year (linear);

$(M_L \times X_a)$ represents M_L interacted with the age dummies.

M_o is estimated:

1. For males and female separately
2. For 5y age groups

This step is present only in the original development of the method and does not need to be applied.

Step 2 – Estimating the AF of lung cancer due to smoking (A_L)

This is done separately for each country-year-sex-age-socioeconomic group using the formula:

$$A_L = \frac{M_L - \lambda_L^N}{M_L} \qquad (2)$$

Where M_L is the observed lung cancer death rate (***from observations in our data***)
λ^N_L is the expected rate among non-smokers (***from CPS-II study***)
When $M_L - \lambda^N_L$ is negative, A_L is set to 0.

Step 3 – Estimating the AF of smoking for mortality from other causes (A_o)

This is done separately for each country-year-sex-age-socioeconomic group using the formula:

$$A_O = 1 - e^{-\beta'_L (M_L - \lambda_L^N)} \tag{3}$$

Where M_L is the observed lung cancer death rate (***from observations in our data***)
λ^N_L is the expected rate among non-smokers (***from CPS-II study***)
β_L is a parameter measuring the impact of smoking on other causes of death (***from Preston et al***).

Step 4 – Estimating the overall AF of deaths from all causes (A)

This is done separately for each country-year-sex-age-socioeconomic group using the formula:

$$A = \frac{A_L D_L + A_O D_O}{D} \tag{4}$$

Where: D_L is the observed number of deaths from lung cancer (***from observations in our data***)
D_O is the observed number of deaths from other causes (***from observations in our data***)
D is the observed number of deaths from all causes combined (***from observations in our data.***