Supplemental Information (S)

Novel Regression Equations Incorporating Aging-specific Contributions of Various Explanatory Variables in Predicting Spirometric Parameters in Non-Obese, Japanese Adults

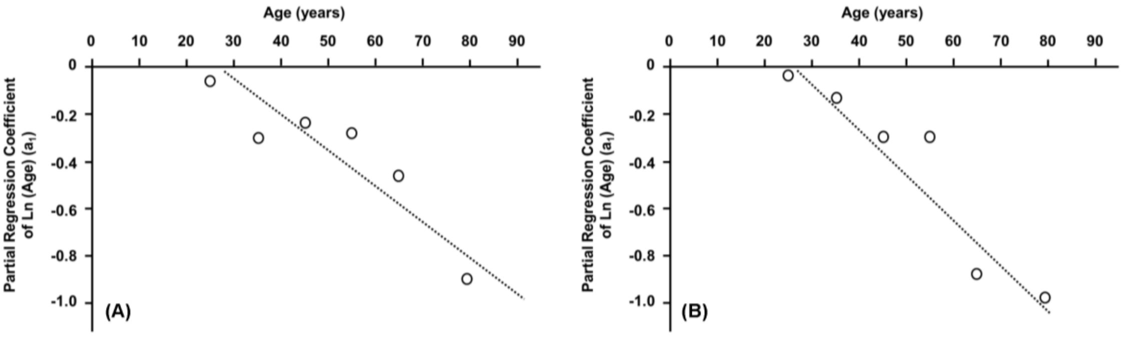
S-1) Determination of a function pertinent to describing an aging-dependent effect of an explanatory variable on a spirometric parameter

To determine the function that allows for describing the aging-dependent change in the contribution of explanatory variable to the prediction of the spirometric parameter (SP), we estimated the partial regression coefficients of a_i (i=1 to 4) in eq. (2) described in the main body of the text from the data harvested for a large cohort of nonsmoking adults stratified at age intervals of 10 years, i.e., 20-29 years, 30-39 years, 40-49 years, 50-59 years, 60-69 years, and more than 70 years. These nonsmokers (total number of subjects was 16,919, in which males and females were 6,116 and 10,803, respectively) were selected from the participants recruited for our earlier study [S1]. Reanalyzing the data obtained from these nonsmokers, we tried to empirically decide the basic function that allows for describing an aging-dependent effect of an explanatory variable on a SP. The numbers of nonsmoking male adults included 124 individuals in the age range of 20-29 years, 1,032 individuals in the range of 30-39 years, 1,644 individuals in the range of 40-49 years, 1,649 individuals in the range of 50-59 years, 1,194 individuals in the range of 60-69 years, and 473 individuals in the range of ≥70 years. The numbers of nonsmoking female adults included 159 individuals in the age range of 20-29 years, 1,519 individuals in the range of 30-39 years, 3,032 individuals in the range of 40-49 years, 3,295 individuals in the range of 50-59 years, 2,115 individuals in the range of 60-69 years, and 683 individuals in the range of ≥70 years. In this estimation, we assumed that the a_i within a group with a narrow age range of 10 years would be constant irrespective of age in an initial approximation.

We demonstrated that the negative impact of a_1, the partial regression coefficient of Ln(age), on Ln(FEV_1) was linearly enhanced as the age of a group was advanced in both genders (S-Fig. 1), which indicates that the aging-dependent contribution of a_1 would be approximately described by a linear function of age.

S-Fig. 1 Aging-specific contribution of a_1 to FEV_1 in either gender

S-Fig. 1

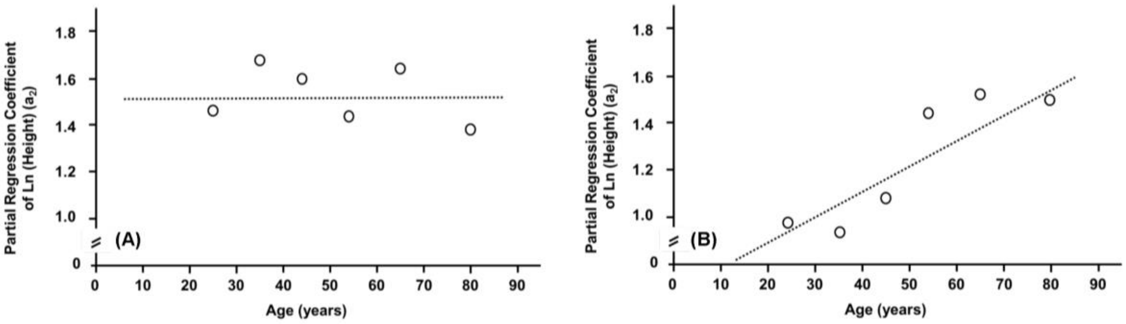


(A): males and (B): females. The a_1 values were estimated for groups stratified by age at an interval of 10 years in nonsmoking adults. In this analysis, a_1 was assumed to be constant within a group with a narrow age range of 10 years. The a_1 value was plotted against the mean age of each group, i.e., 25, 35, 45, 55, 65, and 80 years. This plot demonstrated that an aging-dependent change in a_1 would be approximated by a linear function of age irrespective of the gender, i.e., $a_1 = b_1 + c_1 \cdot \text{age}$, in which c_1 indicates a negative constant.

Regarding the contribution of a_2, the partial regression coefficient of Ln(height), to the SP, it showed a constant, positive effect on the male Ln(FVC) irrespective of age (i.e., an aging-independent contribution of a_2 to the male Ln(FVC)), while it exerted a positive, aging-dependent effect on the female Ln(FVC) in a linear fashion (S-Fig. 2). Again, in general, the aging-dependent effect of a_2 was considered to be described by a linear function of age in an initial approximation.

S-Fig. 2 Aging-specific contribution of a_2 to FVC in either gender

S-Fig. 2



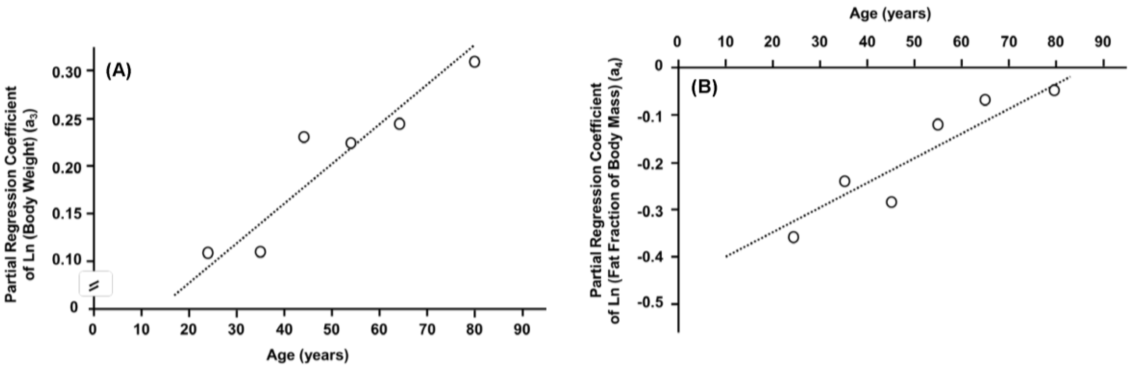
(A): males and (B): females. The groups used for estimations were the same as those defined in S-Fig. 1. The male a_2 was not altered with age, while the female a_2 linearly increased with age. These findings suggested that the a_2 value in either gender would be approximated by a linear function of age, i.e., $b_2 + c_2 \cdot \text{age}$, in which c_2 comprised a positive constant in females but zero in males.

Similar to the a_2, the a_3, which represents the partial regression coefficient of Ln(body weight), had a positive impact on the female Ln(FEF$_{50}$) in an aging-dependent linear fashion (S-Fig. 3).

The effect of a_4, which represents the partial regression coefficient of Ln(fat fraction of body mass), on the SP was manifested in a positive direction with age; thus, its negative impact on the female Ln(FEF$_{75}$) was inhibited in an aging-dependent linear fashion (S-Fig. 3).

S-Fig. 3 Aging-specific contribution of a_3 to female FEF_{50} and that of a_4 to female FEF_{75}

S-Fig. 3



The groups used for estimations were the same as those defined in S-Fig. 1. (A): the body weight had a positive impact on the female Ln(FEF_{50}), which linearly increased with age; this finding indicated that the a_3 would be described by a linear function of age as $a_3 = b_3 + c_3 \cdot age$, in which c_3 comprised a positive constant. (B): the fat fraction of body mass had a negative impact on the female Ln(FEF_{75}). However, its negative impact faded as age advanced, which indicated that the aging-dependent a_4 would be described by a linear function of age, i.e., $b_4 + c_4 \cdot age$, in which b_4 comprised a negative constant, whereas c_4 comprised a positive constant.

Joining the findings provided in S-Figs 1, 2, and 3 together, we hypothesized that the change in a_i (i=1 to 4) with aging, if any, would be approximated by the linear function of age, i.e., $a_i = b_i + c_i \cdot age$.

Reference

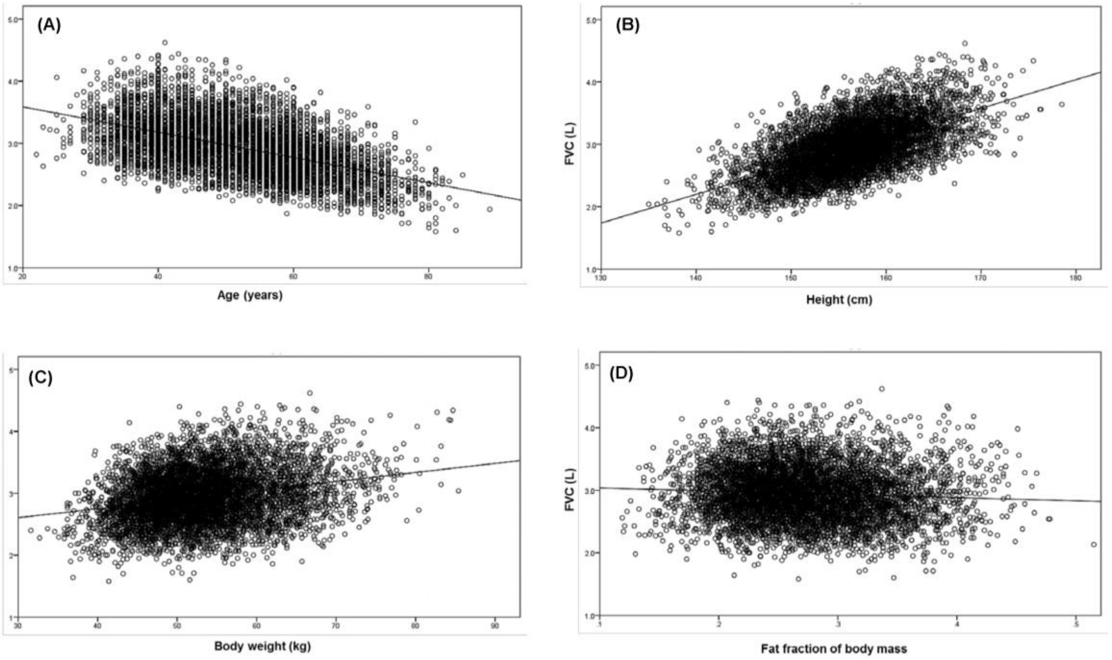
[S1] Omori H, Onoue A, Yamaguchi K (2014) A large cohort study concerning age-dependent impacts of anthropometric variables on spirometric parameters in nonsmoking healthy adults. *PLOS ONE* 9(6): e100733, doi:10.1371/journal.pone.0100733.

S-2) Correlation between measured FVC and explanatory variables in females and males

The single correlation plot of the measured SP against the explanatory variable, including age, height, body weight, or fat fraction of body mass, demonstrated that the major SPs such as FVC and FEV_1 as well as the flow parameters of PEF, FEF_{50}, and FEF_{75} decreased as age advanced while they increased with increasing height irrespective of the gender (S-Figs. 4, 5). On the other hand, the major SPs and flow parameters increased with increasing body weight but decreased with increasing fat fraction of body mass in both genders except for FEF_{75}. In the single correlation analysis, we found no statistical connection of the measured FEF_{75} to body weight in females or to fat fraction of body mass in males. However, in the partial correlation analysis, we found a significant connection of the FEF_{75} to body weight in females or to fat fraction of body mass in males (see Table 2 in the main body of the text). Collectively, these findings obviously suggest that besides age and height, body weight and fat fraction of body mass play a role in deciding most SPs in both genders.

S-Fig. 4 Measured FVC against age, height, body weight, and fat fraction of body mass in female participants

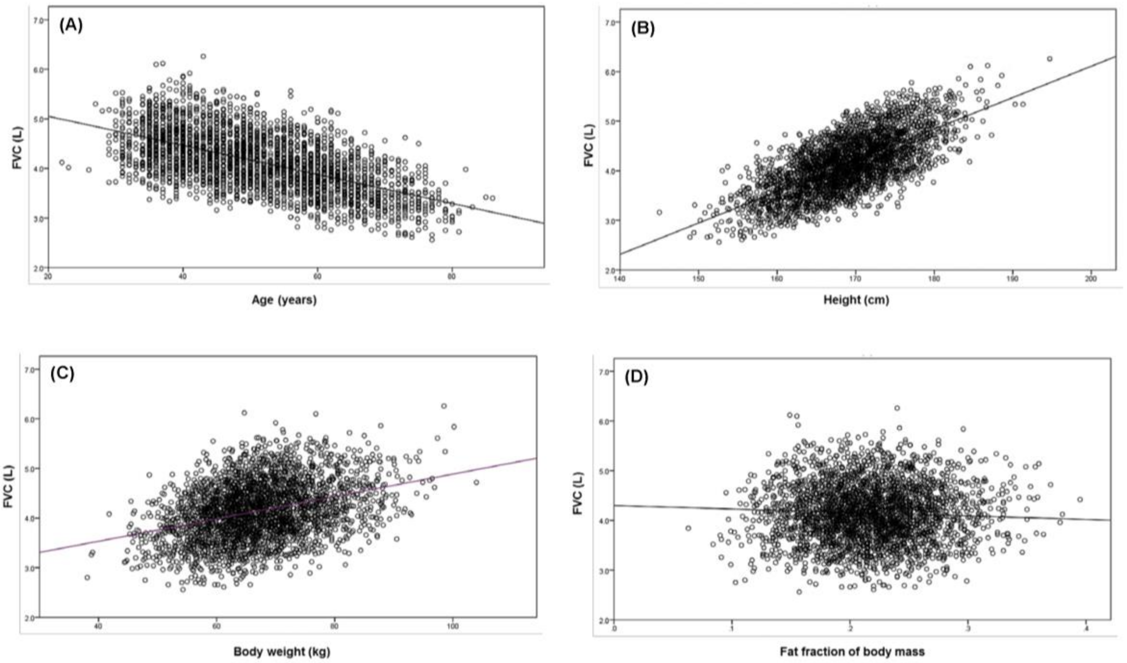
S-Fig. 4



The measured values were depicted by the form with no log-transformation. (A): FVC against age (negative slope, $p<0.0001$), (B): FVC against height (positive slope, $p<0.0001$), (C): FVC against body weight (positive slope, $p<0.0001$), and (D): FVC against fat fraction of body mass (negative slope, $p<0.0001$).

S-Fig. 5 Measured FVC against age, height, body weight, and fat fraction of body mass in male participants

S-Fig. 5



The measured values were depicted by the form with no log-transformation. (A): FVC against age (negative slope, $p<0.0001$), (B): FVC against height (positive slope, $p<0.0001$), (C): FVC against body weight (positive slope, $p<0.0001$), and (D): FVC against fat fraction of body mass (negative slope, $p=0.010$).

S-3) Residual distributions of six spirometric parameters in females and males

Since the classical least-squares minimization adopted in the present analysis belongs to the parametric procedure, it needs the approval of normal distributions of three variables, including an objective variable concerned (a spirometric measure), each explanatory variable (age, height, body weight, or fat fraction of body mass), and a residual between an objective measure and its predicted value [S2]. Among them, the most critical issue is to confirm the normality in the residual distribution, in which the distributional features of the objective measure and explanatory variables are involved.

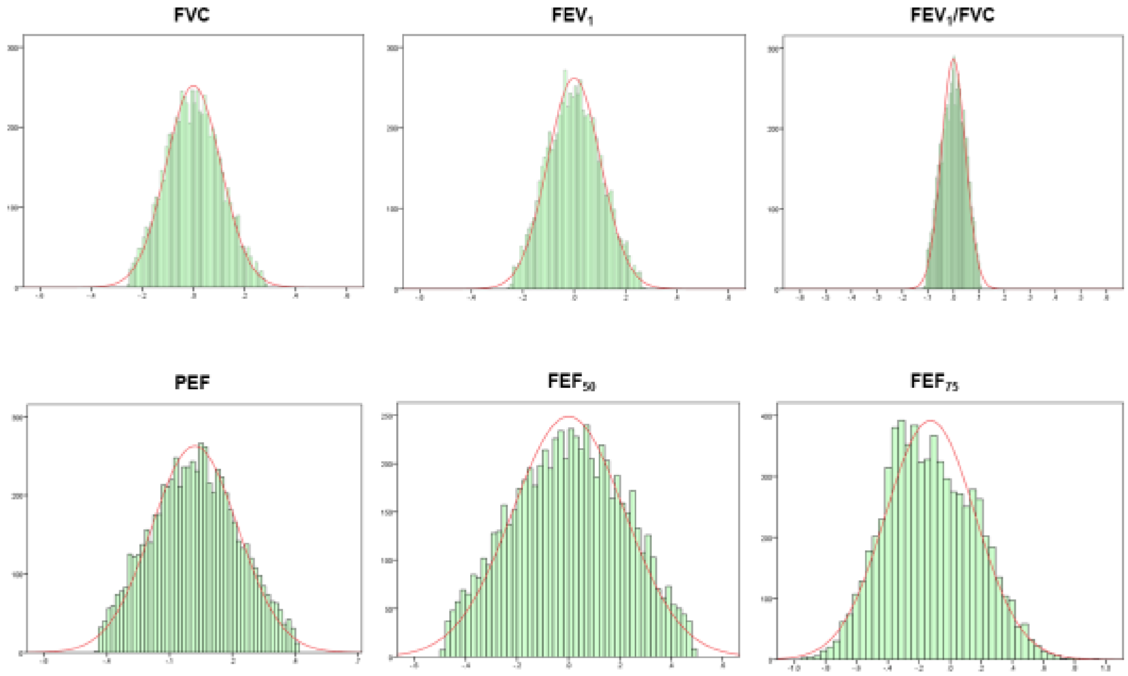
Therefore, we examined the frequency of relative residual represented by (reference mean - measured value)/(reference mean) for each spirometric parameter in either gender (S-Figs. 6, 7). We found the following matters from the frequency plots. (1) In both genders, the dispersion of relative residuals for the major spirometric parameters, including FVC, FEV_1, and FEV_1/FVC, were much smaller than that for the flow parameters, including PEF, FEF_{50}, and FEF_{75}. These findings suggest that the normal ranges of the major spirometric parameters are narrow but those of the flow parameters are wide. (2) In both genders, the residuals of the major spirometric parameters followed the normal distribution in a first approximation, while those of the flow parameters, particularly the female FEF_{75}, deviated from the normal distribution to some degree. These facts indeed indicate that the effect of skewedness (i.e., the disparity from the normal distribution) on the decision of regression equations for the major spirometric parameters is minimal but that for the flow parameters cannot be disregarded. Thus, we concluded that the classical least-squares minimization can be applied for creating the regression equation predicting the reference mean and LLN of the major spirometric parameter, including the FVC, FEV_1, or FEV_1/FVC. However, the reliability of regression equation for the flow parameter, particularly for the female FEF_{75}, is low.

Reference

[S2] Tsushima E (2008) Multivariate Approaches for Medical Data Based on SPSS. Tokyo Tosho Co.

S-Fig. 6 Frequency distributions of relative residuals of six spirometric parameters in females

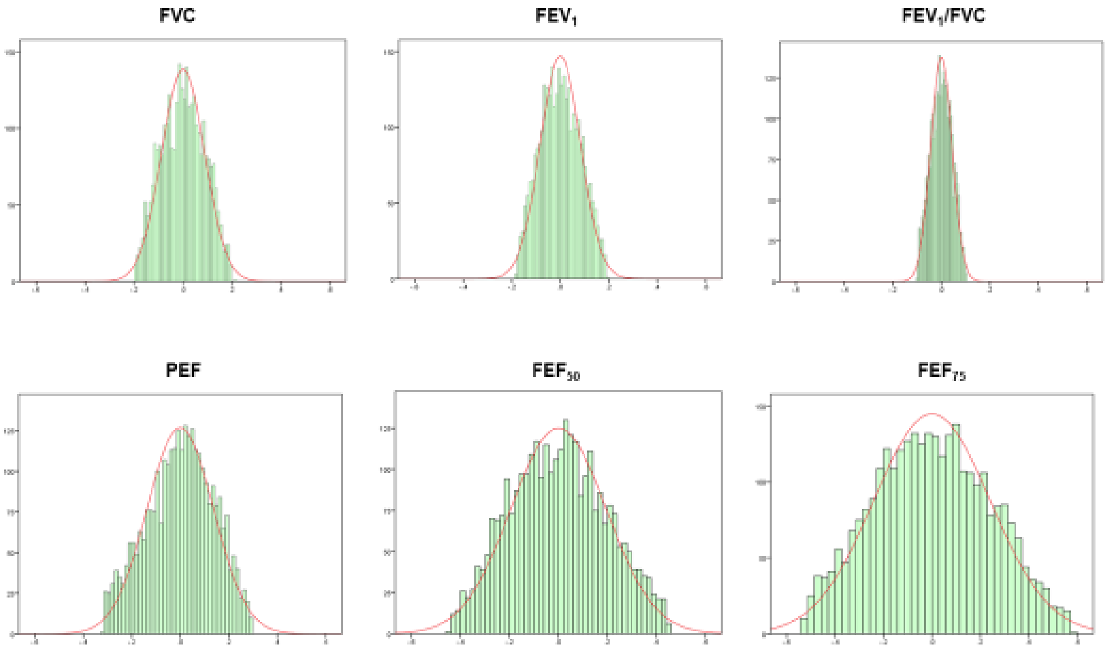
S-Fig. 6



Abscissa: relative residual defined as the ratio of residual to reference mean (residual/reference mean). Vertical axis: frequency with the same relative residual. Red line: expected curve for normal distribution.

S-Fig. 7 Frequency distributions of relative residuals of six spirometric parameters in males

S-Fig. 7



Abscissa: relative residual defined as the ratio of residual to reference mean (residual/reference mean). Vertical axis: frequency with the same relative residual. Red line: expected curve for normal distribution