A simple method for correction of the systematic error in calculating biological age by the multiple regression equation

Anatoly V. Pisaruk

State Institution « D.F. Chebotarev Institute of Gerontology NAMS of Ukraine», 04114, Kiev

Correspondence: avpisaruk54@gmail.com

Received: 8.04.2021; Accepted: 26.04.2021; Published: 26.04.2021

Abstract. There was proposed a simple method of the correction of systematic error of the equation of multiple regression while calculating biological age. The principle of this method is in the vertical transfer of points on the graph "predicted age - chronological age" on a diagonal line. At the same time, the distances from the points to the regression line are maintained. This correction eliminates the systematic error of the multiple regression equation and increases the accuracy of the biological age model.

Keywords: biological age; multiple regression equation; error correction

The calculation of biological age (BA) has been used for a long time to assess the rate of aging and to study the influence of various factors on the aging process [1]. The assessment of BA and the rate of aging is becoming especially important in connection with the need of testing of the new drugs with geroprotective activity. Various mathematical methods are currently used for the assessment of BA. All of them are based on the use of a number of biomarkers in order to calculate the value of the predicted or, as it is called, biological age. The classical method is the calculation of the multiple regression (multiple linear regression, MLR) equation [2-6]. However, this method has a significant disadvantage. When calculating age with the help of MLR based on the values of indicators that are specific to a particular person, a systematic error occurs. At the figure 1 there is an example of such a calculation. It can be seen that for the young people the estimated age (PA) is significantly higher than their chronological age (CA), and for the old people it is lower. In this case, the regression line (Fig. 1) is shifted upward at the origin and passes at an angle less than 45°. The quantity of the age calculation error, defined as PA-CA, is highly correlated with CA (Fig. 2). This error is a mathematical artifact [7]. To avoid the MLR error, other methods for calculating BA are used [8]. The best known are the principal components analysis method (PCA) and the Clemera-Dubal method (KDM). We offer our own simple way of the correction of the BA calculation error. It applies to the BA model based on MLR.

Figure 1. Correlation between chronological and predicted ages.

Figure 2. Dependence of the age calculation error on the chronological age.

To eliminate the MLR systematic error, the points need to be grouped around a line passing through the origin at the angle of 45 degrees. This line satisfies the condition: $Y = X$. The idea of the correction of the age calculation error is to move the points vertically so that their original distances to the regression line (PA vs CA) are preserved (Fig. 3). In this case, the points should be grouped around the diagonal line and their distances to this line should be equal to the original distances.

Figure 3. The principle of transferring points to correct the error of the regression method.

The transformation algorithm is as follows. First, the vertical distance (L) from each point to the regression line is calculated (dash and dash line in Fig. 3).

$$L = PA - Y; \quad Y = k_1 + k_2 CA$$

where: Y is the age calculated by the regression equation at a certain value of CA (Fig. 1), and k_1 and k_2 are the coefficients of this equation.

$$\text{In our example: } Y = 33,96 + 0,385 CA$$

Then we calculate the PA value after correction (PAc) using the following formula:

$$PAc = CA + (PA - Y) \text{ или } PAc = PA + CA(1 - k_2) - k_1$$

Figure 4. Correlation of chronological and predicted ages after correction.

Figure 5. Dependence of the age calculation error on the chronological age (after correction).

Graphically, this is the equivalent to the vertical moving points on the graph. In this case, the points of young people will move down, and the points of old people - up. As a result, the systematic error in the calculating of the age is eliminated, but the original scatter of points is preserved. In fig. 4 the points corresponding to the estimated age of people after correction are represented. As you can see from figure 4, the correlation of estimated age with CA for our example becomes much higher (0.899 vs 0.620), and the standard error of the calculating of the age decreases to 8.9 years, in comparison with 14.4 years before correction. After applying the age calculation correction method, the estimated age error ceases to correlate with CA (Fig. 5). This indicates about the adequacy of the proposed correction method.

Thus, the proposed method for the correcting of the systematic error in calculating BA completely eliminates this error and makes it possible to create more accurate BA models.

Information about Author:

Anatoly V. Pisaruk – DSc (Medicine), Head of the Laboratory for Mathematical Modeling of Aging Processes; <https://orcid.org/0000-0001-5522-0172>

References

1. Anstey, K.J.; Lord, S.R.; Smith, G.A. Measuring human functional age: a review of empirical findings. *Exp Aging Res* **1996**, *22*, pp 245–266. doi: 10.1080/03610739608254010
2. Aykroyd, R.G.; Lucy, D.; Pollard, A.M.; Solheim, T. Technical note: regression analysis in adult age estimation. *Am J Phys Antropol* **1997**, *104*, pp 259–265. doi: 10.1002/(SICI)1096-8644(199710)104:2<259::AID-AJPA11>3.0.CO;2-Z
3. Krøll, J.; Saxtrup, O. On the use of regression analysis for the estimation of human biological age. *Biogerontology* **2000**, *1*, pp 363–368. doi: 10.1023/a:1026594602252
4. Kumari, K.; Yadav, S. Linear regression analysis study - Journal of the practice of Cardiovascular. *Sciences* **2018**, *4*, pp 33-36. doi: 10.4103/jpcs.jpcs_8_18
5. Dubina, T.L.; Dyundikova, V.A.; Zhuk, E.V. Biological age and its estimation. II. Assessment of biological age of albino rats by multiple regression analysis. *Exp Gerontol* **1983**, *18*, pp 5–18. doi: 10.1016/0531-5565(83)90046-3
6. Dubina, T.L.; Mints, A.Y.; Zhuk, E.V. Biological age and its estimation. III. Introduction of a correction to the multiple regression model of biological age in cross-sectional and longitudinal studies. *Exp Gerontol* **1984**, *19*, pp 133–143. doi: 10.1016/0531-5565(84)90016-0
7. Krut'ko, V.N.; Smirnova, T.M.; Dontsov, V.I.; Borisov, S.E. Diagnosing aging: I. Problem of reliability of linear regression models of biological age. *Human Physiol* **2001**, *27*, pp 725–731.
8. Klemra, P.; Doubal, S. A new approach to the concept and computation of biological age. *Mech Ageing Dev* **2006**, *127*, pp 240–248. doi: 10.1016/j.mad.2005.10.004