A Simplified Equation for Adult BMI Growth, and Its Use to Adjust BMI for Age


by

Paul T. Von Hippel, Ph.D.
LBJ School of Public Affairs
University of Texas-Austin
paulvonhippel@utaustin.utexas.edu

Jamie L. Lynch Ph.D.
Department of Sociology
St. Norbert College
Jamie.lynch@snc.edu

*Word count:* 1,153 words excluding References, Figures, Tables, and this title page

*Funding:* None
Although standards for children’s growth in body mass index (BMI) are widely used (1), standards for BMI growth in adulthood are less developed. A year ago in this journal, Østbye, Malhotra, and Landerman (2) made an important contribution to the study of adult BMI growth by describing average BMI growth curves for four latent groups of US adults followed from age 18 to 49 years in the National Longitudinal Survey of Youth, 1979 cohort (NLSY79). Participants in the NLSY79 were born between 1957 and 1965.

In this letter we show that the average growth curves of all four groups can be summarized using one equation with a single parameter. We then use that equation to put the BMIs of adults measured at different ages on a common, age-adjusted scale. Finally, we illustrate how age-adjusted BMIs can be used in epidemiology.

The article by Østbye, Malhotra, and Landerman (2) offer a quadratic approximation for BMI growth from age 18 to age 49. Average or expected BMI at age \( t \) is modeled as

\[
E(BMI|t) = \alpha_0 + \alpha_1 (t - 18) + \alpha_2 (t - 18)^2
\]  

(1)

The authors identify four latent groups in the NLSY97, and estimate the parameters \( \alpha_0, \alpha_1, \alpha_2 \) for each group. Table 1 reproduces their results. Note that each group contains members of both genders and several races and ethnicities.

While the four groups are quite different, we wish to point out a simple relationship among their parameters: groups with a large intercept also have a large linear coefficient and a large negative quadratic coefficient. Across groups, the correlation between the intercept and the linear coefficient is .999, and the correlation between the intercept and the quadratic coefficient is –.993. What this means is that the linear and quadratic coefficients can be very well
approximated as linear functions of the intercept. Using the least-squares criterion, the best linear approximations are

\[
\hat{a}_1 \approx -1.93536 + 0.10326 \hat{a}_0
\]

\[
\hat{a}_2 \approx +0.04288 - 0.002123 \hat{a}_0
\] (2)

(Note that these linear approximations are estimates, since they are based on the estimated quantities \(\hat{a}_0, \hat{a}_1, \) and \(\hat{a}_2).\)

Substituting equation (2) into equation (1), we get an equation that describes the average growth curves of all four groups in terms of a single parameter \(\alpha_0\)

\[
E(BMI|t) = \alpha_0 + (-1.93536 + 0.10326\hat{a}_0)(t - 18) + (0.04288 - 0.002123\hat{a}_0)(t - 18)^2
\] (3)

By substituting each group’s intercept (\(\hat{a}_0 = 19.75, 21.89, 24.45, 28.83\)) into the one-parameter equation (3), we can approximate the average growth curve of each group. The result, in our Figure 1, matches closely the Figure 1 that the authors produced from the three-parameter equation (1).

The one-parameter model allows us to do more than reproduce the average curve of each group. It also allows us to interpolate between the group curves. For example, we might want to predict the average curve for adults who had a BMI of \(\alpha_0 = 27\) at age 18. Using equation (3) we would generate a curve that fits between the curve for the LO group (\(\alpha_0 = 24.45\)) and the EO group (\(\alpha_0 = 28.83\)). Figure 2 displays predicted average curves starting at age \(t=18\) with every intercept from \(\alpha_0 = 18\) to \(\alpha_0 = 30\).
The one-parameter model also lets us do something that can be useful in epidemiology: it lets us adjust adult BMI for age. There are occasions when it is useful to compare the BMIs of adults with different ages, and have the comparison approximate what would be observed if the adults were all measured at the same age. To calculate an age-adjusted BMI, start with an adult’s observed $BMI_{obs}$ at an observed age $t_{obs}$. Then use $BMI_{obs}$ and $t_{obs}$ to estimate $\alpha_0$ from equation (4) below, which is obtained by solving equation (3) for $\alpha_0$:

$$\alpha_0 = \frac{-1.93536(t_{obs} - 18) + 0.04288(t_{obs} - 18)^2 - BMI_{obs}}{-1 - 0.10326(t_{obs} - 18) + 0.002123(t_{obs} - 18)^2}$$

(4)

Finally, use the estimated $\alpha_0$ to predict from equation (5) below (which is just equation (3) with subscripts) what “adjusted” $BMI_{adj}$ the adult would be expected to have at a target age $t_{adj}$.

$$E(BMI_{adj}|t) = \alpha_0 + (-1.93536 + 0.10326\alpha_0)(t_{adj} - 18) + (0.04288 - 0.002123\alpha_0)(t_{adj} - 18)^2$$

(5)

The approximation is likely to work best if the adjusted age $t_{adj}$ is close to the observed age $t_{obs}$, but adjustment can also be attempted if $t_{adj}$ is further from $t_{obs}$.

To validate our age-adjustment method in a different sample, we considered the mothers of participants in the 1997 cohort of the National Longitudinal Survey of Youth (NLSY97). Although NLSY97 participants were a generation younger than participants in the NLSY79, the mothers of NLSY97 participants were approximately the same age as participants in the NLSY79. If our age adjustment method works for the NLSY79, the same adjustment should work for the mothers of the NLSY97.

In the NLSY97 we compared the BMI distribution of female participants and their mothers. This comparison requires age adjustment because NLSY97 participants and their mothers do not report BMI at the same ages. NLSY97 participants reported their BMI every year.
until 2007, when the oldest NLSY97 participants were 27, but their mothers reported BMI only in the first year of the survey, 1997, when the oldest NLSY97 participants were 16 and their mothers’ ages were 39.9 years, on average.

At age 27, female NLSY97 participants had a mean BMI of 27.0, which is the same as the mean of the unadjusted BMIs that their mothers reported 11 years earlier. However, if we adjust the mothers’ BMIs for age, we find that the mean of the mothers’ BMIs, adjusted to age 27, was just 24.2—fully 2.8 kg/m^2 lower than their daughters’ average BMI at the same age. This age-adjusted result comes close to estimates from the National Health and Nutrition Examination Survey (NHANES), which found that the mean BMI of US women aged 20-29 rose from 24.3 in 1988 to 26.8 in 2002 (3).

Age adjustment slightly increases the correlation between the BMIs of mothers and daughters at age 27. The correlation is .31 if mothers’ BMIs are unadjusted, and .32 if mothers BMIs are adjusted to age 27. Age adjustment also increases the correlation between the BMIs of mothers and sons at 27, from .33 without adjustment to .34 with adjustment. These increases in correlation are a good sign; they suggest that the age-adjustment method, although approximate and therefore prone to error, does more good than harm even when t_{adj} is 13 years, on average, from t_{obs}.

The formulas derived here are based on the NLSY79 and are appropriate for Americans of that cohort, including the mothers of the NLSY97. The equations are appropriate through age 49, but may not be appropriate for later ages that were not observed in the NLSY79.

Similar equations could be developed for earlier or later cohorts by analyzing earlier surveys or later surveys such as the NLSY97. The analyses could be carried out like those in this letter, or they could be carried out more simply, by fitting a multilevel growth model (4,5). In a
multilevel growth model, we would fit all of the data to equation (1) and allow the parameters $\alpha_0$, $\alpha_1$, and $\alpha_2$ to vary as correlated random effects across individuals. Given the results in this letter, we would expect $\alpha_0$ and $\alpha_1$ to be positively correlated with each other and negatively correlated with $\alpha_2$. 
Reference List


### Tables

<table>
<thead>
<tr>
<th>Group</th>
<th>% of cohort</th>
<th>Parameter estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Intercept ($\alpha_0$)</td>
</tr>
<tr>
<td>Early adulthood obesity (EO)</td>
<td>4.2%</td>
<td>28.83</td>
</tr>
<tr>
<td>Late adulthood obesity (LO)</td>
<td>19.7%</td>
<td>24.45</td>
</tr>
<tr>
<td>Overweight (OW)</td>
<td>41.2%</td>
<td>21.89</td>
</tr>
<tr>
<td>Normal weight (NW)</td>
<td>35.0%</td>
<td>19.75</td>
</tr>
</tbody>
</table>

Table 1. Average parameters for the four latent groups.
Figure 1. Average curves for all four latent groups, regenerated from the single-parameter equation (3).
Figure 2. Predicted average curves for adults whose BMIs are between 18 and 30 at age 18.