COMPUTATION OF CENTILES AND Z-SCORES FOR HEIGHT-FOR-AGE, WEIGHT-FOR-AGE AND BMI-FOR-AGE

The method used to construct the 2007 WHO references relied on GAMLSS with the Box-Cox power exponential distribution (Rigby and Stasinopoulos, 2004). However, the final selected models simplified to the LMS model (Cole and Green, 1992) since none of the references required adjustment for kurtosis. As a result, the computation of percentiles and z-scores for all three indicators uses formulae based on the LMS method. However, a restriction was imposed on all indicators to enable the derivation of percentiles only within the interval corresponding to z-scores between -3 and 3. The underlying reasoning is that percentiles beyond ±3 SD are invariant to changes in equivalent z-scores. The loss accruing to this restriction is small since the inclusion range corresponds to the 0.135th to 99.865th percentiles.

For all indicators, the tabulated fitted values of Box-Cox power, median and coefficient of variation corresponding to age (or height) *t* are denoted by *L(t)*, *M(t)* and *S(t)*, respectively.

Centiles and z-scores for height-for-age

For this indicator, *L(t)* is equal to 1, simplifying the Box-Cox normal distribution used in the LMS method to the normal distribution. Therefore, differences between adjacent standard deviations (e.g. between 2 SD and 3 SD) were constant for a specific age but varied at different ages.

In this case, the centiles at age *t* can be estimated from:

$$\begin{aligned} C_{100\alpha}(t) &= M(t)\left[1 + L(t)S(t)Z_\alpha\right]^{1/L(t)} = M(t)\left[1 + S(t)Z_\alpha\right] \\ &= M(t) + StDev(t)Z_\alpha, \qquad -3 \le Z_\alpha \le 3 \end{aligned}$$

where Z_α is the normal equivalent deviate for tail area α, C100α(t) is the 100α-th centile, and *StDev(t)* is the standard deviation at age *t* (derived from multiplying *S(t)* by *M(t)*).

The individual z-score for a measurement *y* at age *t* was computed as:

$$z_{ind} = \frac{\left[y / M(t) \right]^{L(t)} - 1}{S(t)L(t)} = \frac{y - M(t)}{StDev(t)}$$

Centiles and z-scores for weight-for-age and BMI-for-age

The weight-based indicators presented right-skewed distributions. When modelled correctly, right skewness in data has the effect of making distances between positive z-scores increase progressively the farther away they are from the median, while distances between negative z-scores decrease progressively. The LMS method fits skewed data adequately by using a Box-Cox normal distribution, which follows the empirical data closely. The drawback, however, is that the outer tails of the distribution are highly affected by extreme data points even if only very few (e.g. less than 1%). Following the same methodology applied to the WHO Child Growth Standards, a restricted application of the LMS method was thus used for the 2007 WHO weight-based indicators, limiting the Box-Cox normal distribution to the interval corresponding to z-scores where empirical data were available (i.e. between -3 SD and 3 SD). Beyond these limits, the standard deviation at each age was fixed to the distance between ±2 SD and ±3 SD, respectively. This approach avoids making assumptions about the distribution of data beyond the limits of the observed values (WHO Multicentre Growth Reference Study Group, 2006).

As a result of this adjustment, the z-score distribution can depart slightly from normality at the extreme tails (beyond ±3 SD), although the expected practical impact is minimal.

The centiles were calculated as follows:

$$C_{100\alpha}(t) = M(t)\left[1 + L(t)S(t)Z_{\alpha}\right]^{1/L(t)}, \qquad -3 \le Z_{\alpha} \le 3$$

The following procedure is recommended to calculate a z-score for an individual child with measurement *y* at age *t*:

1. Calculate

$$z_{ind} = \frac{\left[y/M(t) \right]^{L(t)} - 1}{S(t)L(t)}$$

2. Compute the final z-score $\left(z^{*}_{ind}\right)$ of the child for that indicator as:

$$z^{*}_{ind} = \begin{cases} z_{ind} & if \quad |z_{ind}| \le 3 \\ 3 + \left(\dfrac{y - SD3pos}{SD23pos} \right) & if \quad z_{ind} > 3 \\ -3 + \left(\dfrac{y - SD3neg}{SD23neg} \right) & if \quad z_{ind} < -3 \end{cases}$$

where

$SD3pos$ is the cut-off 3 SD calculated at *t* by the LMS method:

$$SD3pos = M(t)\left[1 + L(t) * S(t) * (3)\right]^{1/L(t)};$$

$SD3neg$ is the cut-off -3 SD calculated at *t* by the LMS method:

$$SD3neg = M(t)\left[1 + L(t) * S(t) * (-3)\right]^{1/L(t)};$$

$SD23pos$ is the difference between the cut-offs 3 SD and 2 SD calculated at *t* by the LMS method:

$$SD23pos = M(t)\left[1 + L(t) * S(t) * (3)\right]^{1/L(t)} - M(t)\left[1 + L(t) * S(t) * (2)\right]^{1/L(t)};$$

and $SD23neg$ is the difference between the cut-offs -2 SD and -3 SD calculated at *t* by the LMS method:

$$SD23neg = M(t)\left[1 + L(t) * S(t) * (-2)\right]^{1/L(t)} - M(t)\left[1 + L(t) * S(t) * (-3)\right]^{1/L(t)}$$

To illustrate the procedure, an example with BMI-for-age for boys is provided below and displayed in Figure 1.

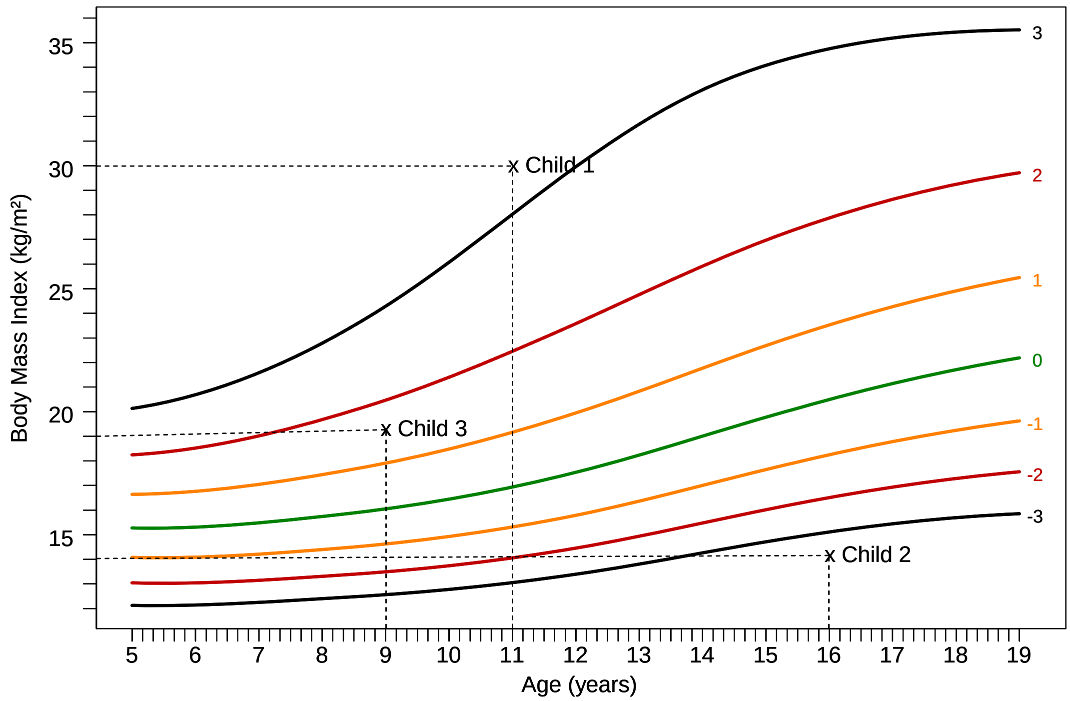

Figure 1 Examples of children/adolescents ranked according to the 2007 WHO BMI-for-age reference

Child 1: 11 year-old boy with BMI=30.

L=-1.7862; M=16.9392; S=0.11070;

$$z_{ind} = \frac{\left[30.0/16.9392\right]^{(-1.7862)} - 1}{0.11070 * (-1.7862)} = 3.24 \quad \mathbf{>3}$$

$$SD3pos = 16.9392 * \left[1 + (-1.7862) * 0.11070 * (3)\right]^{1/(-1.7862)} = 28.03$$
$$SD2pos = 16.9392 * \left[1 + (-1.7862) * 0.11070 * (2)\right]^{1/(-1.7862)} = 22.45$$
$$SD23pos = 28.03 - 22.45 = 5.58$$
$$\Rightarrow \quad z^{*}_{ind} = 3 + \left(\frac{30 - 28.03}{5.58}\right) = 3.35$$

Child 2: 16 year-old boy with BMI=14.

L=-1.3529; M=20.4951; S=0.12579;

$$z_{ind} = \frac{\left[14.0/20.4951\right]^{(-1.3529)} - 1}{0.12579*(-1.3529)} = -3.96$$ **<-3**

$$SD2neg = 20.4951*\left[1+(-1.3529)*0.12579*(-2)\right]^{1/(-1.3529)} = 16.50$$
$$SD3neg = 20.4951*\left[1+(-1.3529)*0.12579*(-3)\right]^{1/(-1.3529)} = 15.11$$
$$SD23neg = 16.50 - 15.11 = 1.39$$
$$\Rightarrow \quad z^*_{ind} = -3 + \left(\frac{14.0-15.11}{1.39}\right) = -3.80$$

Child 3: 9 year-old boy with BMI=19

L=-1.6318; M=16.0490; S=0.10038;

$$z_{ind} = \frac{\left[19/16.0490\right]^{(-1.6318)} - 1}{0.10038*(-1.6318)} = 1.47$$ **≥-3 and ≤3 (LMS z-score)**

Bibliography

Cole TJ, Green PJ (1992). Smoothing reference centile curves: the LMS method and penalized likelihood. *Statistics in Medicine*, 11:1305–1319.

Rigby RA, Stasinopoulos DM (2004). Smooth centile curves for skew and kurtotic data modelled using the Box-Cox power exponential distribution. *Statistics in Medicine*, 23:3053–3076.

WHO Multicentre Growth Reference Study Group (2006). WHO Child Growth Standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development. Geneva: World Health Organization; pp 312.