Measures of body mass and of obesity from infancy to adulthood and their appropriate transformation

Th. Gasser,1 P. Ziegler,1 B. Seifert1, A. Prader2, L. Molinari3 and R. Largo4
1Universität Zürich
2Zentralinstitut für Seelische Gesundheit
3Kinderspital Zürich

Received 17 July 1992; revised 5 March 1993

Summary. In this paper we investigate first the choice of an appropriate index of body mass. The traditional indices weight/height (W/H), W/H² and W/H³ are compared, as well as an approach due to Cole (1986) making the index W/Hᵖ age-dependent, i.e. by allowing the power p to depend on age. While there may be no perfect index reflecting over- and underweight—irrespective of height and width—the Quetelet index W/H² turned out to be a reasonable index from childhood to adulthood. Second, we study the development of the body mass index W/Hᵖ longitudinally from birth onwards, based on structural average distance, velocity and acceleration curves. As a third topic, transformations for obtaining an approximate normal distribution for W/H², weight and skinfolds are compared. The usual log-transformation turned out to be unsatisfactory. While for height and arm skinfolds a transformation −1/ₓ performs rather well across age and sex, a transformation −1/₁ₓ is more appropriate for trunk skinfolds and W/H².

1. Introduction

Weight, skinfolds and circumferences are all related to body mass, fat and over- or underweight. Except for skinfolds they are composite measures, reflecting bone, fat and muscle size. While there are direct methods to estimate total body fat content, skinfolds seem to relate to this quantity reasonably well (Durnin and Rahaman 1967, Forbes and Amirhakimi 1970), despite a substantial inter- and intraobserver variability in these measurements. When using weight to assess body mass, the influence of height has to be accounted for in some appropriate way. A simple and popular approach is to compute indices of the form W/Hᵖ (W = weight, H = height, p = 1,2,3), the body mass index or Quetelet index W/H² being the most popular. A more refined method has been proposed by Cole (1986): for each age, a suitable power p is estimated such that weight relates linearly to height to the power p, resulting in an age-dependent type of body mass index. A bell-shaped function was found to be appropriate for modelling the dependence of the power p on age (Cole 1986).

The risk attributed to overweight for cardiovascular diseases, diabetes, etc. makes the choice of an index of body mass, and the understanding of its properties, a problem of great practical importance. While the validity of such an index has the highest priority, it should be as simple as possible. Its validity throughout development well into adulthood is a particular problem: for comparisons and predictions in scientific investigations it would be desirable to use the same index even if not entirely adequate at all stages of development. Some of the features that we require of such an index are:

(1) It should correlate well with weight.
(2) It should correlate well with skinfolds.
(3) It should by-and-large be uncorrelated with weight.
(4) Its dependence on age and sex should not be too accentuated or too complicated to be taken into account.

(5) It should follow rather closely a normal distribution, or allow a straightforward transformation to an approximately normally distributed variate.

In the work of Rolland-Cachera, Sempé, Guillonl-Bataille, Patois, Péquignot-Guggenbühl and Fautrad (1982) the classical index $W/H^2$ turned out to be best, and they constructed age-related norms over the whole age range (Rolland-Cachera, Deheeger, Guillon-Bataille, Avons, Patois and Sempé 1987). In an investigation including other approaches such as deviations from weight for height regression in narrow age ranges—a less direct and less easy procedure—the $W/H^2$ body mass index also fared well (Spyckerelle, Gueguen, Guilleenot, Tosi and Deschamps 1988). The problem of the approximation to the normal distribution leads directly to the problem of finding an appropriate transformation to achieve this goal. For the body mass index, as well as for weight and skinfolds, a logarithmic transformation has been common. Assessing good transformations for skinfolds across age Van't Hof, Wit and Roede (1985) found transformations close to taking reciprocal values superior, in particular for subscapular and suprailiac skinfolds.

The present investigation is based on the data of the first Zürich longitudinal growth study, which constitutes a representative sample of normal children. However, the lessons to be learnt should also be useful for clinically oriented studies. While there are good reasons to pursue sophisticated approaches, we want to evaluate and compare simple indices, and prefer a single transformation across age for boys and girls rather than age-dependent normalizing transformations. This requires some compromise.

2. Subjects and methods

2.1. Subjects and measurements

In an internationally coordinated study (Falkner 1960), initiated in 1954, participation of 321 children could be obtained for the Zürich sample. In this investigation, $n = 112$ girls and $n = 120$ boys, with rather complete measurements from infancy to adulthood, enter into the analysis. For further details the reader may consult the papers by Gasser, Kneip, Ziegler, Largo and Prader (1990) or Prader, Largo, Molinari and Issler (1989).

Measurements were taken at 1, 3, 6, 9, 12, 18 and 24 months and then annually until age 9 for girls and age 10 for boys; then half-year intervals started. This was again changed to yearly measurements when the annual increment in height was less than 0.5 cm.

2.2. Statistical methods

The different indices related to weight, i.e. $W/H$, $W/H^2$ and $W/H^3$, were compared by computing cross-sectional mean distance curves, and structural average curves for distance, velocity and acceleration. Structural average curves are obtained by first synchronizing individual curves to a common maturational age scale prior to averaging the individual curves (for details see Gasser et al. 1990). Further, rank correlations with height, weight, skinfolds, circumferences and bhumeral and biliac width were obtained throughout development. Following Cole (1986) we also evaluated the construction of an age-dependent and possibly sex-dependent relation $W/H^p$, where the power $p$ depends on age and possibly on sex. According to Cole the
function \( p(\text{age}) \) can be well modelled as a bell-shaped curve as \( p(\text{age}) = c_1 + c_2 \exp(c_3(\text{age} - c_4)^2) \).

Since weight, the body-mass index \( W/H^2 \) and skinfolds do not follow a normal distribution, transformations of the raw data were evaluated which should lead to a better approximation to the normal distribution. The family of power transformations \( y = \text{sign}(\alpha) x^\alpha \) (usually \(-1 \leq \alpha \leq 1\)) has often been effective and comprises a variety of popular transformations: for \( \alpha = 1 \) we get back the raw values, for \( \alpha = 1/2 \), \( 1/3 \) the square and the cube root transformation. For \( \alpha = 0 \) we get the logarithm in the limit, and for \( \alpha = -1 \) the reciprocal transformation. When moving downwards from \( \alpha = 1 \) to \( \alpha = -1 \), and even further, more and more skewness to the right is eliminated. Thus, mild deviations from normality need an \( \alpha \) close to 1 and strong deviations may need up to a reciprocal transformation, close to \( \alpha = -1 \). After a prior analysis across a wide range of \( \alpha \) values we decided to evaluate and compare \( \alpha = 1, 1/2, 0, -1/2, -1 \), which are about equally spaced in the world of transformations.

The log-transformation \( (\alpha = 0) \) has been popular for the variables studied here. However, the optimal transformation is expected to depend on age and perhaps also on sex, making any further statistical analysis more intricate. Therefore, our goal was to find a compromise in the form of a single transformation which produces a reasonable approximation to the normal distribution over the whole age range studied, both for boys and girls. While this may entail deviations from normality at some age, any gross deviation should be more or less eliminated by the transformation. As measures of approximation to the normal distribution we used primarily skewness and kurtosis, which should both be ideally zero. Skewness indicates deviations from symmetry, positive values indicating a thicker right tail, negative ones a thicker left tail. Kurtosis indicates whether the actual distribution has longer tails than the normal distribution (positive values) or shorter ones (negative values). Both measures may become inflated by a few outliers. The Wilk-statistic \( W \) is another measure of deviation from the normal distribution; it was inspected occasionally.

3. Results

3.1. Evaluation of indices of body mass and of obesity

Figure 1 shows cross-sectional mean curves for \( W/H, W/H^2 \) and \( W/H^3 \). Age changes and sex differences should be small for an overall index of obesity. Evidently they were most accentuated for \( W/H \). The age and sex changes of \( W/H^2 \) were rather complex compared to \( W/H^3 \), since an initial increase was followed by a decrease, with a further increase starting before adolescence, earlier for girls than for boys. For the Quetelet index \( W/H^2 \) means and standard deviations were at 4 weeks 13·74 (1·02) for boys and 13·71 (0·96) for girls, and at adulthood 21·14 (2·55) for boys and 20·60 (2·79) for girls. A comparison with cross-sectional mean curves of \( W/H^2 \) obtained by Rolland-Cachera et al. (1982) led to a remarkable agreement (figure 2). Evidently age-dependent norms are needed. Figure 3 gives structural average velocity and acceleration curves for \( W/H^2 \), aligned according to the individual tempo of weight. It demonstrates that the body mass index \( W/H^2 \) reflected not only the state of corpulence, but also developmental changes per se. As for other variables, a pubertal spurt (PS) and a mid-growth spurt (MS) could be clearly identified. The timing of the PS was roughly the same as for weight, with the delay for boys well known from other variables. However, in contrast to most other variables, boys had a less intense PS for
Figure 1. Mean distance curves for $W/H$ (above), for $W/H^2$ (middle) and for $W/H^3$ (below) for boys (dotted line) and for girls (solid line).

Figure 2. Mean distance curves for $W/H^2$ for boys (above) and girls (below) comparing our results (solid line) with those of Rolland-Cachera et al. (1982) (dotted line).
$W/H^2$ in terms of velocity and acceleration. The MS peaked at about 7·5 years in both sexes and its intensity was also independent of sex. Prior to the MS, an accentuated drop in velocity occurred in the first $1\frac{1}{2}$ years, down to distinctly negative values after 1 year. Velocity became positive again at about 6 years. The development was remarkably parallel for boys and girls until 10 years, when the PS for girls started.

Rank correlations with height should be small for a body mass index. They were lowest for $W/H^2$, not far from zero, and higher for the two other indices (figure 4). On the other hand, correlations with weight should be high. They were roughly 0·8 across age for $W/H^2$ and much lower for $W/H^3$ (figure 5). Correlations of $W/H$ with weight were above 0·95 across age so that the following results for $W/H$ hold to a good approximation also for $W$ itself. These indices correlated similarly for the two arm skinfolds, and only those for triceps skinfolds are shown in figure 6. The Quetelet index $W/H^2$ was best, but differences turned out to be small. The correlations of girls were increasing until age 9 to about 0·7-0·8, higher than those of boys, which fluctuated around 0·5. Correlations were a bit better for subscapular skinfolds as compared to suprailiac skinfolds, and are depicted in figure 7. Again the Quetelet index $W/H^2$ was slightly above the other measures, and its rank correlations were increasing until the age of 9-10 years, reaching then 0·7-0·8 for girls and 0·7 for boys.

When comparing rank correlations between $W/H^2$ and all four skinfolds, subscapular skinfolds were highest and biceps skinfolds lowest. The sum of all four skinfolds correlated often about equally with $W/H^2$ as the best single skinfold (subscapular), and at certain ages somewhat worse. Again correlations for $W/H$, $W/H^2$ and $W/H^3$ were rather close together.

To supplement the above analysis, performed according to the criteria set in the introduction, correlations with arm and calf circumferences and with biiliac and bihumeral width were computed. Arm circumference correlated high with
Figure 4. Rank correlation of $W/H$ (solid line), of $W/H^2$ (short dashes) and of $W/H^3$ (long dashes) with height for boys (above) and for girls (below).

Figure 5. Rank correlation of $W/H$ (solid line), of $W/H^2$ (short dashes) and of $W/H^3$ (long dashes) with weight for boys (above) and for girls (below).
Figure 6. Rank correlation of $W/H$ (solid line), of $W/H^2$ (short dashes) and of $W/H^3$ (long dashes) with triceps skinfolds for boys (above) and for girls (below).

Figure 7. Rank correlation of $W/H$ (solid line), of $W/H^2$ (short dashes) and of $W/H^3$ (long dashes) with subscapular skinfolds for boys (above) and for girls (below).
$W/H^2$—about 0.8 in infancy and childhood and 0.9 later on—distinctly higher than calf circumference. Biiliac and bihumeral width correlated with the Quetelet index typically around 0.3–0.4 across age (figure 8). The ratio crown–rump to subischial leg length correlated about 0.2 to 0.3 with the body mass index $W/H^2$, while the ratio bihumeral to biiliac width was approximately uncorrelated with $W/H^2$.

![Figure 8](image)

**Figure 8.** Rank correlation of $W/H^2$ with arm circumference (solid line), with biiliac diameter (short dashes) and with bihumeral diameter (long dashes) for boys (above) and for girls (below).

### 3.2. Age-dependent indices

Finally we evaluated the approach of establishing a linear relation between weight and a suitable power of height (of the form $HF^p$, $p$ age-dependent, following Cole 1986). The appropriate $p$ was determined by computing a linear regression of log $W$ on log $H$. Figure 9 demonstrates that the resulting age-dependent power $p$ had a rather complicated shape, which was furthermore sex-dependent. After a decrease from about $p = 2.4$ to a value slightly above $p = 2$ at 2 years, a slightly curved increase took place until the pubertal spurt, resulting in a sex-dependent shift of the peak. The highest power necessary was $p = 2.56$ at 14 years for boys and $p = 2.75$ at 11.5 years for girls. The optimal $p$ dropped then to about $p = 2$ for girls and even below for boys (roughly $p = 1.8$). The qualitative pattern is reminiscent of the changes of the correlation of the body mass index with height in figure 4. The bell-shaped function suggested by Cole (1986) for fitting the age-dependence of $p$-values was not in particularly good agreement with the data for boys, whereas the approximation was better for girls. For trunk length we would *a priori* expect a less varying relationship
between the index $p$ and age, with values of $p$ closer to $p=3$ than for height. An analysis of sitting height and subischial leg height brought the following results: the index $p$ for the trunk increased from about 2 in infancy to 2.9 in puberty for boys and to 3.4 for girls, and then decreased to about 2.4 for girls and to about 2.0 for boys. Age variation and sex differences were higher than for height, contrary to expectation, whereas the larger size of $p$ was expected. Leg height started at about $p=0.7$ and increased to about $p=1.5$ in puberty, dropping afterwards to values around $p=1$, much smaller than those obtained for height.

3.3. Transformations towards a normal distribution

The transformations $x$ (raw data), $\sqrt{x}$, $\log(x)$, $-1/\sqrt{x}$, $-1/x$ were applied to the variables $x$: weight, body mass index $W/H^2$, triceps skinfold (representing extremity skinfolds) and subscapular skinfold (representing trunk skinfolds) in order to evaluate the approximation to the normal distribution. All transformations proved to be superior to using the raw data, but the popular logarithmic transformation was not sufficient to provide a good approximation to the normal distribution. For the variables considered transformations ranging from $-1/\sqrt{x}$ to $-1/x$ seem to be more appropriate. The following transformations were preferred after careful evaluation:

- Weight: $-1/\sqrt{x}$
- Body mass index $W/H^2$: $-1/x$ (i.e. $-H^2/W$)
- Arm skinfolds: $-1/\sqrt{x}$
- Trunk skinfolds: $-1/x$
The negative sign ensures that the order between small and large values is not reversed.

The rationale for this choice is illustrated in figures 10–13, depicting skewness and kurtosis (measures of approximation to the normal distribution) before and after transformation. For an easier visualization, the ideal value 0 (valid for the normal distribution) and somewhat arbitrary demarcations at ±0.4 (for skewness) and ±1.0 (for kurtosis, without a negative boundary since a negative kurtosis is not too problematic and also uncommon) are also depicted. Figure 10 shows that the raw body mass index \( W/H^2 \) had a distribution with a long right tail (large positive skewness), increasing with age. A transformation \(-1/x\) was able to bring the distribution within reasonable limits, both with respect to skewness and kurtosis (figure 11). The transformation was not entirely satisfactory for girls after 16 years. Subscapular skinfolds (figures 12 and 13) came out with heavy deviations from a normal distribution (kurtosis was clipped at a value 13 for graphical reasons). The transformation \(-1/x\) brought a massive improvement across the whole age span, leaving us with no gross departures from normality. For triceps skinfolds, skewness increased from negative values in infancy and early childhood to distinctly positive values at and after adolescence. This pattern required a delicate compromise: the transformation \(-1/\sqrt{x}\) proved to be good except at young ages. Weight had again an age-dependent pattern of deviations. A reasonable transformation was \(-1/\sqrt{x}\) except for females towards adulthood, where \(-1/x\) would have been more appropriate. For the sum of all four skinfolds, the transformation \(-1/\sqrt{x}\) proved to be the most adequate one.

![Figure 10](image)

**Figure 10.** Skewness for \( W/H^2 \) in raw form (long dashes), transformed as \( \log(x) \) (solid line) and transformed as \(-1/x\) (short dashes) for boys (above) and girls (below).
Body mass and obesity from infancy to adulthood

Figure 11. Kurtosis for \( W/H^2 \) in raw form (long dashes), transformed as \( \log(x) \) (solid line) and as \(-1/x\) (short dashes) for boys (above) and girls (below).

Figure 12. Skewness for subscapular skinfolds in raw form (long dashes), transformed as \( \log(x) \) (solid line) and as \(-1/x\) (short dashes) for boys (above) and girls (below).
4. Discussion

4.1. Indices of body mass

The Quetelet index \( W/H^2 \) is in most respects distinctly superior to indices \( W/H \) or \( W/H^2 \), confirming earlier work by Rolland-Cachera et al. (1982). In comparison, our correlations were often somewhat higher than theirs. The index \( W/H \) correlates strongly with weight and can thus not give much information beyond weight itself. The Quetelet index correlates little with height, and well with weight. Correlations with single skinfolds and with the sum are rather modest until about 10 years and better afterwards, better for trunk than for arm skinfolds, and better for girls than for boys. The latter finding is in agreement with Cronk, Mukherjee and Roche (1983). The somewhat disappointing correlations between \( W/H^2 \) and fat may be partly attributed to the substantial measurement error of skinfolds which is expected to be relatively larger in the lean phases before about age 9-10 (Gasser, Ziegler, Kneip, Prader, Molinari and Largo 1992). Further, it may be due to the fact that the body mass index to a large extent reflects lean mass. Interestingly, body mass correlates highly with arm circumference.

The Quetelet index is markedly age-dependent, and to some extent also sex-dependent. Thus age norms are required, as computed by Rolland-Cachera et al. (1987). It is pleasing to see that our mean curve for \( W/H^2 \) agrees very well with theirs. However, the analysis of the body mass index \( W/H^2 \) by structural average velocity curves shows that this index also reflects truly developmental changes such as the mid-growth spurt and the pubertal spurt. This may be attributed to morphological changes.

Figure 13. Kurtosis for subscapular skinfolds in raw form (long dashes), transformed as \( \log(x) \) and as \(-1/x\) (short dashes) for boys (above) and girls (below).
occurring in these periods, as well as changes in body composition (Gasser et al. 1992 give fat changes during puberty). It may be surprising to see an accentuated pubertal spurt (PS) for $W/H^2$ since skinfolds decrease or hardly increase at that age. This PS is attributed firstly to a massive PS for muscle mass (Tanner, Hughes and Whitehouse 1981, own unpublished results) and secondly to morphological changes in puberty: measures of width show a stronger pubertal spurt than height (Gasser, Kneip, Ziégler, Largo, Molinari and Prader 1991b) which has to increase the body mass index. The more pronounced pubertal growth of the trunk, compared to the legs, points in the same direction. Also at other developmental stages, it is intrinsically difficult to separate sharply normal changes in body composition and morphology from over- or underweight. This is also illustrated in the positive correlation between the ratio trunk to leg length with the body mass index throughout development. The pubertal spurt in $W/H^2$ occurs slightly after the age of peak height velocity, and it is the first measure where the pubertal spurt is more intense in girls. A possible explanation for this sex difference is the substantially negative velocity in skinfolds for boys, in contrast to girls (Tanner et al. 1981, Gasser et al. 1992).

That the body mass index begins to increase again at about age 6, after a period of decrease, was first described by Rolland-Cachera, Deheeger, Bellsle, Sempé, Guilloud-Bataille and Patois (1984) and termed ‘adiposity rebound’. The negative velocity in $W/H^2$ for both sexes from 1 to 6 years may be explained as follows: fat mass hardly increases and muscle mass only to a modest degree during this period. Height velocity, on the other hand, is large during this period, falling from 14 cm/year to about 6 cm/year (Gasser, Kneip, Binding, Prader and Molinari 1991a) and this is mainly due to growth of the legs, which do not contribute much to body mass. As a consequence the body mass index has to decrease. The drop to negative velocity at about 1 year is also associated with behavioural changes after 1 year, when walking starts and eating is no longer so important.

One desirable feature is lacking for $W/H^2$, i.e. a good approximation to a normal distribution. However, this can, to a great extent, be repaired by applying an appropriate transformation (see below), i.e. the negative inverse transformation. The Quetelet index $W/ll$ is a well-established, simple and still useful measure of body mass. From a scientific point of view it would be desirable to account not only for height when computing such an index but, for example, also for bihumeral width and the ratio trunk to leg length. This would eliminate a confounding with morphological characteristics by taking into account body frame (compare e.g. Himes and Bouchard 1985).

4.2. Age-dependent indices

The approach suggested and evaluated by Cole (1986) of choosing an appropriate index of corpulence is quite refined: it involves finding for each age a power $p$ such that height and weight have a linear relationship. This leads then to an age-dependent index $W/H^p$ while a simple form for the dependence of $p$ on age is sought. Our analysis shows that this dependence is not as simple as modelled by Cole, and that it is influenced by such developmental features as the pubertal spurt. Unfortunately, the separate analysis of trunk and leg length did not improve the situation as a priori expected. For many scientific studies it might also be a disadvantage that different transformations are used for different ages, making age-to-age comparisons more difficult. Thus, given the statistical complexity of the approach, and these further problems, we are inclined to favour the use of the Quetelet index in an appropriate
way, i.e. by applying the negative inverse transformation to achieve an approximate normal distribution and by computing then age- and sex-dependent norms.

4.3. Further results

While the body mass index \( W/H^2 \) is roughly uncorrelated with height, it is modestly correlated with breadth, i.e. biliaic width. It may therefore reflect to some extent morphological characteristics of individuals, in addition to body mass. Skinfolds show moderate correlations with arm circumference after 10 years, and good ones with other skinfolds. Tanner (1965) found correlations between radiographically determined skinfolds at different sites of about 0.7. Ours are close to this value, usually somewhat higher. The higher correlations for girls may be associated with the greater importance of fat in body composition. Probably, the moderate correlations before age 9 and 10 are due to a relatively large measurement error. In any case these results put in question the use of skinfold measurements as predictors before that age.

4.4. Transformations to normal distribution

The above evaluation was based on rank correlations and was thus not relying on normally distributed variables. However, a good approximation to the normal distribution is needed for many statistical methods, and transformations achieving this goal reasonably well across age were sought. Rolland-Cachera et al. (1982) used a logarithmic transformation for \( W/H^2 \) and the same transformation is frequently used for weight and skinfolds. Our results demonstrate that the logarithmic transformation is an improvement compared to using raw data, but is not optimal for the variables considered. When searching for age-dependent optimal transformations for skinfolds in childhood, Van’t Hof et al. (1985) also found transformations 'harder' than the logarithmic one preferable. (From a statistical point of view the transformations \(-1/\sqrt{x}\) and \(-1/x\) are 'harder' than the log-transformation since the raw data are changed to a greater degree, e.g. by eliminating more right-tailed skewness). According to our results the transformation \(-1/\sqrt{x}\) is a good compromise for weight and arm skinfolds, and this is true for the transformation \(-1/x\) and trunk skinfolds and the body mass index \( W/H^2 \). For most ages from birth to young adulthood these transformations led to a good approximation to the normal distribution, and for some restricted age ranges this approximation was less good but still satisfactory. Cole (1988) and Cole and Green (1992) studied the statistical properties of age-dependent power transformations and applied them to weight, whereas Rolland-Cachera, Cole, Sempé, Tichet, Rossignol and Charraud (1991) applied the same technique to the body mass index. Their results confirm that inverse transformations are superior to taking logs in these instances.

References


Body mass and obesity from infancy to adulthood


Address for correspondence: Th. Gasser, Abteilung Biostatistik, ISPM, Universität Zürich, Sumatrastrasse 30, CH-8006, Zürich, Switzerland.
