Acta Pædiatrica, 2006; Suppl 450: 76–85

WHO Child Growth Standards based on length/height, weight and age

WHO MULTICENTRE GROWTH REFERENCE STUDY GROUP[1,2]

[1]*Department of Nutrition, World Health Organization, Geneva, Switzerland, and* [2]*Members of the WHO Multicentre Growth Reference Study Group (listed at the end of the first paper in this supplement)*

Abstract

Aim: To describe the methods used to construct the WHO Child Growth Standards based on length/height, weight and age, and to present resulting growth charts. *Methods:* The WHO Child Growth Standards were derived from an international sample of healthy breastfed infants and young children raised in environments that do not constrain growth. Rigorous methods of data collection and standardized procedures across study sites yielded very high-quality data. The generation of the standards followed methodical, state-of-the-art statistical methodologies. The Box-Cox power exponential (BCPE) method, with curve smoothing by cubic splines, was used to construct the curves. The BCPE accommodates various kinds of distributions, from normal to skewed or kurtotic, as necessary. A set of diagnostic tools was used to detect possible biases in estimated percentiles or z-score curves. *Results:* There was wide variability in the degrees of freedom required for the cubic splines to achieve the best model. Except for length/height-for-age, which followed a normal distribution, all other standards needed to model skewness but not kurtosis. Length-for-age and height-for-age standards were constructed by fitting a unique model that reflected the 0.7-cm average difference between these two measurements. The concordance between smoothed percentile curves and empirical percentiles was excellent and free of bias. Percentiles and z-score curves for boys and girls aged 0–60 mo were generated for weight-for-age, length/height-for-age, weight-for-length/height (45 to 110 cm and 65 to 120 cm, respectively) and body mass index-for-age.

Conclusion: The WHO Child Growth Standards depict normal growth under optimal environmental conditions and can be used to assess children everywhere, regardless of ethnicity, socio-economic status and type of feeding.

Key Words: *Body mass index, growth standards, height, length, weight*

Introduction

Nearly three decades ago, an expert group convened by the World Health Organization (WHO) recommended that the National Center for Health Statistics (NCHS) reference data for height and weight be used to assess the nutritional status of children around the world [1]. This recommendation was made recognizing that not all of the criteria the group used to select the best available reference data had been met. The reference became known as the NCHS/WHO international growth reference and was quickly adopted for a variety of applications regarding both individuals and populations.

The limitations of the NCHS/WHO reference are well known [2–5]. The data used to construct the reference covering birth to 3 y of age came from a longitudinal study of children of European ancestry from a single community in the United States. These children were measured every 3 mo, which is inadequate to describe the rapid and changing rate of growth in early infancy. Also, shortcomings inherent to the statistical methods available at the time for generating the growth curves led to inappropriate modelling of the pattern and variability of growth, particularly in early infancy. For these likely reasons, the NCHS/WHO curves do not adequately represent early childhood growth.

The origin of the WHO Multicentre Growth Reference Study (MGRS) [6] dates back to the early 1990s when the WHO initiated a comprehensive review of the uses and interpretation of anthropometric references and conducted an in-depth analysis of growth data from breastfed infants [2,7]. This analysis showed that breastfed infants from well-off households in northern Europe and North America (i.e. the WHO pooled breastfed data set) deviated negatively and significantly from the NCHS/WHO reference [2,7]. Moreover, healthy breastfed infants from Chile, Egypt, Hungary, Kenya and Thailand

Correspondence: Mercedes de Onis, Study Coordinator, Department of Nutrition, World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland. Tel: +41 22 791 3320. Fax: +41 22 791 4156. E-mail: deonism@who.int

ISSN 0803-5326 print/ISSN 1651-2227 online
DOI: 10.1080/08035320500495548

showed similar deviations when compared to the NCHS/WHO reference but not when compared to the WHO pooled breastfed group [2]. Finally, the variability of growth in the pooled breastfed data set was significantly lower than that of the NCHS/WHO reference [2]. It was unclear whether the reduced variability reflected homogeneity in the WHO pooled breastfed group—perhaps because of uniformity in infant feeding patterns—or unphysiological variability in the NCHS/WHO reference. The data for infants used in the NCHS/WHO reference were collected between 1929 and 1975. The majority of these infants were fed artificial milks which, with increasing knowledge about the nutritional needs of infants, changed in formulation over time. It is thus possible that the greater variability in the current international reference reflects responses to formulas of varying nutritional quality over four decades.

The review group concluded from these and related findings that new references were necessary because the current international reference did not adequately describe the growth of children. Under these circumstances, its uses to monitor the health and nutrition of individual children or to derive population-based estimates of child malnutrition are flawed. The review group recommended a novel approach: that a standard rather than a reference be constructed. Strictly speaking, a reference simply serves as an anchor for comparison, whereas a standard allows both comparisons and permits value judgments about the adequacy of growth. The MGRS breaks new ground by describing how children *should* grow when not only free of disease but also when reared following healthy practices such as breastfeeding and a non-smoking environment.

The MGRS is also unique because it includes children from around the world: Brazil, Ghana, India, Norway, Oman and the USA. In a companion paper in this volume [8], the length of children is shown to be strikingly similar among the six sites, with only about 3% of variability in length being due to inter-site differences compared to 70% for individuals within sites. Thus, excluding any site has little effect on the 3rd, 50th and 97th percentile values, and pooling data from all sites is entirely justified. The striking similarity in growth during early childhood across human populations means either a recent common origin as some suggest [9] or a strong selective advantage across human environments associated with the current pattern of growth and development.

The key objectives of this article are 1) to provide an overview of the methods used to construct the standards for length/height-for-age, weight-for-age, weight-for-length/height and BMI-for-age, and 2) to present some of the resulting curves. Complete details and a full presentation of charts and tables pertaining to the standards are available in a technical report [10] and on the Web: www.who.int/childgrowth/en

Methods

Description and design of the MGRS

The MGRS (July 1997–December 2003) was a population-based study taking place in the cities of Davis, California, USA; Muscat, Oman; Oslo, Norway; and Pelotas, Brazil; and in selected affluent neighbourhoods of Accra, Ghana, and South Delhi, India. The MGRS protocol and its implementation in the six sites are described in detail elsewhere [6]. Briefly, the MGRS combined a longitudinal component from birth to 24 mo with a cross-sectional component of children aged 18–71 mo. In the longitudinal component, mothers and newborns were screened and enrolled at birth and visited at home a total of 21 times on weeks 1, 2, 4 and 6; monthly from 2–12 mo; and bimonthly in the second year. In the cross-sectional component, children aged 18–71 mo were measured once, except in the two sites (Brazil and the USA) that used a mixed-longitudinal design in which some children were measured two or three times at 3-mo intervals. Both recumbent length and standing height were measured for all children aged 18–30 mo. Data were collected on anthropometry, motor development, feeding practices, child morbidity, perinatal factors, and socio-economic, demographic and environmental characteristics [11].

The study populations lived in socio-economic conditions favourable to growth, where mobility was low, $\geq$20% of mothers followed WHO feeding recommendations and breastfeeding support was available [11]. Individual inclusion criteria were: no known health or environmental constraints to growth, mothers willing to follow MGRS feeding recommendations (i.e. exclusive or predominant breastfeeding for at least 4 mo, introduction of complementary foods by 6 mo of age, and continued partial breastfeeding to at least 12 mo of age), no maternal smoking before and after delivery, single term birth, and absence of significant morbidity [11].

As part of the site-selection process in Ghana, India and Oman, surveys were conducted to identify socio-economic characteristics that could be used to select groups whose growth was not environmentally constrained [12–14]. Local criteria for screening newborns, based on parental education and/or income levels, were developed from those surveys. Pre-existing survey data for this purpose were available from Brazil, Norway and the USA. Of the 13741 mother–infant pairs screened for the longitudinal component, about 83% were ineligible [15]. A family's low socio-economic status was the most common reason for ineligibility in Brazil, Ghana, India and Oman, whereas parental refusal was the main reason for non-participation in Norway and the USA [15]. For the cross-sectional component, 69% of the 21510 subjects screened were excluded for reasons similar to those observed in the longitudinal component.

Term low-birthweight (<2500 g) infants (2.3%) were *not* excluded. Since it is likely that, in well-off populations, such infants represent small but normal children, and their exclusion would have artificially distorted the standards' lower percentiles. Eligibility criteria for the cross-sectional component were the same as those for the longitudinal component with the exception of infant feeding practices. A minimum of 3 mo of any breastfeeding was required for participants in the study's cross-sectional component.

Anthropometric methods

Data collection teams were trained at each site during the study's preparatory phase, at which time measurement techniques were standardized against one of two MGRS anthropometry experts. During the study, bimonthly standardization sessions were conducted at each site. Once a year, the anthropometry expert visited each site to participate in these sessions [16]. Results from the anthropometry standardization sessions are reported in a companion paper in this volume [17]. For the longitudinal component of the study, screening teams measured newborns within 24 h of delivery, and follow-up teams conducted home visits until 24 mo of age. The follow-up teams were also responsible for taking measurements in the cross-sectional component involving children aged 18–71 mo [11].

The MGRS data included weight and head circumference at all ages, recumbent length (longitudinal component), height (cross-sectional component), and arm circumference, triceps and subscapular skinfolds (all children aged $\geq$3 mo). However, here we report on only the standards based on length or height and weight. Observers working in pairs collected anthropometric data. Each observer independently measured and recorded a complete set of measurements, after which the two compared their readings. If any pair of readings exceeded the maximum allowable difference for a given variable (weight 100 g; length/height 7 mm), both observers once again independently measured and recorded a second and, if necessary, a third set of readings for the variable(s) in question [16].

All study sites used identical measuring equipment. Instruments needed to be highly accurate and precise, yet sturdy and portable to enable them to be carried back and forth on home visits. Length was measured with the Harpenden Infantometer (range 30–110 cm for portable use, with digit counter readings precise to 1 mm). The Harpenden Portable Stadiometer (range 65–206 cm, digit counter reading) was used for measuring both adult and child heights. Portable electronic scales with a taring capability and calibrated to 0.1 kg (i.e. UNICEF Electronic Scale 890 or UNISCALE) were used to measure weight. Length and height were recorded to the last completed unit rather than to the nearest unit. To correct for the systematic negative bias introduced by this practice, 0.05 cm (i.e. half of the smallest measurement unit) was added to each measurement before analysis. This correction did not apply to weight, which was rounded off to the nearest 100 g. Full details of the instruments used and how measurements were taken are provided elsewhere [16].

Criteria for including children in the sample used to generate the standards

The total sample size for the longitudinal and cross-sectional studies from all six sites was 8440 children. A total of 1743 children were enrolled in the longitudinal sample, six of whom were excluded for morbidities affecting growth (four cases of repeated episodes of diarrhoea, one case of repeated episodes of malaria and one case of protein-energy malnutrition), leaving a final sample of 1737 children (894 boys and 843 girls). Of these, the mothers of 882 children (428 boys and 454 girls) complied fully with the MGRS infant-feeding and no-smoking criteria and completed the follow-up period of 24 mo. The other 855 children contributed only their birth records, as they either failed to comply with the study's criteria or dropped out before 24 mo. The total number of records for the longitudinal component was 19 900. The cross-sectional sample comprised 6697 children. Of these, 28 were excluded for medical conditions affecting growth (20 cases of protein-energy malnutrition, five cases of haemolytic anaemia G6PD deficiency, two cases of renal tubulo-interstitial disease and one case of Crohn disease), leaving a final sample of 6669 children (3450 boys and 3219 girls) with a total of 8306 records.

Data cleaning procedures and exclusions applied to the data

The MGRS data management protocol [18] was designed to create and manage a large databank of information collected from multiple sites over a period of several years. Data collection and processing instruments were prepared centrally and used in a standardized fashion across sites. The data management system contained internal validation features for timely detection of data errors, and its standard operating procedures stipulated a method of master file updating and correction that maintained a clear trail for data-auditing purposes. Each site was responsible for collecting, entering, verifying and validating data, and for creating site-level master files. Data from the sites were sent to the WHO every month for master file consolidation and more extensive quality-control checking. All errors identified were communicated to the site for correction at source.

After data collection was completed at a given site, a period of about 6 mo was dedicated to in-depth data quality checking and master file cleaning. The WHO produced detailed validation reports, descriptive statistics and plots from the site's master files. For the longitudinal component, each anthropometric measurement was plotted for every child from birth to the end of his/her participation. These plots were examined individually for any questionable patterns. Query lists from these analyses were sent to the site for investigation and correction, or confirmation, as required. As with the data collection process, the site data manager prepared correction batches to update the master files. The updated master files were then sent to the WHO, and this iterative quality assurance process continued until both the site and WHO were satisfied that all identifiable problems had been detected and corrected. The rigorous implementation of what was a highly demanding protocol yielded very high-quality data.

To avoid the influence of unhealthy weights for length/height, prior to constructing the standards, observations falling above +3 SD and below −3 SD of the sample median were excluded. For the cross-sectional sample, the +2 SD cut-off (i.e. 97.7 percentile) was applied instead of +3 SD as the sample was exceedingly skewed to the right, indicating the need to identify and exclude high weights for height. This cut-off was considered to be conservative given that various definitions of overweight all apply lower cut-offs than the one we used [19,20]. The procedure by which this was done is described in the technical report outlining the construction of the standards [10]. The number of observations excluded for unhealthy weight-for-length/height was 185 (1.4%) for boys and 155 (1.1%) for girls, most of which were in the upper end of the cross-sectional sample distribution. In addition, a few influential observations for indicators other than weight-for-height were excluded when constructing the individual standards: for boys, four (0.03%) observations for weight-for-age and three (0.02%) observations for length/height-for-age; and for girls, one (0.01%) and two (0.01%) observations for the same indicators, respectively.

Statistical methods for constructing the WHO child growth curves

The construction of the child growth curves followed a careful, methodical process. This involved a) detailed examination of existing methods, including types of distributions and smoothing techniques, in order to identify the best possible approach; b) selection of a software package flexible enough to allow comparative testing of alternative methods and the actual generation of the curves; and c) systematic application of the selected approach to the data to generate the models that best fit the data.

A group of statisticians and growth experts met at the WHO to review possible choices of methods and to define a strategy and criteria for selecting the most appropriate model for the MGRS data [21]. As many as 30 methods for attained growth curves were examined. The group recommended that methods based on selected distributions be compared and combined with two smoothing techniques for fitting its parameter curves to further test and provide the best possible method for constructing the WHO child growth standards.

Choice of distribution. Five distributions were identified for detailed testing: the Box-Cox power exponential [22], the Box-Cox t [23], the Box-Cox normal [24], the Johnson's SU [25] and the modulus-exponential-normal [26]. The first four distributions were fitted using the GAMLSS (Generalized Additive Models for Location, Scale and Shape) software [27] and the last using the "xriml" module in the STATA software [28]. The Box-Cox power exponential (BCPE) with four parameters—μ (for the median), σ (coefficient of variation), ν (Box-Cox transformation power) and τ (parameter related to kurtosis)—was selected as the most appropriate distribution for constructing the curves. The BCPE is a flexible distribution that simplifies to the normal distribution when $\nu=1$ and $\tau=2$. Also, when $\nu\neq1$ and $\tau=2$, the distribution is the same as the Box-Cox normal (LMS method distribution). The BCPE is defined by a power transformation (or Box-Cox transformation) having a shifted and scaled (truncated) power exponential (or Box-Tiao) distribution with parameter τ [22]. Apart from other theoretical advantages, the BCPE presents as good as or better goodness of fit than the modulus-exponential-normal or the SU distribution.

Choice of smoothing technique. Two smoothing techniques were recommended for comparison by the expert group: cubic splines and fractional polynomials [21]. Using GAMLSS, comparisons were carried out for length/height-for-age, weight-for-age and weight-for-length/height. The cubic spline smoothing technique offered more flexibility than fractional polynomials in all cases. For the length-for-age and weight-for-age standards, a power transformation applied to age prior to fitting was necessary to enhance the goodness of fit by the cubic splines technique.

Choice of method for constructing the curves. In summary, the BCPE method, with curve smoothing by cubic splines, was selected as the approach for constructing the growth curves. This method is included in a broader methodology, the GAMLSS [29], which offers a general framework that includes a wide range of known methods for constructing growth curves. The GAMLSS allows for modelling the mean (or location) of the growth variable under consideration as well as other parameters of its distribution that determine scale and shape. Various kinds of distributions can be assumed for each growth variable of interest, from normal to highly skewed and/or kurtotic distributions. Several smoothing terms can be used in generating the curves, including cubic splines, lowess (locally weighted least squares regression), polynomials, power polynomials and fractional polynomials.

Process and diagnostic criteria for selecting the best model to construct the curves. The process for selecting the best model to construct the curves for each growth variable involved selecting first the best model *within* a class of models and, second, the best model *across* different classes of models. The Akaike Information Criteria [30] and the generalized version of it [22] were used to select the best model *within* a considered class of models. In addition, worm plots [31] and Q-tests [32] were used to determine the adequate numbers of degrees of freedom for the cubic splines fitted to the parameter curves. In most cases, it was necessary to transform age before fitting the cubic splines to "stretch" the age scale during the neonatal period when growth is rapid and the rise in percentile curves is steep. Thus, selecting the best model within the same class of models involved finding the best choice for degrees of freedom for the parameter curves, determining whether age needed to be transformed and finding the best power (λ). In selecting the best model *across* different classes of models, we started from the simplest class of models (i.e. the normal distribution) and proceeded to more complex models when necessary. The goal was to test the impact of increasing the model's complexity on its goodness of fit. The same set of diagnostic tools/tests was used at this stage.

Two diagnostic tools were used to detect possible biases in estimated percentile or z-score curves. First, we examined the pattern of differences between empirical and fitted percentiles; second, we compared observed and expected proportions of children with measurements below selected percentiles or z-score curves.

A more detailed description of the statistical methods and procedures that were followed to construct the WHO Child Growth Standards is provided elsewhere [10].

Types of curves generated

Percentile and z-score curves were generated ranging from the 99th to the 1st percentile and from +3 to −3 standard deviations, respectively. Due to space constraints, we present in this article only the z-score curves for the following lines: 3, 2, 1, 0, −1, −2 and −3 standard deviations. An extensive display of the standards' charts and tables containing such information as means and standard deviations by age and sex, percentile values and related measures is provided in the technical report [10] and on the Web: www.who.int/childgrowth/en

Results

The specifications of the BCPE models that provided the best fit to generate specific standards are summarized in Table I. These are specific values for the age power transformation and the degrees of freedom for the cubic spline functions fitting the four parameters that define the BCPE distribution selected for each standard. Age needed to be transformed for boys and girls except for weight-for-length/height and BMI curves from 24 to 60 mo. There was wide variability in the degrees of freedom that were necessary for the cubic splines to achieve the best fit for modelling the median (μ) and its coefficient of variation (σ). In the case of length/height-for-age for boys and girls, the normal distribution (i.e. when ν takes the value of 1 and τ is 2) proved to be the parsimonious option. In all other cases, it was necessary to model skewness (ν) but not kurtosis (i.e. τ was 2 for all standards), which simplified the model considerably. One to three degrees of freedom for the ν parameter were sufficient in all cases where the distribution was skewed (Table I). The degrees of freedom chosen for boys and girls were often the same or similar.

It was possible to construct both length-for-age (0 to 2 y) and height-for-age (2 to 5 y) standards fitting a unique model, yet still reflecting the difference between recumbent length and standing height. The cross-sectional component included the measurement of both length and height in children 18 to 30 mo old ($n = 1625$ children), and from these data it was estimated that length was the larger measure by 0.7 cm [10]. To fit a single model for the whole age range, 0.7 cm was therefore added to the cross-sectional height values. After the model was fitted, the final curves were shifted downwards by 0.7 cm for ages 2 y and above to create the height-for-age standards. Coefficient of variance values were adjusted to reflect this back transformation using the shifted medians and standard deviations. The length-for-age (0 to 24

Table I. Degrees of freedom for fitting the parameters of the Box-Cox power exponential (BCPE) distribution for the models with the best fit to generate standards based on age, length and weight in children 0–60 mo of age.

Standards	Sex	λ[a]	df(μ)[b]	df(σ)[c]	df(ν)[d]	τ[e]
Length/height, 0–60 mo	Boys	0.35	12	6	0[f]	2
Length/height, 0–60 mo	Girls	0.35	10	5	0[f]	2
Weight, 0–60 mo	Boys	0.35	11	7	2	2
Weight, 0–60 mo	Girls	0.35	11	7	3	2
Weight-for-length/height, 0–60 mo	Boys	None	13	6	1	2
Weight-for-length/height, 0–60 mo	Girls	None	12	4	1	2
BMI, 0–24 mo	Boys	0.05	10	4	3	2
BMI, 0–24 mo	Girls	0.05	10	3	3	2
BMI, 24–60 mo	Boys	None	4	3	3	2
BMI, 24–60 mo	Girls	None	4	4	1	2

[a] Age transformation power.
[b] Degrees of freedom for the cubic splines fitting the median (μ).
[c] Degrees of freedom for the cubic splines fitting the coefficient of variation (σ).
[d] Degrees of freedom for the cubic splines fitting the Box-Cox transformation power (ν).
[e] Parameter related to the kurtosis fixed ($\tau = 2$).
[f] $\nu = 1$: normal distribution.

mo) standard was derived directly from the fitted model. A similar approach was followed in generating the weight-for-length (45 to 110 cm) and weight-for-height (65 to 120 cm) standards. In the generation of the length/height-for-age standards, data up to 71 mo of age were used and the fitted model truncated at 60 mo in order to control for edge effects. For the weight-for-length/height standards, data up to 120 cm height were used to fit the model to prevent the fitting from being influenced by the portion of the data presenting instability [10].

In addressing the differences between length and height, a different approach for the BMI-for-age standards was followed because BMI is a ratio with length or height squared in the denominator. After adding 0.7 cm to the height values, it was not possible, after fitting, to back-transform lengths to heights. The solution adopted was to construct the standards for younger and older children separately based on two sets of data with an overlapping range of ages below and above 24 mo. To construct the BMI-for-age standard using length (0–2 y), the longitudinal sample and the cross-sectional height data up to 30 mo were used after adding 0.7 cm to the height values. Analogously, to construct the standard from 2 to 5 y, the cross-sectional sample plus the longitudinal length from 18–24 mo were used after subtracting 0.7 cm from the length values. Thus, a common set of data from 18 to 30 mo was used to generate the BMI standards for younger and older children.

The concordance between smoothed percentile curves and observed or empirical percentiles was remarkably good. As examples, we show comparisons for the 3rd, 10th, 50th, 90th and 97th percentiles for length-for-age for boys (Figure 1) and for weight-for-height for girls (Figure 2). Overall, the fit was best for length and height-for-age standards, but it was almost as good for the standards based on combinations of weight and length [10]. The average absolute difference between smoothed and empirical percentiles was small: 0.13 cm for length-for-age in boys 0 to 24 mo (Figure 1) and 0.16 kg for weight-for-height for girls 65 to 120 cm (Figure 2). Taking the sign into account, the average differences are close to zero: -0.03 cm and -0.02 kg in Figures 1 and 2, respectively, which indicates lack of bias in the fit between smoothed and empirical percentiles.

Z-score curves are given for length/height-for-age for boys and girls from birth to 60 mo of age (Figures 3 and 4), weight-for-age for boys and girls from birth to 60 mo (Figures 5 and 6), weight-for-length for boys and girls 45 to 110 cm (Figures 7 and 8), weight-for-height for boys and girls 65 to 120 cm (Figures 9 and 10) and BMI-for-age for boys and girls from birth to 60 mo (Figures 11 and 12). The last are in addition to the previously available set of indicators in the NCHS/WHO reference.

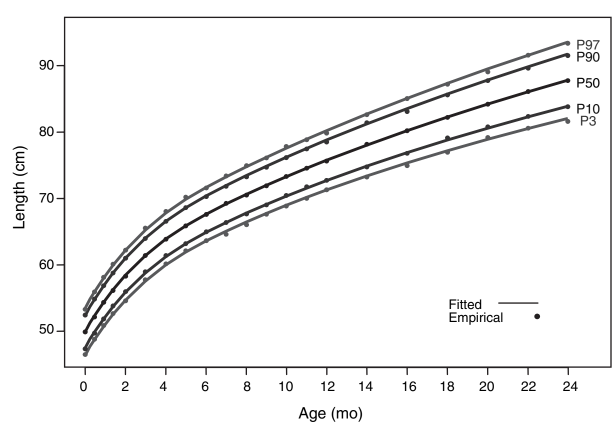

Figure 1. Comparisons between 3rd, 10th, 50th, 90th and 97th smoothed percentile curves and empirical values for length-for-age for boys.

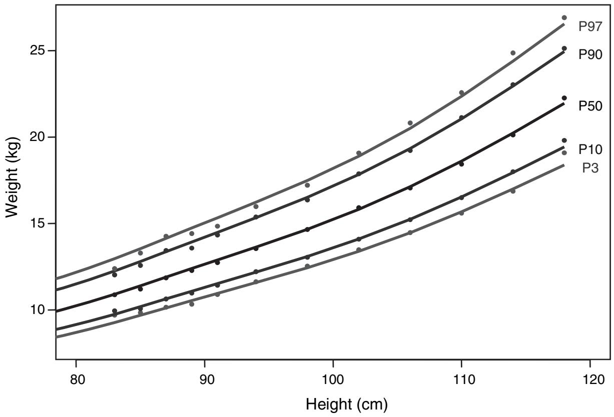

Figure 2. Comparisons between 3rd, 10th, 50th, 90th and 97th smoothed percentile curves and empirical values for weight-for-height for girls.

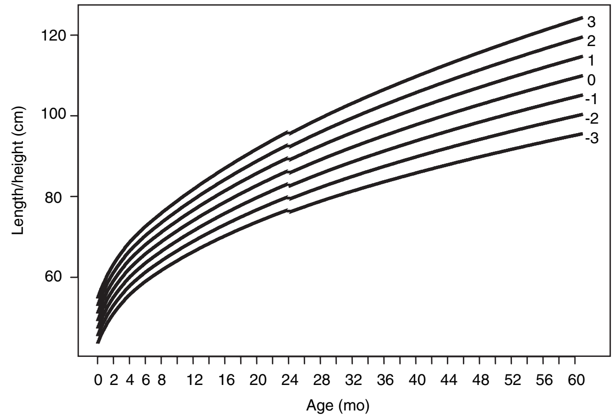

Figure 4. Z-score curves for length/height-for-age for girls from birth to 60 mo. Length from birth to 23 completed months; height from 24 to 60 completed months.

Discussion

The goal of the MGRS was to describe the growth of healthy children. Criteria were applied in the study design to achieve this aim. Screening at enrolment using site-specific socio-economic criteria and maternal non-smoking status excluded children likely to experience constrained growth. Morbidities that affect growth (e.g. repeated bouts of infectious diarrhoea and Crohn disease) were identified, and affected children were excluded from the sample. Application of these criteria resulted in no evidence of under-nutrition in either the longitudinal or cross-sectional samples.

In the longitudinal sample, the behavioural criteria of breastfeeding through 12 mo and its close monitoring throughout data collection yielded a sample of children with no evidence of over-nutrition (i.e. no excessive right skewness). In the cross-sectional sample, however, despite the criterion of at least 3 mo of any breastfeeding, the sample was exceedingly skewed to the right, indicating the need to identify and exclude excessively high weights for heights if the goal of constructing a standard was to be satisfied. A similar prescriptive approach was taken by the developers of the 2000 CDC growth charts for the USA when excluding data from the last national survey (i.e. NHANES III) for children aged $\geq$6 y from the revised weight and BMI growth charts [33]. Without this exclusion, the 95th and 85th percentile curves of the CDC charts would have been higher, and fewer children would have been classified as overweight or at risk of overweight.

Rigorous methods of data collection, standardized across sites, were followed during the entire study. Sound procedures for data management and cleaning were applied. As a result, the anthropometric data available for analysis were of the highest possible quality. A process of consultation with experts in statistical methods and growth was followed, and methodical, state-of-the-art statistical methodologies were employed to generate the standards [21]. The fit between the smoothed curves and empirical or observed percentiles was excellent and free of bias at

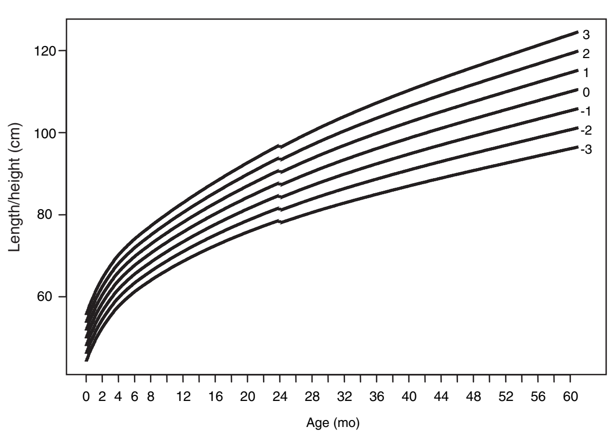

Figure 3. Z-score curves for length/height-for-age for boys from birth to 60 mo. Length from birth to 23 completed months; height from 24 to 60 completed months.

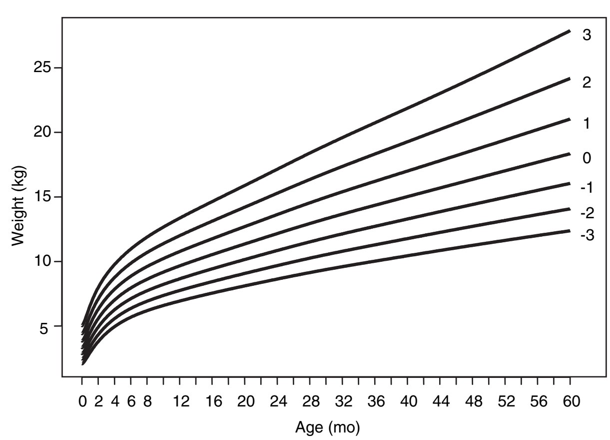

Figure 5. Z-score curves for weight-for-age for boys from birth to 60 mo.

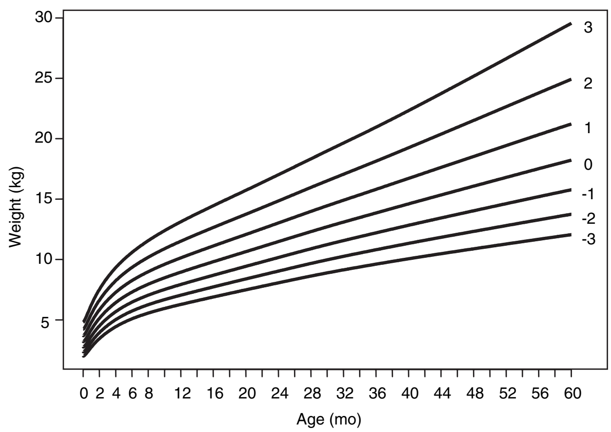

Figure 6. Z-score curves for weight-for-age for girls from birth to 60 mo.

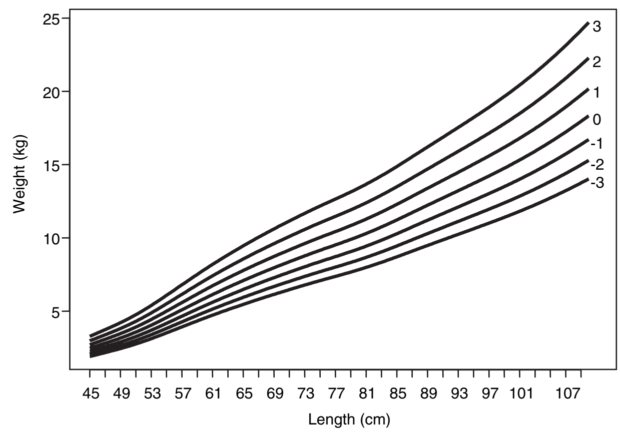

Figure 8. Z-score curves for weight-for-length for girls from 45 to 110 cm.

both the median and the edges, indicating that the resulting curves are a fair description of the true growth of healthy children. Thus, the MGRS can serve as a model of how studies of this type should be carried out and analysed.

The technical report, of which this article is a summary, includes a comparison of the new WHO standards to the previously recommended NCHS/WHO international reference [10]. As expected, there are important differences. However, these vary—by anthropometric measure, sex, specific percentile or z-score curve, and age—in ways that are not easily summarized. Differences are particularly important in infancy. Impact on population estimates of child malnutrition will depend on age, sex, anthropometric indicator considered and population-specific anthropometric characteristics. Thus, it will not be possible to provide an algorithm that will convert new prevalence values from old ones. A notable effect is that stunting will be greater throughout childhood when assessed using the new WHO standards compared to the previous international reference. The growth pattern of breastfed infants compared to the NCHS/WHO reference will result in a substantial increase in underweight rates during the first half of infancy (i.e. 0–6 mo) and a decrease thereafter. For wasting, the main difference between the new standards and the old reference is during infancy (i.e. up to about 70 cm length) when wasting rates will be substantially higher using the new WHO standards. With respect to overweight, use of the new WHO standards will result in a greater prevalence that will vary by age, sex and nutritional status of the index population.

The WHO Child Growth Standards were derived from children who were raised in environments that minimized constraints to growth such as poor diets and infection. In addition, their mothers followed healthy practices such as breastfeeding their children and not smoking during and after pregnancy. The standards depict normal human growth under optimal environmental conditions and can be used to assess children everywhere, regardless of ethnicity, socio-economic status and type of feeding. It would be as inappropriate to call for separate standards to be developed for children whose mothers smoked during pregnancy as it would be for children who are fed a

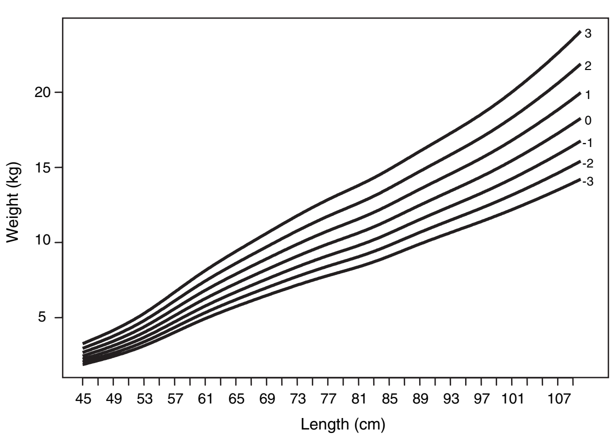

Figure 7. Z-score curves for weight-for-length for boys from 45 to 110 cm.

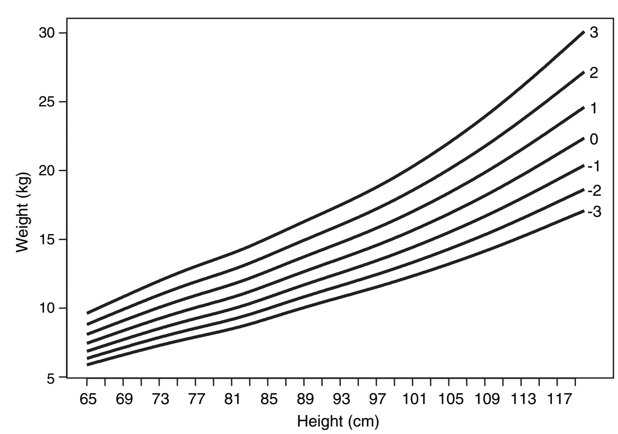

Figure 9. Z-score curves for weight-for-height for boys from 65 to 120 cm.

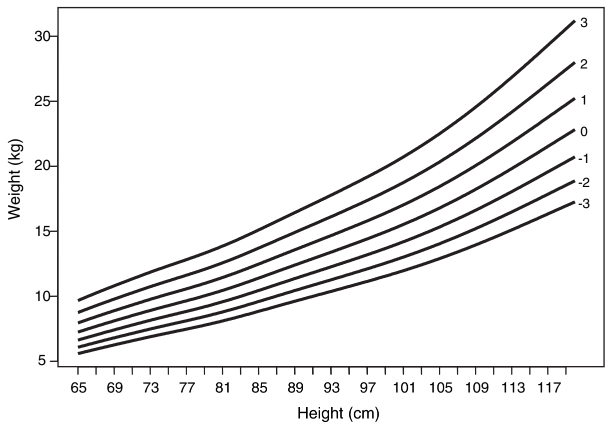

Figure 10. Z-score curves for weight-for-height for girls from 65 to 120 cm.

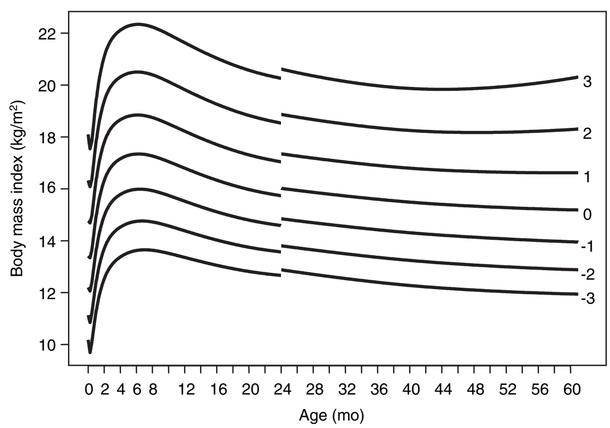

Figure 11. Z-score curves for BMI-for-age for boys from birth to 60 mo. BMI based on length from birth to 23 completed months; BMI based on height from 24 to 60 completed months.

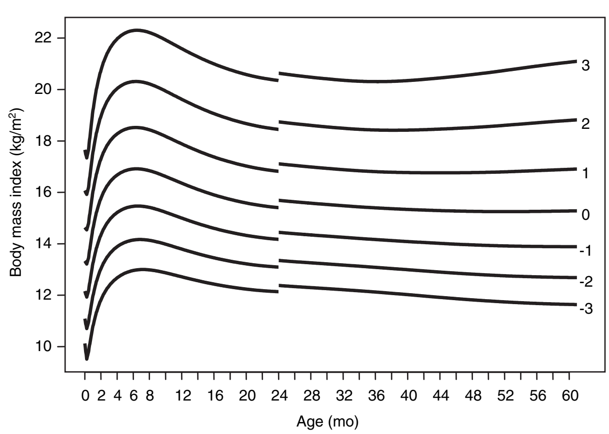

Figure 12. Z-score curves for BMI-for-age for girls from birth to 60 mo. BMI based on length from birth to 23 completed months; BMI based on height from 24 to 60 completed months.

breast-milk substitute. Rather, deviations from any area in the world in the patterns described by the standards, such as a high proportion of children with short heights or high weight-for-heights, when properly assessed and interpreted, should be seen as representing abnormal growth and taken as evidence of stunting and obesity, respectively, in these examples.

Acknowledgements

This paper was prepared by Mercedes de Onis, Reynaldo Martorell, Cutberto Garza and Anna Lartey on behalf of the WHO Multicentre Growth Reference Study Group. The statistical analysis was conducted by Elaine Borghi.

References

[1] Waterlow JC, Buzina R, Keller W, Lane JM, Nichaman MZ, Tanner JM. The presentation and use of height and weight data for comparing the nutritional status of groups of children under the age of 10 years. Bull World Health Organ 1977;55: 489–98.
[2] WHO Working Group on Infant Growth. An evaluation of infant growth. Geneva: World Health Organization; 1994.
[3] de Onis M, Yip R. The WHO growth chart: historical considerations and current scientific issues. Bibl Nutr Dieta 1996;53:74–89.
[4] de Onis M, Habicht JP. Anthropometric reference data for international use: Recommendations from a World Health Organization Expert Committee. Am J Clin Nutr 1996;64: 650–8.
[5] Garza C, de Onis M, for the WHO Multicentre Growth Reference Study Group. Rationale for developing a new international growth reference. Food Nutr Bull 2004;25 Suppl 1:S5–14.
[6] de Onis M, Garza C, Victora CG, Bhan MK, Norum KR, editors. WHO Multicentre Growth Reference Study (MGRS): Rationale, planning and implementation. Food Nutr Bull 2004;25 Suppl 1:S1–89.
[7] WHO Working Group on Infant Growth. An evaluation of infant growth: The use and interpretation of anthropometry in infants. Bull World Health Organ 1995;73:165–74.
[8] WHO Multicentre Growth Reference Study Group. Assessment of differences in linear growth among populations in the WHO Multicentre Growth Reference Study. Acta Paediatr Suppl 2006;450:56–65.
[9] Rosenberg NA, Pritchard JK, Weber JL, Cann HM, Kidd KK, Zhivotovsky LA, et al. Genetic structure of human populations. Science 2002;298:2381–5.
[10] WHO Multicentre Growth Reference Study Group. 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; 2006.
[11] de Onis M, Garza C, Victora CG, Onyango AW, Frongillo EA, Martines J, for the WHO Multicentre Growth Reference Study Group. The WHO Multicentre Growth Reference Study: Planning, study design and methodology. Food Nutr Bull 2004;25 Suppl 1:S15–26.
[12] Owusu WB, Lartey A, de Onis M, Onyango AW, Frongillo EA. Factors associated with unconstrained growth among affluent Ghanaian children. Acta Paediatr 2004;93:1115–9.
[13] Bhandari N, Bahl R, Taneja S, de Onis M, Bhan MK. Growth performance of affluent Indian children is similar to that in developed countries. Bull World Health Organ 2002;80:189–95.

[14] Mohamed AJ, Onyango AW, de Onis M, Prakash N, Mabry RM, Alasfoor DH. Socioeconomic predictors of unconstrained child growth in Muscat, Oman. East Mediterr Health J 2004;10:295–302.

[15] WHO Multicentre Growth Reference Study Group. Enrolment and baseline characteristics in the WHO Multicentre Growth Reference Study. Acta Paediatr Suppl 2006;450:7–15.

[16] de Onis M, Onyango AW, Van den Broeck J, Chumlea WC, Martorell R, for the WHO Multcentre Growth Reference Study Group. Measurement and standardization protocols for anthropometry used in the construction of a new international growth reference. Food Nutr Bull 2004;25 Suppl 1:S27–36

[17] WHO Multicentre Growth Reference Study Group. Reliability of anthropometric measurements in the WHO Multicentre Growth Reference Study. Acta Paediatr 2006;Suppl 450:38–46.

[18] Onyango AW, Pinol AJ, de Onis M, for the WHO Multicentre Growth Reference Study Group. Managing data for a multicountry longitudinal study: Experience from the WHO Multicentre Growth Reference Study. Food Nutr Bull 2004;25 Suppl 1:S46–52.

[19] Daniels SR, Arnett DK, Eckel RH, Gidding SS, Hayman LL, Kumanyika S, et al. AHA scientific statement. Overweight in children and adolescents: Pathophysiology, consequences, prevention and treatment. Circulation 2005;111:1999–2012.

[20] Koplan JP, Liverman CT, Kraak VI, editors. Preventing childhood obesity: Health in the balance. Washington, DC: National Academies Press; 2005.

[21] Borghi E, de Onis M, Garza C, Van den Broeck J, Frongillo EA, Grummer-Strawn L, et al., for the WHO Multicentre Growth Reference Study Group. Construction of the World Health Organization child growth standards: selection of methods for attained growth curves. Stat Med 2006;25:247–65.

[22] Rigby RA, Stasinopoulos DM. Smooth centile curves for skew and kurtotic data modelled using the Box-Cox power exponential distribution. Stat Med 2004;23:3053–76.

[23] Rigby RA, Stasinopoulos DM. Box-Cox t distribution for modelling skew and leptokurtotic data. Technical report 01/04. London: STORM Research Centre, London Metropolitan University; 2004.

[24] Cole TJ, Green PJ. Smoothing reference centile curves: The LMS method and penalized likelihood. Stat Med 1992;11:1305–19.

[25] Johnson NL. Systems of frequency curves generated by methods of translation. Biometrika 1949;36:149–76.

[26] Royston P, Wright EM. A method for estimating age-specific reference intervals ('normal ranges') based on fractional polynomials and exponential transformation. J R Stat Soc Ser A 1998;161:79–101.

[27] Stasinopoulos DM, Rigby RA, Akantziliotou C. Instructions on how to use the GAMLSS package in R. Technical report 02/04. London: STORM Research Centre, London Metropolitan University; 2004.

[28] Wright E, Royston P. Age-specific reference intervals ("normal ranges"). Stata Technical Bulletin 1996;34:24–34.

[29] Rigby RA, Stasinopoulos DM. Generalized additive models for location, scale and shape. J R Stat Soc Ser C Appl Stat 2005;54:507–44.

[30] Akaike H. A new look at the statistical model identification. IEEE Trans Automat Contr 1974;19:716–23.

[31] van Buuren S, Fredriks M. Worm plot. A simple diagnostic device for modelling growth reference curves. Stat Med 2001;20:1259–77.

[32] Royston P, Wright EM. Goodness-of-fit statistics for age-specific reference intervals. Stat Med 2000;19:2943–62.

[33] Kuczmarski RJ, Ogden CL, Guo SS, Grummer-Strawn LM, Flegal KM, Mei Z, et al. 2000 CDC growth charts for the United States: Methods and development. National Center for Health Statistics. Vital Health Stat Series 11 No. 246, 2002:1–190.