2. METHODOLOGY

2.1 Design of the WHO Multicentre Growth Reference Study

The Multicentre Growth Reference Study (MGRS) (July 1997–December 2003) was a population-based study that took 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 (de Onis et al., 2004a). Briefly, the MGRS combined a longitudinal component from birth to 24 months with a cross-sectional component of children aged 18–71 months. 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 months; and bimonthly in the second year. In the cross-sectional component, children aged 18–71 months were measured once, except in the two sites (Brazil and USA) that used a mixed-longitudinal design in which some children were measured two or three times at three-month intervals. Both recumbent length and standing height were measured for all children aged 18–30 months. Data were collected on anthropometry, motor development, feeding practices, child morbidity, perinatal factors, and socioeconomic, demographic and environmental characteristics (de Onis et al., 2004b).

The study populations lived in socioeconomic conditions favourable to growth and where mobility was low, ≥20% of mothers followed WHO feeding recommendations and breastfeeding support was available (de Onis et al., 2004b). 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 months, introduction of complementary foods by the age of 6 months, and continued partial breastfeeding up to at least 12 months), no maternal smoking before and after delivery, single term birth, and absence of significant morbidity (de Onis et al., 2004b).

As part of the site-selection process in Ghana, India and Oman, surveys were conducted to identify socioeconomic characteristics that could be used to select groups whose growth was not environmentally constrained (Owusu et al., 2004; Bhandari et al., 2002; Mohamed et al., 2004). 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 13 741 mother-infant pairs screened for the longitudinal component, about 83% were ineligible (WHO Multicentre Growth Reference Study Group, 2006e). Families’ low socioeconomic 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 (WHO Multicentre Growth Reference Study Group, 2006e). For the cross-sectional component, 69% of the 21 510 subjects screened were excluded for reasons similar to those observed in the longitudinal component.

Term low-birth-weight (<2500 g) infants (2.3%) were not excluded. Since it is likely that in well-off populations such infants represent small but normal children, 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 three months of any breastfeeding was required for participants in the study’s cross-sectional component.

2.2 Anthropometry 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 (de Onis et al., 2004c). Results from the anthropometry standardization sessions have been reported elsewhere (WHO Multicentre Growth Reference Study Group, 2006e).
Methodology

Growth Reference Study Group, 2006f). For the longitudinal component of the study, screening teams measured newborns within 24 hours of delivery, and follow-up teams conducted home visits until 24 months of age. The follow-up teams were also responsible for taking measurements in the cross-sectional component involving children aged 18–71 months (de Onis et al., 2004b). 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 ≥3 months). This report presents only the standards based on head circumference, arm circumference and skinfolds. The standards based on length or height and weight are presented in an earlier publication (WHO Multicentre Growth Reference Study Group, 2006d; web site http://www.who.int/childgrowth/publications/technical_report_pub/en/index.html).

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 (e.g. head circumference, 5 mm; arm circumference, 5 mm; skinfold thickness, 2 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 (de Onis et al., 2004c).

All study sites used identical measuring equipment. Instruments needed to be highly accurate and precise. A self-retracting, 0.7 cm-wide, flat metal tape with blank lead-in strip (range, 0–200 cm, calibrated to 1 mm), was used to measure circumferences. Metal tapes were chosen because they are more robust and accurate, and stay in a single plane around the head. They were replaced on a regular basis when the grading marks faded. The Holtain/Tanner-Whitehouse skinfold caliper (jaw face area, 35 mm²; pressure between the jaws, 10 ± 2 g/mm²; range, 0–40 mm; calibrated to 0.2 mm) was used to measure skinfolds. The skinfold calipers, being particularly fragile, were checked before each use with calibration blocks of various widths for accuracy and to ensure that the needle moved smoothly and continuously with the opening of the caliper jaws. Full details of the instruments used and how measurements were taken are provided elsewhere (de Onis et al., 2004c).

2.3 Sample description

The total sample size for the longitudinal and cross-sectional components 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 (4 cases of repeated episodes of diarrhoea, 1 case of repeated episodes of malaria, and 1 case of protein-energy malnutrition) leaving a 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 months (96% of compliant children completed the 24-month follow-up). The other 855 either failed to comply with the study's infant-feeding and no-smoking criteria or dropped out before 24 months. These children, whose size at birth was similar to that of the compliant sample, contributed only birth measurements. The increased sample size at birth served to minimize the left-edge effect in the head circumference-for-age curves. For arm circumference and the skinfolds, which were measured starting at age 3 months, the data did not allow for this correction in the corresponding standards. 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). The total number of records in the cross-sectional component was 8306 as some children in Brazil and the USA were measured two or three times at three-month intervals. A full description of the MGRS sample with regard to screening,
recruitment, sample attrition and compliance, as well as the baseline characteristics of the study sample is provided elsewhere (WHO Multicentre Growth Reference Study Group, 2006e).

2.4 Data cleaning procedures and exclusions

Data cleaning

The MGRS data management protocol (Onyango et al., 2004) 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 WHO/HQ 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 months was dedicated to in-depth data quality checking and master file cleaning. Detailed validation reports, descriptive statistics and plots were produced 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 WHO/HQ and this iterative quality assurance process continued until all identifiable problems had been detected and corrected. The rigorous implementation of what was a highly demanding protocol yielded very high-quality data.

Data exclusions

In addition to exclusions of data based on weight-for-height (WHO Multicentre Growth Reference Study Group, 2006d), a few influential observations outside ±4 SD were excluded when constructing the individual standards included in this report. These were, for head circumference-for-age: boys, all 21 observations belonging to one boy in the longitudinal sample (0.15%) and one single observation (0.01%) among girls; for triceps skinfold-for-age: boys, 7 (0.06%) and girls, 9 (0.08%); for subscapular skinfold-for-age: boys, 17 (0.16%) and girls, 19 (0.17%) observations. These observations were set to missing in the final data set and therefore did not contribute to the construction of the corresponding standards. The final number of observations used in the construction of the WHO child growth standards is shown in Table 1.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Girls</th>
<th>Boys</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head circumference-for-age</td>
<td>13 798</td>
<td>13 541</td>
<td>27 339</td>
</tr>
<tr>
<td>Arm circumference-for-age</td>
<td>10 970</td>
<td>10 770</td>
<td>21 740</td>
</tr>
<tr>
<td>Triceps skinfold-for-age</td>
<td>10 943</td>
<td>10 762</td>
<td>21 705</td>
</tr>
<tr>
<td>Subscapular skinfold-for-age</td>
<td>10 934</td>
<td>10 757</td>
<td>21 691</td>
</tr>
</tbody>
</table>
2.5 Statistical methods for constructing the growth curves

The underlying methodology used for constructing the head circumference-for-age, arm circumference-for-age, triceps skinfold-for-age and subscapular skinfold-for-age standards was the same used to construct the standards for length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age. The growth curve fitting method and diagnostic tools used to select the best models for each of the indicators are described in detail in the report of the first set of standards (WHO Multicentre Growth Reference Study Group, 2006d, 2006g).

The Box-Cox-power-exponential (BCPE) method (Rigby and Stasinopoulos, 2004), with curve smoothing by cubic splines was selected for constructing the WHO child growth curves. The BCPE accommodates various kinds of distributions, from normal to skewed or kurtotic. All the indicators in this second set standards required a power-transformation to stretch the age scale (x-axis) as a preliminary step to fitting the curves. For each set of curves, the search for the best model specification began by examining various combinations of degrees of freedom to fit the median and variance estimator curves. When data had a non-normal distribution, degrees of freedom for parameters to model skewness and kurtosis were added to the initial model and adequacy of fit was evaluated. Apart from head circumference-for-age, which followed a normal distribution, the other standards required the modelling of skewness, but not kurtosis. The diagnostic tools used iteratively to detect possible model misfits and biases in the fitted curves included various tests of local and global goodness of fit, like Q-tests (Royston and Wright, 2000), worm plots (van Buuren and Fredriks, 2001) and residual plots. Patterns of differences between empirical and fitted percentiles were also examined, as were proportions of observed versus expected percentages of children with measurements below selected percentiles. The curves were constructed using all available data (i.e. from birth to 71 months) but the final standards were truncated at 60 completed months to avoid the right-edge effect (Borghi et al., 2006).

The GAMLSS package (Stasinopoulos et al., 2004) was used for the construction of the growth curves.