Field-Testing the WHO Child Growth Standards in Four Countries

Adelheid W. Onyango,2* Mercedes de Onis,2 Margherita Caroli,3 Uzma Shah,4 Yanina Sguassero,5 Nora Redondo,6 and Berenise Carroli3

2Department of Nutrition, World Health Organization, CH-1211 Geneva 27, Switzerland; 3Nutrition Unit, Department of Prevention AUSL BR/1, Brindisi, Italy; 4Pediatric Gastroenterology, Hepatology and Nutrition, The Aga Khan University Hospital, Karachi, Pakistan; 5Centro Rosarino de Estudios Perinatales (CREP), Rosario, Argentina; and 6Department of Health, Municipality of Rosario, Santa Fe, Argentina

Abstract

In April 2006 the WHO released a set of growth standards for children from birth to the age of 5 y. Prior to their release, the standards were field-tested in 4 countries. The main objective was to compare children’s length/height-for-age and weight-for-length/height based on the new standards with clinician assessments of the same children. The study sampled children <5 y–old attending well-child clinics in 2 affluent populations (Argentina and Italy) and 2 less-affluent ones (Maldives and Pakistan). Length/height and weight were measured by doctors and epidemiologists who also recorded a clinical assessment of each child’s length/height in relation to age and weight relative to length/height. Anthropometric indicators of nutritional status were generated based on the WHO standards. As expected, Pakistan and the Maldives had higher rates of stunting, wasting, and underweight than Italy and Argentina, and the reverse was true for overweight and obesity. Where stunting was prevalent, the children classified as short were a mean <−2 SD for height-for-age. In all sites, the children classified as thin were indeed wasted (<−2 SD for weight-for-height) and a positive association in trend was evident between weight-for-height and the line-up of groups from thin to obese. The overall concordance between clinical assessments and the WHO standards-based indicators attested to the clinical soundness of the standards. J. Nutr. 137: 149–152, 2007.

Introduction

In April 2006, the WHO released a set of standards for assessing the growth of children <5 y of age (length/height-for-age, weight-for-age, weight-for-length/height, and BMI-for-age) (1). Preliminary standards were field-tested in 4 countries (Argentina, Italy, the Maldives, and Pakistan) at the same time as the technical report describing the standards’ construction methodology (2) was undergoing peer review. The objective of field testing was to evaluate the construct validity of the standards against clinician assessments of children’s attained linear growth and size. Nutritional status classifications, based on length/height-for-age (hereafter, height-for-age), were evaluated against clinician classification of children’s height as short, normal, or tall. Similarly, weight-for-length/height (hereafter, weight-for-height) was evaluated against clinician classification of children’s weight as thin, lean, normal, heavy, overweight, or obese. This article presents results of the field testing in the 4 sites.

Sample and Methods

The sites were selected to provide baseline nutritional status profiles with adequate representation of both undernutrition and overweight. Two affluent populations (Argentina and Italy) and 2 others with high rates of undernutrition (the Maldives and Pakistan) were selected. The field-testing exercise was carried out between January and August of 2005. In Argentina (January–May) and Italy (February–June) assessments were performed by the children’s regular pediatricians, whereas in the Maldives (April) and Pakistan (March–August), the study was conducted by research teams of qualified medical doctors and epidemiologists. All children <5 y of age receiving routine pediatric care from the participating health facilities or clinics were eligible to participate if their parents gave their consent. Sick or preterm-born children were excluded as were those with chromosomal abnormalities, physical disabilities that interfered with anthropometric measurements, endocrine disorders, chronic neurological disease, and dehydration.

The key data collected were weight and length/height, which were measured according to the WHO Multicentre Growth Reference Study protocols (3). The UNISCALE (UNICEF) was used to measure weight in Argentina and the Maldives, whereas different brands of electronic scales were used in Italy and Pakistan. In Argentina, Italy, and Pakistan, length and height were measured on conventional stadiometers or locally constructed equipment verified to be reliable before use in the study.
whereas, in the Maldives, both length and height were measured on Shorr boards (Olney).

For the specific purposes of the study, we developed visual classification codes to describe a child's length/height relative to age (short, normal, or tall) and weight relative to length/height. The latter variable had 6 categories: thin (wasted or too thin), lean (fleshed out, no noticeable fat), normal (rounded contours, no noticeable excess fat), heavy (sturdy, mostly muscular), overweight (noticeable fat), and obese (excess fat). The field exercise in Argentina began before these codes had been defined, so the clinical assessment in that site was open-ended. The teams at the other sites discussed the meaning of each code to standardize the approach to clinical assessment. If necessary, they palpated tissues to determine whether, for example, what appeared to be fat was actually muscular tissue.

The sample in Argentina was recruited from 8 health facilities spread throughout the Municipality of Rosario, in the Province of Santa Fe. During the preparatory phase, data collection instruments were standard-ized and questionnaires translated from English to Spanish with some site-specific adaptations. Data collection forms were returned weekly to the local coordinating center where they were checked for errors before they were double-data entered into the computer. The duplicate files were compared for consistency, and then the data were validated and cleaned on site before being sent to the central coordinating center at the WHO in Geneva.

In Italy, the study was carried out in the southern province of Puglia and involved newborns at Brindisi Hospital and children <5-y-old receiving care from 12 pediatricians throughout the province. The chief nurse at Brindisi Hospital nursery ward measured newborns' weight and length at birth, at 4 d, and at 15 d. These data are excluded from the present analysis because the newborn measurements did not include a clinical assessment. A computer program (Datasus SSLF) developed by a local company was used by the pediatricians to capture and store child health monitoring information gathered during routine pediatric visits. A special subroutine was created within this program to alert the pediatrician when the child visiting was within the age range (1–60 mo) targeted for the survey and to activate the screen for capturing the field-testing data. The field-testing data (ID, date of birth, sex, date of visit, length/height, weight, clinical assessment of weight and length/height, local nutritional status classification, and any known chronic disease) were routed into a special file, ready to be used to compile the study master file. The data relevant to the study were thus automatically extracted (every 2 wk) to compile the master file, carry out quality control checks, and monitor sample accumulation. The principal investigator carried out quality control checks and interacted with the pediatricians and the chief nurse to correct or confirm queried data. The clean master file was sent to the WHO in Geneva.

The Maldivian sample was taken from 5 islands of Raa Atoll. The health center on each island had a list of all children <5-y-old on the island, arranged by age from oldest to youngest. The lists included children displaced from 1 other island by the tsunami in December 2004. Family health workers went through the list systematically, inviting every 2nd child, to achieve the goal of covering 50% of the population on each island. The lists were entered into a template in Excel, checked, and cleaned on the same day they were collected (by A.W.O. and M. O.).

In Pakistan, the sample was recruited at 2 centers in Karachi: the pediatric ambulatory clinic of the Aga Khan University (AKU) Hospital and a clinic serving the slum community of Bilal Colony. Participants targeted for recruitment were randomly selected from among all the <5-y-olds visiting the participating facilities for well-child visits. Upon completing the consultation with primary physicians, participants were assessed by research staff. Undernourished children [defined locally as <5th centile weight-for-age on the National Center for Health Statistics reference (4) at AKU or on the Road to Health chart (5) at Bilal] were booked for re-evaluation after 4 mo of 2-wk follow-ups at home by lady health visitors. A senior medical officer oversaw the study's implementation and quality control. Data forms were returned to the study supervisor daily to be checked for completeness and accuracy. The forms were submitted for data entry at the research coordinating center in the Department of Pediatrics at AKU. Data were entered by 2 operators separately using a special dual-entry program in Visual FOXPRO 6. The program allowed discrepancies and inconsistencies to be identified and corrected. Records flagged in the validation process were retrieved by the data manager who was responsible for checking and correcting any errors. Cleaned data files were converted to SPSS 13 for Windows and sent to the WHO in Geneva.

**Results**

The study captured representative samples of children <5 y of age using the well-child clinic services at each site (Table 1). The median age in the pooled sample was 16.5 mo (mean 21.0 ± 17.1 mo, range 3 d to 60.9 mo). The largest sample was obtained in Pakistan (n = 1673 from AKU and n = 1705 from Bilal), and the smallest in the Maldives (n = 443). The latter, however, represented 51% of the <5-y-old population on the surveyed islands. There was a predominance of males in the sample, especially in the Maldives (55%) and Pakistan (57%).

Nutritional status Z-scores (mean ± SD), for 4 indicators of the WHO standards, indicate a clear dichotomy between Argentina and Italy on the one hand, and Pakistan and the Maldives on the other (Table 1). The Argentinean height-for-age mean was slightly below the WHO standard median, whereas the mean in Italy was equal to the WHO median. The weight-based indicator means for these 2 countries were above the standard median, by nearly half a Z-score for Argentina (weight-for-height and BMI-for-age) and slightly lower for Italy. For the Maldives and Pakistan, the site means on all indicators were below the WHO standard median, in most cases below –1 SD. It is worth noting that weight-for-height and BMI-for-age means were similar within samples, with differences ranging between 0.02 and 0.1 Z-score units.

Prevalence estimates of undernutrition were high in the Maldives (37% stunting, 40% underweight, and 20% wasting) and Pakistan (27% stunting, 20% underweight and 15% wasting) (Fig. 1). Overweight was more prevalent in Argentina (8%) and Italy (6%) compared with the other 2 sites (not >1%). The graphs for wasting, severe wasting, overweight and obesity in Figure 1 were based on weight-for-height. Corresponding prevalence estimates based on BMI-for-age were the same, with the exception of wasting in the Maldives where the prevalence based on BMI-for-age was lower (13.8 vs. 19.6%).

The majority of children in the Maldives (60%) were classified as short during clinical assessment, compared with 19% in Pakistan, 8% in Argentina and 7% in Italy (Table 2). Italy had by far the largest proportion of children described as tall (37%) compared with only 7% in Argentina and <3% in the other 2

**TABLE 1** Basic characteristics and nutritional status indicator Z-scores by site1

<table>
<thead>
<tr>
<th></th>
<th>Argentina</th>
<th>Italy</th>
<th>Maldives</th>
<th>Pakistan</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>1810</td>
<td>2997</td>
<td>443</td>
<td>3378</td>
</tr>
<tr>
<td>Males, %</td>
<td>51.3</td>
<td>52.9</td>
<td>55.1</td>
<td>56.9</td>
</tr>
<tr>
<td>Age, mo</td>
<td>21.7 ± 17.4</td>
<td>22.0 ± 17.5</td>
<td>29.1 ± 17.0</td>
<td>18.8 ± 16.1</td>
</tr>
<tr>
<td>Height-for-age</td>
<td>−0.33 ± 1.06</td>
<td>0.02 ± 1.02</td>
<td>−1.66 ± 1.07</td>
<td>−1.10 ± 1.46</td>
</tr>
<tr>
<td>Z-score</td>
<td>0.18 ± 1.05</td>
<td>0.26 ± 1.00</td>
<td>−1.72 ± 1.07</td>
<td>−1.18 ± 1.29</td>
</tr>
<tr>
<td>Weight-for-age</td>
<td>0.52 ± 1.07</td>
<td>0.40 ± 1.04</td>
<td>−1.11 ± 1.04</td>
<td>−0.78 ± 1.29</td>
</tr>
<tr>
<td>Weight-for-height</td>
<td>0.54 ± 1.07</td>
<td>0.36 ± 1.05</td>
<td>−1.01 ± 0.97</td>
<td>−0.72 ± 1.29</td>
</tr>
</tbody>
</table>

1 Values are means ± SD unless otherwise noted.
sites. The scores for weight relative to length/height described only ~5% of the Maldivian and Pakistani children as heavy, overweight, or obese, compared with 16% in Argentina and 29% in Italy. As for short length/height, the Maldivian sample had the largest proportion of thin (31%) or lean (39%) children, followed by Pakistan (2% thin and 41% lean).

The association between clinical assessments and corresponding anthropometric Z-scores based on the WHO standards are presented in Figures 2 and 3. The mean height-for-age Z-score for the group classified as normal height in Argentina (−0.32) was equal to the Z-score for the full sample in that site (−0.33). In the Maldives and Pakistan, the groups classified as normal had mean Z-score lengths/heights 0.5 and 0.3 taller than the respective site mean height-for-age, whereas their counterparts in Italy were ~0.3 SD below the site mean. Mean height-for-age for the groups classified as short was lowest in Pakistan (−2.69), followed by the Maldives (−2.06), Argentina (−1.47), and Italy (−1.27). Interestingly, the Pakistani tall group was the tallest (height-for-age 1.22 SD) followed by Argentina (0.80) and Italy (0.75), whereas the Maldivian tall group was barely equal to the WHO standard median (height-for-age −0.06 SD). The Pakistani plot had the steepest gradient, reflecting the large SD around the mean height-for-age Z-score for that site (Table 1). The Maldivian plot reflected the overall shift to the left of the <5-year-old population’s height distribution (sample mean −1.66 SD).

The weight code classifications in all sites had a positive association in trend with weight-for-height Z-scores (Fig. 3). The thin groups in all 4 sites had mean weight-for-height <−2 SD, the cut-off for wasting. The lean groups in all sites were close to −1 SD weight-for-height. The groups classified as normal weight had weight-for-height clustered within ±0.5 SD of the standard median, with Argentina and Italy above (0.42 and 0.23 SD, respectively) and Maldives and Pakistan below the median (−0.34 and −0.40 SD, respectively). Those described as heavy were clustered at −1 SD, overweight −2 SD, whereas those described as obese approached 3 SD in weight-for-height. In Argentina, where the clinical assessment was open-ended, mean weight-for-height for the heavy and overweight groups was equal. Overall, all 4 sites showed a consistent linear trend in the means from thin to obese (only 2 children in the Maldivian sample and 1 in Pakistan were classified as obese).

**Discussion**

The 4 sites provided a variety of nutritional profiles with adequate proportions of undernourished and overweight children to permit the classification of observations in the tails of the height-for-age and weight-for-height distributions. The 2 indicators of

---

**Figure 1** Malnutrition prevalence estimates by site.

**Figure 2** Height-for-age mean ± SE by height code by site.

**Figure 3** Weight-for-height mean ± SE by weight code by site.

---

**TABLE 2** Clinical assessment based on length/height relative to age and weight relative to length/height

<table>
<thead>
<tr>
<th></th>
<th>Argentina</th>
<th>Italy</th>
<th>Maldives</th>
<th>Pakistan</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>n</strong></td>
<td>1795</td>
<td>2897</td>
<td>443</td>
<td>3378</td>
</tr>
<tr>
<td><strong>Height classification</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short</td>
<td>143 (8.0)</td>
<td>207 (6.9)</td>
<td>264 (59.6)</td>
<td>632 (18.7)</td>
</tr>
<tr>
<td>Normal</td>
<td>1527 (85.1)</td>
<td>1672 (55.8)</td>
<td>166 (37.5)</td>
<td>2655 (78.6)</td>
</tr>
<tr>
<td>Tall</td>
<td>125 (7.0)</td>
<td>1117 (37.3)</td>
<td>13 (2.9)</td>
<td>91 (2.7)</td>
</tr>
<tr>
<td><strong>Weight classification</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thin</td>
<td>2 (0.1)</td>
<td>9 (0.3)</td>
<td>135 (30.5)</td>
<td>65 (1.9)</td>
</tr>
<tr>
<td>Lean</td>
<td>111 (6.2)</td>
<td>453 (15.1)</td>
<td>173 (39.1)</td>
<td>1374 (40.7)</td>
</tr>
<tr>
<td>Normal</td>
<td>1401 (78.1)</td>
<td>1659 (55.4)</td>
<td>111 (25.1)</td>
<td>1764 (52.2)</td>
</tr>
<tr>
<td>Heavy</td>
<td>54 (3.0)</td>
<td>694 (23.2)</td>
<td>15 (3.4)</td>
<td>165 (4.9)</td>
</tr>
<tr>
<td>Overweight</td>
<td>211 (11.8)</td>
<td>158 (5.3)</td>
<td>7 (1.6)</td>
<td>9 (0.3)</td>
</tr>
<tr>
<td>Obese</td>
<td>16 (0.9)</td>
<td>24 (0.8)</td>
<td>2 (0.5)</td>
<td>1 (0.03)</td>
</tr>
</tbody>
</table>

1 Values are n (%).
2 Thin (wasted or too thin); lean (fleshed out, no noticeable fat); normal (rounded contours, no noticeable excess fat); heavy (sturdy, mostly muscular); overweight (noticeable fat); and obese (excess fat).
weight relative to length/height (i.e., weight-for-height and BMI-for-age) gave similar results. As might be expected, the 2 affluent sites had a higher prevalence of overweight than the less-developed country sites, which had higher rates of stunting, underweight, and wasting than the former. Clinical assessments of weight-for-height were better matched with the WHO standard classification than were assessments of length/height relative to age. The latter assessments were likely complicated by the fact that age is not a self-evident physical attribute and appear, moreover, to have been influenced by the overall size of the index population. Thus, in the population with the highest prevalence of stunting (the Maldives), those classified as tall were barely of mean height according to the WHO standard.

The field test, as part of the development of growth curves, is an innovative feature of the WHO Child Growth Standards. References based on American children have documented the methods used to construct them but none appear to have been field-tested as part of their development (6–9). The U.K. 1990 growth references have been validated following publication but always in comparison with other standards (10–12). Similarly, the Euro-Growth reference curves were compared with a selection of other published references but were not otherwise externally validated (13).

This field-testing exercise complemented the internal validation described in the technical methodology report of the WHO standards (2). Clinical assessment is inherently subjective and length/height assessments were more prone to be influenced by the distribution of attained linear growth in individual sites. However, clinical assessments of weight relative to length/height classified wasting and overweight remarkably well across the 4 sites, and where stunting was a public health problem, groups described as short were indeed stunted (with mean height-for-age Z-score < −2). These findings, and the assessment of the technical reviewers of the WHO standards’ construction methodology, confirmed that the standards were technically and clinically sound. Therefore, the final standards were constructed using models very similar to those used for the preliminary standards (a few of the original models required very minor adjustments in statistical properties) with negligible clinical impact on the resulting curves.

Acknowledgments
The authors thank Cuthberto Garza and Edward Frongillo for reviewing drafts of the field-testing protocol; Unni Silkoset for collaboration in the study, and UNICEF for funding the field work in the Maldives. We also thank Jorge Luna, WHO representative, and Laila Ali for administrative and logistic coordination of the field testing in the Maldives. We also thank the study teams in Pakistan (Drs Erum, Ummai Salwa, Rahila, Salwa, and Huda), Italy (Drs Laura Dell’Edera, Gianfranco Geronimo, Elisabetta Intini, Guiseppe Mattei, Carmela De Robertis, Roberto Gianfredi, Giovanni Simeone, Angela De Pasquale, Oriana Marzano, Maria Rosaria Filograna, Teresa Cazzato, Antonietta D’Oria and Donata Filippo) and Argentina (Drs Humberto Fain, Alicia Orellano, Sara Krupitzky, Alicia Aronna, Mónica Liborio, Mirta Ruiz, Norma Villanova, Mónica García, Marina Duarte, Silvina Feroci, Claudia Constanzo, Maria Gabriela Capiello, Carolina Ares, Andrea Montaner, Laura Fontana, Regina Delgado Pastor, Carolina Binner, Roxana A. Juan, Marcela Scuizzato, Marta Di Benedetto, and Gabriel Chaquí).

Literature Cited