Development of a Research Child Growth Reference and Its Comparison With the Current International Growth Reference

Zuguo Mei, MD, MPH; Ray Yip, MD, MPH; Laurence M. Grummer-Strawn, PhD; Frederick L. Trowbridge, MD, MSc

Objective: To better characterize childhood growth and further assess potential limitations of the current National Center for Health Statistics and World Health Organization international growth reference.

Design: The LMS method was used for curve fitting to summarize the changes in height and weight distributions by 3 curves representing the skewness (L), median (M), and coefficient of variation (S). A series of polynomial regression procedures was applied to smooth the L, M, and S curves.

Setting: Subset data from 18 states contributing clinic data to the Centers for Disease Control and Prevention Pediatric Nutrition Surveillance System were used for this research reference.

Methods: We chose only those clinics in which the height and weight distributions of children closely matched with those of the first and second National Health and Nutrition Examination Surveys.

Results: Unlike the current international growth reference, the new reference has no disjunction at 24 months of age because it is based on a single data source for children aged 0 to 59 months. The reference also better characterizes the growth for infants than the current international reference, a fact we demonstrated with data from the National Health and Nutrition Examination Surveys, Pediatric Nutrition Surveillance System 1995, and the Davis Area Research on Lactation, Infant Nutrition, and Growth studies.

Conclusions: The current National Center for Health Statistics and World Health Organization international growth reference needs to be updated. The methods used in this study will be useful to evaluate other data sets and to evaluate future modifications of growth references.


Editor's Note: Periodically, evaluation tools also need to be evaluated and updated, if indicated. After all, the only sure things in life are death, a polite IRS, and change.

Catherine D. DeAngelis, MD

GROWTH REFERENCES are used widely in the developed and developing worlds, both as a clinical tool to monitor growth in individual children and as a public health indicator to compare nutritional status over time or among different populations.1,2 It is widely accepted that an international reference is useful, since the growth in height and weight of well-fed, healthy children younger than 5 years from different ethnic backgrounds and different continents is reasonably similar.3-5 In 1978, the World Health Organization adopted the growth reference developed by the National Center for Health Statistics based on the US children as an international growth reference both for monitoring individual child growth and for assessing the nutritional status of populations.2 This international growth reference has served a valuable purpose by providing a common basis for the analysis of growth data.

Although the current National Center for Health Statistics and World Health Organization international growth reference continues to serve a unique and valuable role, it has some important technical limitations. The current international reference was developed from 2 data sources. For children younger than 36 months, the length-for-age, weight-for-age, and weight-for-length references were based on data from a longitudinal growth study from 1929 to 1975 of the Fels Research Institute.6-8 For children aged 2 to 18 years, the height-for-age, weight-for-age, and weight-for-height references were based on 3 nationally representative surveys: National...
SUBJECTS AND METHODS

SELECTION OF REFERENCE DATA SET

The NHANES I and NHANES II were conducted by the US National Center for Health Statistics to provide representative data from samples of the civilian, noninstitutionalized US population.16-17 The NHANES I examined persons aged 1 to 74 years during 1971 through 197416,17. NHANES II examined persons aged 6 months to 74 years during 1976 through 1980.17 The problem with using NHANES I and II to create a national reference is that no children younger than 6 months were included in the data sets. Furthermore, the sample size for children aged 6 to 11 months is rather small (only 356 infants).

The CDC PedNSS monitors the general health and nutritional characteristics of low-income US children who participate in publicly funded health and nutrition programs. It provides a framework for tabulating and interpreting state-specific information on the nutritional characteristics of low-income children.16-20 Data for the majority of the infants and children monitored by the PedNSS come from clinic service records of the Special Supplemental Nutrition Program for Women, Infants, and Children Program. This program was initiated in 1972 and is administered by the Food and Nutrition Service of the US Department of Agriculture.21,22 Since its inception, the number of states participating in the PedNSS has increased from 5 in 1973 to 38 states, plus the District of Columbia, Puerto Rico, and seven Indian reservations, in 1995.

According to the protocols used by the Women, Infants, and Children Program, height or length is measured to the nearest 1/8 in or 0.1 cm. A measuring board is used to measure the child’s recumbent length if the child is younger than 24 months; a standing height is measured for children older than 24 months. Weight is measured to the nearest 1/4 lb or 0.1 kg by means of a pediatric scale or other beam balance scale. All the height and weight measurements require trained public health nurses, nutritionists, or dietitians. Also, the protocols require that 2 measurements of length or height of a child agree within 1/4 in, and 2 respective readings in weight agree within 1/4 lb.20,23 All records are entered onto a standardized paper form or onto an automated computer system in the clinics. Once the records are computerized at the state level, they are transferred to the Centers for Disease Control and Prevention for inclusion in the PedNSS database.

A drawback of using the PedNSS database for creating a reference is that the data are not nationally representative. Also, since the data were collected in public health clinics, there is no guarantee that the protocols for measurement techniques were always followed, and thus the accuracy of the measurements may vary. To counteract these drawbacks, we based the new reference on a subset of the PedNSS data from 1975 to 1995 for which we chose only clinics in which the growth of children matched closely with those of national samples for ages 12 to 59 months. Clinics submitting data to the CDC PedNSS were selected only if they met 3 criteria: First, the mean height and weight of the children in the clinic’s population were within ±0.5 cm or kg of the mean from the combined NHANES I and II for every 6-month interval from 12 to 59 months of age. Second, the selected clinics’ population had to have an SD within ±0.2 cm or kg of the SD of the combined NHANES I and II. Third, the skewness in weight distribution of the selected clinics’ population had to be within ±0.3 of the skewness in weight of the combined NHANES I and II. We assumed that the growth of infants aged 0 to 11 months in these clinics was reflective of national patterns.

The final data set for this analysis included 162 163 observations coming from 52 clinics in 18 states (Alabama, Connecticut, Florida, Georgia, Illinois, Indiana, Iowa, Kansas, Kentucky, Massachusetts, Minnesota, New Hampshire, North Carolina, North Dakota, New Jersey, New York, Pennsylvania, and Washington). The sample size for each single month of age in infancy exceeds 2000. Both the height and the weight curves from the PedNSS closely match the NHANES curves for children 12 to 39 months of age and consistent basis for evaluating and comparing growth performance with those of the current international growth reference. The research reference provides a relatively stable and consistent basis for evaluating and comparing growth characteristics of children in research studies. It therefore should help investigators better understand the biologic characteristics of child growth. The research reference may also be useful in evaluating further the limitations of the current international growth reference as well as evaluating future modifications of growth references.
The difference between the medians for the 2 weight-for-age references is relatively small, especially for children younger than 2 years (Figure 3). The 2 references are also similar below the median. Above the median, however, our reference accounts for a greater degree of skewness in the weight distributions. The upper z score curves (except +1 SD curve) of the new reference are slightly higher than those of the current international reference. Also, the current international reference shows an obvious discrepancy at 24 months at upper z score curves, which is consequent to the combination of 2 data sources. Our new curve has no disjunction.

The difference between the 2 references in weight-for-height (length) for boys is shown in Figure 4. The new reference starts at 45 cm rather than 49 cm because the average length of younger infants in developing countries is sometimes less than 49 cm. For weight-for-length (45-84 cm), the new reference curves are consistently higher at the upper z score curves than those of the current international reference for both sexes. For weight-for-height (85-115 cm), the new reference curves are higher at the lower z scores and slightly lower at the
upper z scores compared with those of the current international reference for both sexes. The new curve has a small disjunction at 85 cm to account explicitly for the difference in measured recumbent length and standing stature. However, it is less marked than the disjunction in the current international reference.

We transferred the smoothed L, M, and S curve parameter estimates into statistical analysis system (SAS) codes for height- (length-)for-age, weight-for-age, and weight-for-height (length) for children from birth to 59 months (Table 2). Using these SAS codes, researchers can easily assign z scores and/or percentiles to each child in their studies. In addition, the equations given in the “Subjects and Methods” section of this article can be used to generate percentile or z score growth charts.

VALIDITY OF THE RESEARCH REFERENCE

To compare the research growth reference and the current international reference, we used each reference to assign z scores to the heights and weights of children in 3 data sets: the combined NHANES I and II data, the PedNSS 1995 data, and the DARLING study data.

Comparison of the US National Data

Although the combined NHANES I and II data were used in selecting the PedNSS clinics for this reference, they can also be used to validate the research reference. Figure 5 summarizes the mean height-for-age, weight-for-age, and weight-for-height for the combined NHANES I and II data. Mean z scores based on the research reference show a pattern that more consistently adheres to
the expected 0 $z$ score line. The fact that the combined NHANES I and II data do not adhere to the current international reference is surprising, given that it was created on the basis of NHANES I data and there are no trends in height or weight between NHANES I and NHANES II.

**Comparison of the CDC PedNSS Data**

The second comparison of the current international reference and the research reference was based on the entire year of data of the CDC PedNSS collected in 1995 (Figure 6). The large sample size in PedNSS 1995 data allows us to show mean $z$ score by month of age and thus better demonstrates the marked disjunctions of height-for-age and weight-for-height at 24 months of age for the international reference. Growth status based on the current international reference gives a lower length-for-age $z$ score before 24 months of age but shifts dramatically upward at 24 months. This same shift is also evident for weight-for-height; mean $z$ scores based on the current international reference are higher than our new reference means before 24 months of age. All 3 mean $z$ score curves based on the new reference are more stable across age than are the curves based on the current international reference.

**Comparison of Breast-Fed and Formula-Fed Infants of the DARLING Study**

The infants of the DARLING study were taller and heavier at birth compared with either reference (Figure 7). In the comparison of length-for-age, the research reference consistently resulted in a higher mean $z$ score than the current international reference. Interestingly, there was an important difference in how each reference characterized the mean weight-for-age $z$ score of the 2 infant feeding groups. The current international reference gave a higher $z$ score at early months but became lower by 8 months. This crossover behavior occurred for both breast-fed and formula-fed infants, giving the impression of growth faltering, especially among the breast-fed infants. The absolute decline for breast-fed infants in mean weight-for-age $z$ score from birth to 12 months of age was about 1.2 $z$ scores based on the current international reference, and about 0.7 $z$ scores assigned from the research reference. This finding indicates that neither reference adequately reflects the weight pattern of breast-fed infants, although the research reference is better. In the case of formula-fed infants, the weight-for-age $z$ score from the current international reference implied growth faltering with a net reduction of 0.6 $z$ scores, but the research reference shows no growth faltering. This finding indicates that the current international reference generates an impression of growth faltering among formula-fed infants, whereas the research reference does not.

We also observed a growth faltering in weight-for-length among breast-fed infants, although it was more pronounced for the current international reference (Figure 7). For formula-fed infants, the $z$ scores assigned by both references were generally stable, even though the research reference gave a slightly increasing trend.

**COMMENT**

A reference is defined as a tool for grouping and analyzing data and provides a common basis for comparing populations. In practice, however, researchers use a reference as a standard and, in most cases, for judging the nutritional status of individuals and populations. There-
are not as tightly controlled as in a survey or special study, (2) measurement techniques and equipment calibration for birth to 5 years of age. The PedNSS data have 2 main drawbacks: (1) they represent only low-income children and (2) measurement techniques and equipment calibration are not as tightly controlled as in a survey or special study, and thus the accuracy of measurements may vary. However, by choosing only clinics whose distributions matched the NHANES I and II, we approximated nationally representative data (Figure 1). This similarity supports the validity of the data.

It has been suggested that the current international reference is inadequate for capturing the growth of breastfed infants. In fact, we have shown that it is also in-

Table 2. Statistical Analysis System Code for the Research Reference

<table>
<thead>
<tr>
<th>Variables</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>*AGE</td>
<td>Child's age in months;</td>
</tr>
<tr>
<td>*SEX</td>
<td>Child's gender;</td>
</tr>
<tr>
<td>*HEIGHT</td>
<td>Child's height/length in centimeters; *assuming measure is</td>
</tr>
<tr>
<td></td>
<td>recurrent length before 24 months of age and measure is standing</td>
</tr>
<tr>
<td></td>
<td>height at or above 24 months of age; *if child's age is unknown,</td>
</tr>
<tr>
<td></td>
<td>then measure is recurrent length before 85 cm and standing height</td>
</tr>
<tr>
<td></td>
<td>at or above 85 cm;</td>
</tr>
<tr>
<td>*WEIGHT</td>
<td>Child's weight in kilograms;</td>
</tr>
<tr>
<td>*HAZ</td>
<td>Child's height/length-for-age z score; *assuming measure is</td>
</tr>
<tr>
<td></td>
<td>recurrent length before age 24 months of age and measure is</td>
</tr>
<tr>
<td></td>
<td>standing height before age 24 months of age; if child's age is</td>
</tr>
<tr>
<td></td>
<td>unknown, then measure is standing height at or above 24 months of</td>
</tr>
<tr>
<td></td>
<td>age; *if child's age is unknown, then measure is recurrent length</td>
</tr>
<tr>
<td></td>
<td>before 85 cm and standing height at or above 85 cm;</td>
</tr>
<tr>
<td>*HAPCTILE</td>
<td>Child's height/length-for-age percentile; *assuming measure is</td>
</tr>
<tr>
<td></td>
<td>recurrent length before age 24 months of age and measure is</td>
</tr>
<tr>
<td></td>
<td>standing height before age 24 months of age; if child's age is</td>
</tr>
<tr>
<td></td>
<td>unknown, then measure is standing height at or above 24 months of</td>
</tr>
<tr>
<td></td>
<td>age; *if child's age is unknown, then measure is recurrent length</td>
</tr>
<tr>
<td></td>
<td>before 85 cm and standing height at or above 85 cm;</td>
</tr>
<tr>
<td>*WAZ</td>
<td>Child's weight-for-age z score;</td>
</tr>
<tr>
<td>*WHPCTILE</td>
<td>Child's weight-for-height/length percentile; *assuming measure is</td>
</tr>
<tr>
<td></td>
<td>recurrent length before age 24 months of age and measure is</td>
</tr>
<tr>
<td></td>
<td>standing height before age 24 months of age; if child's age is</td>
</tr>
<tr>
<td></td>
<td>unknown, then measure is standing height at or above 24 months of</td>
</tr>
<tr>
<td></td>
<td>age; *if child's age is unknown, then measure is recurrent length</td>
</tr>
<tr>
<td></td>
<td>before 85 cm and standing height at or above 85 cm;</td>
</tr>
</tbody>
</table>

For Weight-for-Age z Score and Percentile Computations:

IF SEX = 'M' THEN DO;
L = -0.9626601449 - 0.0145968533*(AGE);
M = 3.5768246640 + 0.9754237895*(AGE)
- 0.0039678954*(AGE)**2 + 0.0016168898*(AGE)**3
- 2.184338717E-05*(AGE)**4 + 1.2033018315E-07*(AGE)**5;
S = 0.1556938761 - 0.0044300594*(AGE)
+ 0.001385447*(AGE)**2 - 1.2721539463E-06*(AGE)**3;
END;

IF SEX = 'F' THEN DO;
L = 1.1500832005 - 0.0256066515*(AGE);
M = 3.4282715045 + 0.8474169893*(AGE)
- 0.0385115428*(AGE)**2 + 0.0010697738*(AGE)**3
- 1.4155772287E-05*(AGE)**4 + 7.201043046E-08*(AGE)**5;
S = 0.1450000863 - 0.0022409844*(AGE)
+ 8.715271625E-05*(AGE)**2 - 7.328416562E-07*(AGE)**3;
END;

For Weight-for-Height/Length z Score and Percentile Computations:

IF SEX = 'M' THEN DO;
L = 1.6924354155 + 0.0066400232*(AGE);
M = 50.9260631251 + 3.1504976754*(AGE)
- 0.1527046168*(AGE)**2 + 0.0044238653*(AGE)**3
- 6.370234941E-05*(AGE)**4 + 3.5317693394E-07*(AGE)**5;
S = 0.0527829498 - 0.0015567494*(AGE)
+ 4.4179125095E-05*(AGE)**2 - 4.2738761143E-07*(AGE)**3;
END;

IF SEX = 'F' THEN DO;
L = 1.6954506163 + 0.0018343182*(AGE);
M = 50.1162086261 + 3.1509767754*(AGE)
- 0.1334465191*(AGE)**2 + 0.0037282402*(AGE)**3
- 5.2286854967E-05*(AGE)**4 + 2.854038496E-07*(AGE)**5;
S = 0.0252975448 - 0.0012025441*(AGE)
+ 3.9131714747E-05*(AGE)**2 - 3.7628266632E-07*(AGE)**3;
END;

IF HEIGHT LT 45 OR HEIGHT GT 120 OR AGE = . THEN DO;
HAZ = .; HAPCTILE = .; FOR missing and biologically implausible values;
END;
ELSE DO;
IF (-0.01 < L < 0.01) THEN HAZ = LOG(ADJHT/M)/S;
ELSE HAZ = (WEIGHT/M)**L/(L*S);
HAPCTILE = PROBNORM(HAZ); +PROBNORM is an SAS function for probability from the standard normal distribution;
END;

DROP ADJHT L M S;
adequate for capturing the growth of formula-fed infants. The formula-fed infants in the DARLING study showed a faltering in the weight-for-age $z$ score after 6 months when interpreted with the current international reference. However, when the research reference was applied, the mean weight-for-age $z$ score exhibited a stable pattern without the faltering seen during infancy. The faltering with the current international reference is probably related to problems with the Fels data. The Fels data do not accurately describe the characteristics of infant growth, perhaps because there were too few data observations during infancy (the children were measured at birth and at 1, 3, 6, 9, and 12 months) to fully capture the rapid growth pattern.

Previous studies show that breast-fed children have a rapid decline in the mean weight-for-age compared with formula-fed children during infancy. Our analysis from the DARLING study observed the same growth pattern regardless of which reference was used. However, the current international reference shows a more dramatic decline in the mean weight-for-age $z$ score, giving the impression that breast-fed infants exhibit greater “growth faltering” than formula-fed infants (Figure 7).

The LMS method has been used in Europe to construct several anthropometric references. It has several advantages: the skewness can be incorporated while retaining many of the advantages of a normal distribution; it can generate any required centiles, not just the conventional set of 7 (fifth, 10th, 25th, median, 75th, 90th, and 95th); and individual measurements of height and weight can be accurately and directly converted to $z$ scores or centiles.

The research childhood growth reference was based on a single data source with a large sample. The data were

Figure 5. Mean $z$ scores for the research reference (gray line) and for the current international reference (dashed line), combined National Health and Nutrition Examination Survey I (1971-1974) and II (1976-1980) data.

Figure 6. Mean $z$ scores for the research reference (gray line) and for the current international reference (dashed line), the Pediatric Nutrition Surveillance System 1995 data.
chosen to match the nationally representative US data sets. It eliminates the marked disjunction of height-for-age and weight-for-height z score at the 24-month age point of the current international reference, and it more accurately characterizes growth of formula-fed infants and better captures growth of breast-fed infants.

The principal purpose of making this research reference available is to provide researchers who wish to study changes of growth in infants and young children across age a more accurate tool to characterize such changes. This research reference is not intended to replace the current international reference for routine clinical-based growth monitoring or international comparison. An international working group under the auspices of the World Health Organization is working toward a new international reference to substitute for the current international reference adopted from the United States. This research reference will be useful for the evaluation of the new international reference. The methods and the data source for the research reference development can potentially be useful for the future development of a new US national reference. The recently completed NHANES III survey will be able to provide a more expanded national sample for older infants and children. However, there will still be a gap from birth to 3 months when no national representative data are available, and the methods developed in this article may be useful to fill this gap.

Accepted for publication November 12, 1997.

We extend thanks to all state nutrition programs participating in the CDC PedNSS, especially those that provided the data for this study. We also thank Ellen Borland and Jimmy Simmons for providing data management support and Kelley Scanlon, PhD, and Bettylou Sherry, PhD, for reviewing the document. We are grateful to Kathryn Dewey, PhD, for providing the data from the DARLING study.
REFERENCES


44. Guo S. The Development of Age-Specific Percentiles and the Comparison of Smoothing Procedures: The Revision of the National Center for Health Statistics/ Centers for Disease Control Growth Charts. Hyattsville, Md: National Center for Health Statistics; 1996.


