Scientific and Practical Issues in the Development of the US Childhood Growth Reference


*Maternal and Child Nutrition Branch, Centers for Disease Control and Prevention, Atlanta, Georgia, USA, and Division of Health Examination Statistics and †Office of Research and Methodology, National Center for Health Statistics, Centers for Disease Control and Prevention, Washington, DC, USA

In 1977, the US National Center for Health Statistics (NCHS) published a childhood growth reference for the heights, weights, and head circumferences of children and adolescents from birth to 18 years of age (1,2). This reference, along with the nearly equivalent normalized version of it (3), has become the international reference for assessing the nutritional status of children worldwide.

However, over the years, various limitations of these growth curves have become well recognized (4). First, because the charts for children up to age 3 years were derived from a select sample from Ohio (the Fels Research Institute) and the charts for children and young adults aged 2–20 years were derived from nationally representative data, the curves are not comparable between the ages of 24 and 36 months (5). This disjunction has made it virtually impossible to track the growth of individual children or of populations throughout the first 5 years of life. Second, owing to peculiarities of the Fels sample or the way in which the Fels children were measured or to the curve-fitting procedures in use in 1977, the reference implies a faster rate of growth during infancy than is currently found in other samples of infants (6–8). Finally, the reference does not provide an indicator for assessing relative bodyweight after the age of 10 or 11. With the recent rise in pediatric overweight in the United States (9–11), a national reference for defining overweight is needed.

There is now considerable effort to develop new or revised references that correct these limitations. The World Health Organization (WHO) is developing standard growth curves based on economically advantaged breast-fed children from six countries worldwide (12,13). The US Centers for Disease Control and Prevention (CDC), on the other hand, is developing curves based on representative samples of the US population. The CDC effort is the focus of this chapter.
In 1985–86, the NCHS began the process of revising the US growth curves, by building into the design of the third National Health and Nutrition Examination Survey (NHANES III) the necessary anthropometric measurements and an oversampling of preschool children (14). Workshops were held in 1992, 1994, 1995, and 1997 (15–17), drawing upon the expertise of many governmental and nongovernmental researchers and public health officials. The work of the revision has been carried out by CDC’s NCHS in collaboration with the Division of Nutrition and Physical Activity at the National Center for Chronic Disease Prevention and Health Promotion, with the significant input of Alex Roche and Shumei Guo. The new charts include revision of the set of infant (0–35 months) and child (2+ years) curves developed in 1977: weight for age, weight for length or stature, stature or length for age, and head circumference for age (1). In addition, the new charts add a reference for body mass index (BMI) for age.

This chapter highlights some of the scientific and practical issues faced in creating a growth reference and some of the solutions decided upon by the working group charged with conducting the revision. It does not provide full documentation of all the decisions, evaluations, and the methodology for creating the curves, as this is to be published elsewhere.

USE OF GROWTH CHARTS

Childhood growth curves are used in a wide variety of settings. Pediatric clinicians use them to evaluate the appropriate growth and development of children in their care. Health planners use them to assess the overall health and well-being of populations. Researchers use them to better understand growth and its association with various predictors and consequences. Each of these groups places different requirements for the childhood growth curves. For clinical applications, the development of z-score curves may be unnecessary; however, for the researcher, they are an important tool for describing the status of individual children. For researchers, the differences between curves developed from the general population versus select subsets may be unimportant, as researchers generally examine relative differences between individuals; however, for the health planner, understanding the magnitude of a health problem requires the curves to define appropriately what constitutes abnormality.

The creation of the US growth reference is complicated by the need to address the requirements of these various groups. The working group considered a variety of perspectives and consulted with various outside experts at each juncture in the revision process to ensure that the various requirements were being satisfied.

REFERENCES VERSUS STANDARDS

In developing growth charts, one can take a descriptive approach to generate a “reference,” which describes how children do grow, or one can take a prescriptive approach to generate a “standard,” which describes how children should grow. In a prescriptive approach, one might use well nourished, healthy children who adhere to
established feeding and health care recommendations. Because of this approach, it is assumed that the resulting growth curves will define the bounds of normal, healthy growth and can therefore be used to define growth problems. In a descriptive approach, a representative sample of children in the entire population would be used to describe how children in the population actually grow. As there is no selection of children to ensure that they are healthy, receiving adequate health care, or being fed appropriately, it is impossible to say whether the curves reflect normal growth or not.

The distinction between these two approaches is, in practice, not so clear-cut. Identifying a group of children who adhere to all relevant recommendations and who live in absolutely optimal conditions is practically impossible. Compromises are necessary to find an adequate population on which a growth standard can be developed. Furthermore, health recommendations change from time to time, and so what may be defined as "optimal behavior" now may be different in the future. Also, the distinction between a descriptive reference and a prescriptive standard is not always well understood by users such that the two would be likely to be used in the same way.

The WHO approach is essentially prescriptive, in that it is collecting data on healthy populations that adhere (at least loosely) to international feeding recommendations, most notably exclusive or predominant breast-feeding for 4–6 months and at least 12 months of breast-feeding. The CDC approach, on the other hand, is explicitly descriptive. This descriptive approach drives many of the decisions to be made in creating the curves. In general, children with conditions that might affect their growth are not removed from the charts. For example, data are included irrespective of the economic status of the child or how that child was fed.

However, recognizing that the charts are often used in a prescriptive manner in clinical settings, the working group did make two compromises on developing a "true reference." First, very low-birthweight infants (<1,500 g) were excluded from the curves, as the growth of these infants differs from that of normal-birthweight infants, and alternative charts are available for such infants. Second, weight data for older children from the most recent national survey (NHANES III) were excluded because the inclusion of these data shifted the upper centile curves upward. Also, certain gaps in data around birth have required the use of datasets that are not nationally representative, and so the data around birth do not always represent the entire US population of children.

DATA SOURCES

To develop a truly representative reference of the size of children in the USA, it is necessary to have representative data of the US population. Such data exist in the Health Examination Surveys (HESs) and the NHANESs conducted by the NCHS. These surveys used nationally representative samples of children selected from the civilian, noninstitutionalized population. The children's length or stature, weight, and head circumference were measured in a standardized setting by highly qualified technicians. The second and third HESs (1963–65 and 1966–70, respectively) (18,19) contributed data for children and adolescents of 6–17 years of age.
the first NHANES (1971–74) (20) were used for children of 12 months to young adults of 20 years, and data from the second NHANES (1976–80) (21) were used for the age range 6 months to 20 years. Finally, the third NHANES (1988–94) (14) contributed data for the age range 2 months to 20 years. The sample design for NHANES III included an oversampling of children up to 6 years old.

However, in spite of the use of multiple datasets for creation of the curves, still no nationally representative survey data are available before 2 months of age. Thus, other datasets were incorporated to address the limitations of the survey data.

**SOCIOECONOMIC STATUS**

As the curves are designed to be a reference describing children in the USA, they include children representative of all socioeconomic levels. However, there are important differences in growth by income. Table 1 shows the prevalence of underweight and overweight as well as shortness and tallness among children under age 10 in the NHANES III sample, as calculated using the 1977 normalized reference (3). The prevalence of shortness is increased in the lower-income group, whereas thinness is increased in the upper-income group.

**RACE AND ETHNICITY**

There are also observed differences in growth by race and ethnicity. Mexican–American children show a higher prevalence of overweight, and non-Hispanic blacks show a higher prevalence of tallness (Table 1). Because of these observed race–ethnic differences, the working group considered whether or not to produce separate curves for different race–ethnic groups. The option not to produce separate charts was selected for various reasons. First, current knowledge and expert opinion indicate that almost all children have a similar genetic potential for growth. Second, observed racial–ethnic differences in growth appear to be attributable primarily to

<table>
<thead>
<tr>
<th>TABLE 1. Prevalence of underweight, overweight, shortness, and tallness by ethnicity and poverty status in NHANES III, children aged 2 months to 10 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight for height</td>
</tr>
<tr>
<td>&lt;5th centile</td>
</tr>
<tr>
<td>------------------</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Percent of poverty level</td>
</tr>
<tr>
<td>&lt;100%</td>
</tr>
<tr>
<td>100–199%</td>
</tr>
<tr>
<td>&gt;200%</td>
</tr>
<tr>
<td>Race/ethnicity</td>
</tr>
<tr>
<td>Non-Hispanic white</td>
</tr>
<tr>
<td>Non-Hispanic black</td>
</tr>
<tr>
<td>Mexican–American</td>
</tr>
</tbody>
</table>

NHANES III, third National Health and Nutrition Examination Survey.
environmental influences. Third, the use of such curves would be difficult, if not completely impractical, given significant admixtures of races in the United States.

BREAST-FEEDING

The growth of infants who are breast-fed differs from that of infants fed on formula (7). Breast-fed infants tend to grow more quickly in the first few months but then more slowly for the remainder of the first year of life. This disparity raises the question of how infant feeding should be handled in the growth charts. A growth reference could be based solely on breast-fed children or on a mixture of breast-fed and formula-fed children, or separate charts could be developed for breast-fed and formula-fed children. To represent all children in the USA, both breast-fed and formula-fed infants must be included. Creating separate curves for breast-fed and formula-fed children introduces several practical difficulties. First, what duration and intensity of breast-feeding should be used to define the children in the breast-fed curves? Second, how would a clinician decide which curve to use for children who receive some breast milk and some formula over the first year of life (currently the majority of children)? Finally, unless the breast-feeding curves were to be based on a minimal definition of breast-feeding (e.g., ever breast-fed), there would not be a sufficient sample size in the NCHS surveys to develop stable curves.

It should be noted that whereas the 1977 charts are based on almost exclusively formula-fed infants, the new charts include a large proportion of infants who were ever breast-fed. However, the duration of breast-feeding is relatively short, and less than a quarter of children are exclusively breast-fed for the first 4 months of life.

LOW BIRTHWEIGHT

The growth pattern of low-birthweight infants (<2,500 g) differs from that of normal-birthweight infants (22–25). Numerous investigators have described patterns of catchup growth (26). Because the growth curves might be skewed by the inclusion of low-birthweight infants and because the measurements of such infants can be plotted on curves specially designed for them, the working group considered whether or not the reference curves should exclude low-birthweight infants.

The primary disadvantage of excluding low-birthweight infants is that clinicians must have data on birthweight to decide which curve should be used to plot the growth of children. Often, such information is not readily available, especially for recent immigrants or for children who have repeatedly changed medical care providers. A second problem is that the exclusion of low-birthweight infants creates a truncated distribution of weights at birth, which cannot be described with a parametric form. This implies that the creation of z scores of weight for age is problematic in the first few months of life. Third, alternative charts for low-birthweight infants do not merge with the national curves, partly because catchup growth is not complete and partly because the population and smoothing techniques differ. As a result, there would be a disjunction between the low-birthweight charts and the normal-birthweight charts.
FIG. 1. Weight-for-age z score across age for children born in different birthweight groups, linked natality, and pediatric surveillance data, Tennessee.
at the point at which the low-birthweight charts stop, and the low-birthweight infants would show a sudden change in centile. Fourth, existing low-birthweight charts are for preterm, low-birthweight infants, whereas a high percentage of low-birthweight infants are born full term (27). Finally, the use of two distinct curves implies that children just below the cutoff would be assigned very different centiles compared with children just above the cutoff. For example, an infant who weighed 2,450 g at birth would have a high percentile on a low-birthweight chart, but an infant who weighed 100 g more would have a low percentile on a normal-birthweight chart. In addition to creating an odd discontinuity for clinical practice, this discrepancy would considerably complicate the computation of rates of undernutrition in a population. In the USA, 7.5% of children are classified as having low birthweight (28) and thus would need to use the alternative curves.

To examine the degree to which growth of low-birthweight infants differs from that of very low-birthweight infants (<1,500 g), the working group examined linked natality and pediatric surveillance data in Tennessee. The mean z score for weight for age from birth through 5 years was computed, categorizing children by their birthweight group (Fig. 1). The catchup growth pattern is far more pronounced among very low-birthweight infants than among moderately low-birthweight infants (1,500–2,499 g).

The group also examined the effects of excluding low-birthweight infants, by constructing infant curves with and without these infants included. The effect, even in the lowest centiles, is barely visible. This result is not surprising because many low-birthweight infants experience catchup growth and many normal-birthweight infants do not gain weight as rapidly and thus drop down centiles. These results indicate that the inclusion or exclusion of low-birthweight infants would have a minimal effect on the general curves after the first month of life.

Ultimately, the working group decided to exclude very low-birthweight infants (<1,500 g) from the curves but not infants born at 1,500–2,499 g. In this way, the population affected by the problems described above (roughly 1% in 1998) is minimized. At the same time, this is the population for whom growth patterns in the first year of life are most different, indicating a need to use other curves.

SECULAR TRENDS IN OVERWEIGHT

Over the last several decades, the USA has experienced considerable increases in pediatric obesity (9,10). Increasing weights in children were particularly pronounced between the second and third NHANESs. Because of these trends, there was a concern that growth curves based on the most recent NHANES would show considerably higher centile lines at the 85th and 95th centiles for BMI and weight and that, as a result, clinicians would underdiagnose overweight and obesity in children.

On the other hand, one of the key reasons for revising the NCHS/CDC growth curves was to update the curves with more recent data. Children under the age of 6 years were oversampled in NHANES III, with the explicit purpose of updating the growth curves. Moreover, to directly exclude obese children from the curves would create a truncated distribution of weight status.
A decision made early in the revision process was to include data from all the national surveys that included height and weight measurements on children. To examine the final impact of including NHANES III data in the curves, the working group examined plots of weight for age, with and without the NHANES III data included (Fig. 2). The impact of the NHANES III data is seen for girls at all ages above 48 months but appears only after about age 6 for boys.

**FIG. 2.** Comparison of weight-for-age centiles with (solid) and without (dashed) third National Health and Nutrition Examination Survey (NHANES III) data.
Ultimately, the working group decided to exclude the NHANES III data on weight for children older than 6 years. This cutoff point takes advantage of the oversampling in NHANES III for 3–71 month olds. Starting at age 6, data available from the second and third HESs provide additional data to maintain the requisite sample size for creating precise curves. In this way, the major impact on the final curves of trends in pediatric obesity in the USA has been minimized.

**BMI FOR AGE**

The use of the BMI, defined as the weight (kg) divided by the height squared (m$^2$), has been recommended for use in children and adolescents (29,30), with cutoff points being age specific. The 1977 NCHS charts extend weight for stature only up to 145 cm for boys and 137 cm for girls. In practice, the charts used for girls can be used only up to 10.5 years and for boys only up to 11 years, because children reach sexual maturity around these ages and the relation between weight and height is different before and after maturity. A BMI-for-age reference was developed by Must et al. (31,32) based on NHANES I by year of age from 6 to 19 years. In the new charts, the working group decided that a BMI-for-age reference should also be developed for ages 2–20 years. But should the BMI reference replace the weight-for-stature reference in the younger years, and how young should the BMI reference start?

Smoothing of BMI for ages younger than 2 years is problematic because BMI changes so rapidly, rising during infancy to a peak at 9 months of age and then falling quickly. For children older than 2, there are essentially three options: produce only a BMI-for-age curve; produce only a weight-for-stature curve up to a given age and then switch to BMI for age; produce both weight-for-stature and BMI-for-age curves and let users decide which they prefer to use. The primary problem with the first option is that clinicians caring for younger children are generally not familiar with the BMI, and its calculation requires extra work before plotting values on a chart. The calculation introduces the potential for an additional source of error. Researchers have used BMI as an indicator of adiposity in children, but there has been surprisingly little work comparing its performance with that of weight for stature or its relation to adiposity in preschool children. Requiring that clinicians switch from weight for stature (which is in use now) to BMI for age for assessing adiposity among preschool-age children may be imprudent unless there is evidence that BMI for age is a better indicator. The primary problem with the second option is that switching from one adiposity indicator to another at a given age limits the ability to track adiposity over time in a child. Examination of a child’s changes in BMI during the preschool years might identify children experiencing early adiposity rebound (33,34), who are at increased risk of adult overweight. If only weight-for-stature curves are available for preschool children, assessment of early adiposity rebound is very difficult. The primary problem with the third option is that some clinicians may choose one indicator and some the other, and parents could receive conflicting advice as a result.
CURVE FITTING

Because the data for the growth charts are obtained from sample surveys, it is necessary to smooth the data to create smooth curves that describe the growth of children. Two basic approaches to curve fitting were examined. The first approach, referred to here as the "empirical percentile" approach, entails calculating the centiles observed in the dataset at a series of age intervals. For example, the 5th centile is computed for specific age groups such as 12–15, 15–18, and 18–21 month olds, and so on. This series of numbers is then fitted with either a parametric or a nonparametric smoothing algorithm to generate a smooth curve. In this way, each individual centile curve is fitted independently of the other centile curves (the fitting of the 5th centile is completely unaffected by the fitting of the 10th centile). Because of this independence, it is theoretically possible that the smoothed centile curves could cross over one another.

The second approach smooths the data along the age axis but also along the centile axis. That is to say, a functional form is used to smooth across centiles at a given age. This is, in fact, the approach implied by the normalization of growth curves by Dibley et al. (3). Cole (35,36) has developed a method for curve fitting that simultaneously smoothes across age and centiles. This method, called the LMS method (lambda–mu–sigma, for the three parameters in the normalization function), assumes that the anthropometric variables can be described by a skewed normal distribution. Once the data are transformed by a Box–Cox transformation, they can then be fitted by a normal distribution at each age. The functional form for the LMS is as follows: 

Centile = \mu(1 + \lambda \sigma Z_\alpha)^{1/\lambda},

where \( Z_\alpha \) is the normal equivalent deviate corresponding to the centile and \( \lambda, \mu, \) and \( \sigma \) are parameters to be estimated for each age. The parameter \( \lambda \) reflects the degree of skewness in the distribution, \( \sigma \) reflects the coefficient of variation, and \( \mu \) reflects the median. The key difference between the normalization method of Dibley et al. (3) and the LMS method is that Dibley et al. calculated a standard deviation below the median and another standard deviation above the median, such that there was a distinct change in the spread at a specific point (the 50th centile). The LMS method, on the other hand, changes the spread of the distribution gradually across all the centiles. The most important advantage of the LMS approach is that z scores and exact centiles are readily available.

Both approaches to curve fitting were used and compared, and each showed certain advantages. The empirical centile approach is not dependent on distributional assumptions and is thus more flexible. On the other hand, because the LMS uses data over the whole distribution, the position of the outer centiles is partially controlled by the position of the inner centiles. As a result, the LMS curves in general show a more regular pattern of differences between centiles but missed certain anomalies in the data. The working group took advantage of the strengths of both methods in its final curve-fitting procedure.

DISJUNCTIONS BETWEEN LENGTH AND STATURE

Because children have different heights when measured recumbent versus standing, there is a disjunction between the length-for-age and stature-for-age curves and
between the weight-for-length and weight-for-stature curves. The relation between recumbent length and stature is not constant for all children, so one approach to the two measurements is to treat them as independent measures. Under this approach, the length-for-age curves (0–3 years) and the stature-for-age curves (2–20 years) can be smoothed independently. Likewise, the weight-for-length and weight-for-stature curves can be smoothed independently. As a result of this independence, the difference between length for age and stature for age in the range in which they overlap (24–36 months) is not consistent across the 12-month period or across the centiles.

A second approach is to assume that the difference between recumbent length and stature is fixed for all children and does not vary across the centiles. The amount of difference can be estimated from the second and third NHANESs, which measured children aged 24–36 months both recumbent and supine. The working group is considering an approach similar to that published by Mei et al. (37).

CONCLUSIONS

At first consideration, one might expect that describing the height, weight, and head circumference of US children should be a straightforward task. In practice, however, decisions about what data should be used, who should be included or excluded, whether subpopulations should be lumped together or split apart, what curves should be produced, and how to appropriately smooth data, all have important ramifications for users. In many cases, additional data analyses have been necessary to clarify the issues. Each decision must be considered from a variety of viewpoints to ensure that the needs of all users can be best met. This process is time consuming and has resulted in a later release of the US growth charts than was originally anticipated.

Ultimately, the US growth charts are likely to be used for all children born in the USA, currently nearly 4 million a year. Additionally, many older children will receive new growth charts and be assessed against the new reference. However, discussion about the issues raised in this chapter will continue. With new data available in the coming years from the continuing NHANES (1999+), the appropriateness of the curves will probably be reconsidered by various researchers. The development of new curves by the WHO will also open new discussions in the USA about what reference (or standard) is most appropriate for US children. Thus, the development of the US growth charts must be viewed as a step in a long process rather than an end in itself.

REFERENCES


DISCUSSION

Dr. Ulijaszek: Could you comment on the relative advantages and disadvantages of using the hybrid method? What would Tim Cole say now if he were in this room?

Dr. Grummer-Strawn: I am not sure what Tim Cole would say. He has his own perspective on how the statistics ought to be generated. I think we are talking more of a pragmatic approach than a statistical argument. We felt that before going to the LMS, there were advantages in taking out some of the random variation seen in the LMS parameters. This is very much driven by the view that the centile curves themselves are pre-eminent and the distribution that we find from the data is somewhat secondary—there is a presumption that we are not going to trust the distributional parameters as much as we are going to trust the empirically derived centiles. I think that ultimately we end up with a product that is very similar: a set of parameters that allows us to use the curves in the same way as Tim Cole would be able to use them.

Dr. Koletzko: You pointed out your concern about the disjunction between infant and child curves in the previous version. When I looked at your data carefully, it appears to me that there is now also a disjunction in the 95th centile of BMI between the childhood curves and the accepted standard for defining obesity in adults—that is, a BMI of 30. I suppose the same would apply to weight for age. Have you considered this disjunction? I am concerned that it might lead to underdiagnosis of obese children or adolescents by clinicians using your charts.

Dr. Grummer-Strawn: Yes, we have considered it, and yes, we are concerned about it! Unfortunately, we don’t have any easy answers. One consideration was to draw in a dashed line on the published charts showing the adult standards, so that there is a caveat against continuing to use the 95th centile when it exceeds 30. The feeling of the group in general was that these references are supposed to be a reference, not a clinical guide. How clinicians use them becomes a clinical issue. Another possibility is not to publish the charts beyond 18 years of age, so that a disjunction does not occur. There would be a recommendation that the adult standards from age 20 onward should be applied from age 18 onward. In our clinical guidance that comes out of the Division of Nutrition and Physical Activity, that’s probably the way we are going to go. I think our assumption is that the best thing to do is to continue to use the growth reference until you have reached those adult standards and then switch over. In clinical practice, that is most likely to be the case anyway.

Dr. Molinari: I have looked at many reference curves, but I never see an assessment of the variability of these curves, which is important for creating charts of the different subgroups—the variability of the extreme centiles, for example. Thus, if you look at the curve of the 3rd centile, you don’t know how far away from that curve is the true 1st centile.
Dr. Grummer-Strawn: In other words, when we look at the 95th centile, we don’t know how much error there is associated with drawing that centile. That is true; it’s a good criticism. I expect that one of the papers that will come out after the release of the growth charts will address that issue, but it will not be part of the initial release of these charts. Everything has been done to try to minimize that standard error on each of the charts. First of all, we have included multiple datasets to get the sample size to be as large as possible. Second, the use of the LMS methodology allows us to borrow some of the data from other centiles—whereas the standard error about the 95th centile was initially somewhat wider, this has the impact of making it smaller, because we are borrowing data from other points in the distribution. But I don’t have the data available to tell you how much the variability is. I expect that will be forthcoming in the next couple of years.

Dr. Räähä: You showed low-birthweight infants growing much slower than full-term infants. In our experience, if they are fed human milk that is fortified with optimal amounts of protein and energy, their growth is not slower. When they get near to term, they grow like normal infants. I wonder how your low-birthweight infants were fed.

Dr. Grummer-Strawn: The chart shows a catchup in growth, so there is more rapid growth over the first year. However, these infants still do not catch up completely, even by 5 years of age. They don’t stay at their initial low z score, and they do have more rapid growth, especially during the first year of life, but they never catch up fully. The data I showed here are from, I believe, the late 1980s, so it’s certainly possible that changes in the care of low-birthweight infants will have resulted in somewhat different patterns, but I don’t have those data available.

Dr. Haschke: I’m still not clear about the exclusion of the NHANES III weight data above 6 years of age. You explained that the group feared that the inclusion of the data would result in an underestimation of obesity, but isn’t it cheating not to show what the growth really is? Also, you showed us weight-for-age data with and without the NHANES III included, and these showed some differences in the 95th centile. Have you also got data on BMI? How is the 95th centile influenced in that case?

Dr. Grummer-Strawn: The same set of curves that I showed here for weight for age was also generated for the BMI curves. It gives a very similar picture. Obviously, the shape of the curve is different, but it is the same conclusion: that it is at the upper centiles that you can really see the separation in the BMI. As to your first question, are we cheating ourselves? Yes, that could be said, but this is not purely a reference describing the growth of children in the USA. There has been a very explicit movement away from that initial endpoint of a purely descriptive reference toward more of a standard. I agree it is somewhat odd to pick a dataset and exclude certain data in certain age ranges to generate a different set of curves. But when you consider the different datasets that are included for each of the curves, that’s an oddity in itself. We have different time periods from the USA being used for different ages of children all combined into one standard. I don’t want to defend this as truly the best reference. It was the best that we could do with the data at hand. We felt that, because of its likely clinical use, the standards should not be descriptive of the population as it is. So we did “cheat” somewhat by excluding certain data points, thus pulling the curve down artificially, so that clinical management would be improved.
Dr. Rivera: I really liked the idea that you presented that there is a continuum between a standard and a reference. Being a nutritional epidemiologist in a developing country, and being required often to evaluate the adequacy of growth in the population, I am really glad that you decided to move from a perfect reference population toward a standard population by eliminating the weight data after 6 years from the NHANES III and also the very low-birthweight group. In developing countries, we will be able to consider these charts as prescriptive rather than just descriptive.

Dr. Grummer-Strawn: It is certainly the case now that the WHO growth reference is used as a standard throughout the world as a prescriptive way to describe how children ought to grow. I don’t think the new US growth reference is going to turn out to be substantially different from the current reference. There certainly are some changes in the infant curves: If anything, I think they are probably closer to a prescriptive reference than what we had before. There are certainly more breast-fed children and the duration of breast-feeding is longer, but there is some tendency toward more overweight in these growth curves than there was previously. In the latter sense, it is less prescriptive than it was before. But there has not been a substantial trend toward overweight in the younger population, and the trend toward overweight in the older population was between the second and third NHANESs.

There are more practical questions that we need to consider about changing over to a new reference throughout the world. The WHO is in the midst of a process of developing curves to be as prescriptive as possible. With that on the horizon, it may make more sense, from a practical perspective, to wait to change over to that WHO reference, rather than changing to the CDC reference. I know the WHO will not be sanctioning the CDC reference for international use.

Dr. Martorell: You showed us data that convinced me that you shouldn’t develop reference data for different poverty levels or for different ethnicities, at least the three major ones that you dealt with—non-Hispanic whites, non-Hispanic blacks, and Mexican Americans—but the data you showed us by ethnicity were for preschool children aged 3-5. Have you done this analysis for adolescents? Mexican Americans, for example, end up at the 25th centile at the end of adolescence. I want to know how this issue was dealt with in the adolescent period. There should be enough data on Mexican Americans to allow that to be done.

Dr. Grummer-Strawn: I have not looked at the issues in adolescence. What I showed you were data up to 10 years of age. I think we come back to the same issues we discussed in relation to Dr. Ulijaszek’s paper. Even if there are differences, are they due to the environment or are they due to genetic potential? If they are due to the environment, are we in some sense legitimizing a poor environment by advocating different curves for different subpopulations? This is part of the thinking behind saying that we really need to have one reference for all, because we don’t believe that genetic factors are strong enough to generate large differences.

Dr. Guessry: In one of your first slides, you mentioned optimal growth. I don’t think that the concept of optimal growth is an optimal concept, because when you speak of optimality, you should always refer to function. Optimal for playing professional basketball and optimal for sailing in a pocket submarine are not the same. For us as pediatricians and nutritionists, I think another concept that may be more productive would be the percentage of achievement of your growth potential.

Dr. Grummer-Strawn: The concept defined here was really with regard to inputs rather than outcomes. We are not defining optimal growth in the sense of
consequences. The way in which it gets tied to consequences is that it is tied to practice. It's tied to optimal feeding because we say that breast-feeding is the best way for children to grow, for children's health in general. Therefore, we expect it is best for growth as well. I agree that we do have to ask: Good for what? Is it for health, for working in a factory, for playing basketball? Those are interesting questions. We have defined it just on the basis of the inputs, not the outputs.

**Dr. Bhan:** In a more international sort of perspective, what about children whose families are well-to-do and who don't have any obvious environmental constraints to growth, but who don't consume any animal foods; in other words, predominantly vegetarian populations?

**Dr. Grummer-Strawn:** There was some brief consideration of restrictions based on diet, for example, excluding people on vegetarian diets. But for the reasons already set out, we wanted to create a reference rather than a standard. With that being the overarching starting point, we did not get into the specifics of issues like that.

**Dr. Pelletier:** I found it very useful to know exactly what is in and what is out. The decision to include NHANES III before 72 months and also the decision to use hybrid LMS methodology clearly indicate your desire to preserve the descriptive nature below 72 months. But then after 72 months, you've moved toward more of a prescriptive notion, and it seems to me that this is a policy decision that needs to be reconciled with public health priorities, where obesity is clearly recognized as a priority. We know that BMI tracks from before 72 months on into later childhood and into adulthood, so it seems to me that there are very serious implications of preserving that progression. You mentioned the possibility of putting in a dotted line at the 85th centile, but that seems to me to be a weaker step than what could have been done before 72 months. What was the nature of the policy that essentially legitimizes the overweight/obesity that does exist before 72 months?

**Dr. Grummer-Strawn:** The BMI reference exists from 24 months onward, not just from 72 months, and the possible dotted line really deals only with the 17 to 19 year olds—that is, how to get a continuous movement from the pediatric growth reference into what we recommend for the adult assessment of BMI. One of the rationales for using BMI is to have that continuity of measurement, and yet we don't have continuity of standards. That was the issue I was addressing with regard to the dotted line. Before 72 months, the curves clearly are continuous from 24 months onward; there is no discontinuity at 72 months because of the change in population. The curves are smooth throughout the age range despite the changes in the data that are coming in at each point. Sample sizes are accounted for through appropriate weighting, and essentially at each age we come to a certain description of what is going on at that age and then smooth across the ages. So, we don't have this kind of sudden movement in population from one time point to another.

From a public health perspective, we are very much concerned about pediatric overweight throughout the age range, particularly from age 2 onward, and that is why there is a focus on BMI and tracking. As to the reason for choosing 72 months, it was really a pragmatic question about what was going to affect the curves. As I showed, there is very little impact of including or excluding the NHANES III data before 6 years of age. We needed those data for sample size purposes to reduce the sampling error at that time; we did not feel we were introducing any bias, but we were reducing the variance.