Contents

The WHO Multicentre Growth Reference Study (MGRS): Rationale, planning, and implementation

Mercedes de Onis (WHO), Cutberto Garza (UNU), Cesar G. Victora (Brazil), Maharaj K. Bhan (India), and Kaare R. Norum (Norway), guest editors

Food and Nutrition Bulletin

Food and Nutrition Bulletin, vol. 25, no. 1 (supplement 1)

United Nations University Press
Published by the International Nutrition Foundation for The United Nations University
53-70 Jingumae 5-chome, Shibuya-ku, Tokyo 150-8925, Japan
Tel.: (03) 3499-2811 Fax: (03) 3406-7345
E-mail: mbox@hq.unu.edu
ISSN 0379-5721
Design and production by Digital Design Group, Newton, MA USA
Printed on acid-free paper by Webcom Ltd., Toronto, ON Canada

Preface

"A future of sustainable development begins with safeguarding the health of every child."

Kofi A. Annan
Secretary-General of the United Nations

Among the indisputable rights of children is the right to health. Without respecting this right and providing the necessary resources to secure it, we cannot hope to achieve any of the major development goals the world has united around in the United Nations Millennium Declaration. Human capital is essential to all development. Without basic health and nutrition, the potential of our children goes to waste.

Growth references are among the most valuable and widely used tools we have to measure how we manage to fulfill children's basic physical needs. Of course, assessing growth alone is not enough to adequately evaluate an individual's health status. But his or her physical development is a core element.

The usefulness of growth references, however, stretches far beyond that. Because so many physiological processes must "go right" and so many needs must be met in childhood if growth is to proceed normally, divergences and variations within populations and strata can give useful indications of how supportive the children's surroundings are and even help us track our progress in attaining "health for all." Data collected throughout populations over time can give us important information about their medium- and long-term social and economic development.

Thus, not surprisingly, United Nations and governmental agencies responsible for promoting, securing, and sustaining children's well-being rely on growth references for a wide range of tasks, such as assessing general health status, promoting equity, formulating health and related policies, planning interventions, and monitoring the effectiveness of their efforts and those of others who share commitments and responsibilities to children.

Under the leadership of the World Health Organization (WHO), the United Nations in 1993 undertook a comprehensive review of the uses and interpretation of anthropometric references. As a result of this review, the World Health Assembly (WHA) endorsed the development of a new set of tools to assess infant and young child growth. The Assembly also stressed the need to move beyond past approaches designed to describe *how children grow in a particular region and time* to the more desirable goal of describing *how all children should grow* when their needs are met. In setting this more ambitious goal, the WHA moved beyond recommending the construction of a *reference,* i.e., a device for grouping and analyzing data, to the development of a *standard* (or as close to one as possible), i.e., a device that embodies the concept of a norm or target, thus enabling a value judgment.

To accomplish this more ambitious goal, WHO and its principal partner, the United Nations University, undertook the Multicentre Growth Reference Study (MGRS). At its core was the recruitment of children who met rigorous standards of health. These children not only had to be free of debilitating diseases, but also had to come from families that had conformed with health recommendations in areas such as breastfeeding and smoking cessation.

Emboldened by WHA's commitment, this effort went two steps further. It recruited children from all of the world's major regions to underscore that all children, regardless of ethnic background or regional origin, grow similarly when their needs are met. Moreover, it linked growth measurements to the assessment of motor development. The latter component was facilitated by key support from UNICEF.

By replacing the present international reference, which is based on children from a single country, with one based on an international group of children, we are significantly strengthening the hand of those working to extend the right to health to all children. Similarly, by linking physical growth to motor development, we highlight the very important point that although normal physical growth is a necessary enabler of human development, it is not sufficient. Attention also must be focused on the functional capacities that normal growth makes possible, but does not assure. Together, these three new elements—the "prescriptive" approach that moves beyond the development of growth references to the approximation of standards, the inclusion of children from around the world, and links to motor development—provide us with a much

better instrument to use in our efforts to meet the needs of the world's children. But it also significantly raises expectations of what we should achieve.

This supplement documents the planning, methods, and implementation of the MGRS. The challenges of its adaptation in six distinct sites—Brazil, Ghana, India, Norway, Oman, and the United States—and the creative approaches used to meet them are evident in its contents. Covering five areas, the supplement:

» reviews the rationale for developing a new international set of tools to assess infant and child growth;
» describes the planning, study design, and methodologies adopted to meet the aims of the MGRS;
» reviews the protocols developed to obtain and standardize anthropometric measurements and motor milestones;
» outlines the comprehensive and rigorous data management system designed to assure optimal data quality; and
» systematically considers the site-specific implementation of this global activity.

The outcomes of the MGRS will be scientifically more robust tools to assess child growth than the ones currently available to the international community. Perhaps equally important, these will also be powerful tools for purposes of child health advocacy.

We firmly believe that having tools that provide approximate standards and that are based on children from all of the world's major regions sends crucial messages about aspects of human development that bind all children, political commitments that enable the biological/physical development of individuals and their communities, and responsibilities that are imposed by the last century's remarkable achievements in health, food and agricultural sectors, and information technology.

This project has been a model example of cooperation and collaboration within the UN family and with its external bilateral partners and civil society, and we take pleasure in that fact. Special recognition is due to literally thousands of volunteers and their families who gave freely of their time to this international effort, the principal investigators and their staffs at each of the study sites, and the hundreds of scientists who served as reviewers and in other advisory roles. Special recognition is also due to the WHO Department of Nutrition for Health and Development for its leadership and day-to-day coordination of this activity; the UNU Food and Nutrition Program for its constant support, leadership, and commitment; and the multiple donors who provided vital financial support, encouragement, and intellectual resources for this activity.

LEE Jong-wook
Director-General
World Health Organization

Hans van Ginkel
Rector
United Nations University

Rationale for developing a new international growth reference

Cutberto Garza and Mercedes de Onis, for the WHO Multicentre Growth Reference Study Group

Abstract

The rationale for developing a new international growth reference derived principally from a Working Group on infant growth established by the World Health Organization (WHO) in 1990. It recommended an approach that described how children should grow rather than describing how children grow; that an international sampling frame be used to highlight the similarity in early childhood growth among diverse ethnic groups; that modern analytical methods be exploited; and that links among anthropometric assessments and functional outcomes be included to the fullest possible extent. Upgrading international growth references to resemble standards more closely will assist in monitoring and attaining a wide variety of international goals related to health and other aspects of social equity. In addition to providing scientifically robust tools, a new reference based on a global sample of children whose health needs are met will provide a useful advocacy tool to health-care providers and others with interests in promoting child health.

Key words: Anthropometry, breastfeeding, child health, child nutrition, growth monitoring, growth references, growth standards, human rights

Cutberto Garza is affiliated with the United Nations University and Cornell University, Ithaca, New York, USA. Mercedes de Onis is affiliated with the Department of Nutrition, World Health Organization, Geneva, Switzerland.

Please direct queries to: Mercedes de Onis, Study Coordinator, Department of Nutrition, World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland. Telephone: 41-22-791 3320; fax: 41-22-791 4156; e-mail: deonism@who.int.

Members of the WHO Multicentre Growth Reference Study Group and Acknowledgments are listed at the end of this paper.

Introduction

Growth references are among the most commonly used and most valuable tools for assessing the general well-being of individuals, groups of children, and the communities in which they live, and for tracking progress in reaching a range of health and other, broader goals related to social equity. The value of growth references resides in the fact that numerous physiological processes must proceed normally and many needs must be met in fetal life and childhood if growth is to proceed normally. Thus, although assessing growth is insufficient as a means of adequately evaluating the health status of an individual or a population, normal physical development is a necessary aim of any strategy that includes aspects of well-being as key outcomes. The marked vulnerability of the health of infants and young children also makes assessments of child growth a "sentinel" indicator in evaluations of the health and socioeconomic development of communities in which they live.

The 1993 WHO Expert Committee on the use and interpretation of anthropometry

Given the importance of normal growth as a summary indicator for health, it is clearly within the set of responsibilities of the World Health Organization (WHO) to establish norms for it. In keeping with this normative function, the WHO has periodically convened Working Groups and Expert Committees to examine issues related to anthropometry. The most recent WHO Expert Committee to review this topic was convened in 1993 [1].

In the past, WHO's attention to this topic focused principally on the anthropometry of infants and young children [2–4]. The 1993 Expert Committee, however, was asked to address the use and interpretation of anthropometry throughout the life cycle. This broadened interest signaled an increased appreciation of the utility of anthropometric measurements and indicators

for the screening and evaluation of the health status of individuals and populations of all ages, and the value of changes in anthropometric measurements in the assessment of progress in meeting health, equity, and other societal goals.

WHO organized seven Working Groups in preparation for the 1993 expert consultation to review issues specifically relevant to pregnant and lactating women, the fetus and newborn infant, infants and young children through 2 years of age, children 2 to 10 years of age, adolescents, adults, and the elderly. The Multicentre Growth Reference Study (MGRS) was a direct result of the deliberations of the Working Group on Infant Growth [5].

The WHO Working Group on Infant Growth

The Working Group on Infant Growth was charged with developing recommendations for the appropriate uses and interpretation of anthropometry in infants and young children, i.e., for individuals and populations in diverse operational settings; identifying and/or developing reference data for anthropometric indicators; providing guidelines on their uses; and identifying crucial issues and gaps in knowledge in need of further development. From the beginning of its deliberations, the Working Group focused on incongruities presented by the apparent poor growth of healthy breastfed infants of well-nourished women living in favorable environments. This apparent poor growth was inconsistent with the multiple health benefits associated with breastfeeding and other health behaviors associated with these demographic groups and the environments in which they resided. These inconsistencies focused the Working Group's attention on an evaluation of the current international reference, the US National Center for Health Statistics (NCHS)/WHO International Growth Reference [6], and on a systematic review of the growth performance of breastfed infants studied under relatively highly controlled conditions.

Brief history of the current international reference

The history of the current international growth reference was reviewed in 1996 by de Onis and Yip [7]. This reference was based on a framework initially recommended by the Food and Nutrition Board of the United States National Academy of Sciences [8, 9] and implemented by a joint NCHS and Centers for Disease Control task force [6, 10]. The task force compiled longitudinal data (0 to 23 months) collected by the Ohio Fels Research Institute from various groups of children studied before 1975 and cross-sectional data collected by the US Health Examination Surveys conducted from 1960 to 1975 in children 2 to 18 years of age. The Fels data were collected from predominantly formula-fed infants who resided in a restricted geographic area and were of relatively high socioeconomic backgrounds. The US Health Examination Surveys were designed to reflect representative samples of US children.

During the same period, WHO convened an expert group in 1975 to advise it on the use of anthropometric indicators in nutrition surveys and surveillance activities [11]. This group recommended the use of reference data for these purposes and outlined specific criteria that such data should meet. Although none of the growth data available at that time met the recommended criteria, the NCHS data were recommended by this group for use as the international reference [12]. The major limitations of the infant portion of these data and the reference constructed from them were that the sample was limited to infants of European descent residing in the United States, measurements were taken only every three months, and the analytical methods available at the time were inadequate for the task and were likely to inappropriately depict the pattern and variability of normal growth [7]. The latter two shortcomings contributed to a mischaracterization of the shape of the growth curve, particularly during the first six months when rapid growth occurs, and whose accurate characterization is crucial for effective lactation management.

Summary of the analyses of the Working Group

To review the growth performance of healthy breastfed infants, the Working Group assembled published and unpublished growth data from infants who were exclusively or predominantly breastfed to at least 4 months and who continued breastfeeding for the first 12 months. The Working Group applied fairly conservative criteria to data selection to maximize the likelihood that the growth pattern of the selected sample was not constrained by environmental factors, the nutritional status of the mother, the index pregnancy, or inadequate lactation support. A sample of 226 infants (109 boys and 117 girls) who met the feeding and other criteria outlined above was selected from the larger set of published and unpublished data available to the Working Group. Although this sample had a broader geographic base than the Fels sample, the "pooled breastfed data set" was also from children predominantly of European background and of relatively high socioeconomic status. Additional details of these analyses have been published elsewhere [5, 13].

Among the more salient findings from these analyses are three results particularly relevant to this discussion. First, it was clear that the growth of this conservatively

selected sample of infants deviated negatively from the current international reference and that the magnitude of the deviation was sufficiently large to interfere with nutritional management. The mean Z scores for length-for-age, weight-for-age, and weight-for-length of children 1 to 12 months of age, calculated on the basis of the current international reference, are summarized in figure 1. Rather than the anticipated approximate tracking of early growth trajectories, weight-for-age Z scores fell progressively from months 2 through 12, Z scores for weight-for-length showed a similar pattern, and those for length-for-age fell through 8 months.

To further evaluate the patterns of growth represented by the current international reference and the pooled breastfed data set, the Working Group examined a data set from a WHO Human Reproduction Programme (HRP) study conducted in five countries: Chile, Egypt, Hungary, Kenya, and Thailand [14]. The HRP data set included 1,273 infants whose geographic origins and socioeconomic status were more diverse than those of infants who comprised either the current international reference or the pooled breastfed data set. The Working Group compared the growth of a subset (n = 382) of those infants—those who were either exclusively or predominantly breastfed for various lengths of time through the first year—with both the current international reference and the Working Group's pooled breastfed data set.

The results of the weight-for-age comparison are summarized in figure 2. The Z scores of healthy HRP infants fell from approximately month 3 to months 11 or 12 when the basis for comparison was the current international growth reference, or were sustained or slightly increased when the pooled breastfed data set was used as the reference. The HRP group's declining Z scores relative to the current international reference and its sustained tracking of early Z scores relative to the pooled breastfed data set supported the view that the present international reference was inappropriate for assessing the growth of healthy infants, at least through 12 months of age, and that the growth pattern followed by the pooled breastfed data probably reflected "physiological growth" more closely than did the current international reference.

The third finding was that the variability of growth in the pooled breastfed data set appeared to be significantly smaller than that of the present international reference. These differences were sustained throughout the first 12 months for length and weight in both males and females. The consequences of the decreased variability are illustrated in figure 3. As is evident from this figure, narrowing the distance between the means and the commonly used statistical cutoffs (±2 SD) to identify children at significant risk for either inadequate or excessive growth significantly influences the classification of individual children into either category and estimates of the prevalence of either condition. The narrower variation in the pooled breastfed data set may have resulted from its conservative selection criteria. Alternatively, the wider variation depicted by the current international growth reference may reflect the apparently broad definition of health used to select the Fels population, i.e., the absence of observable illness and the lack of feeding criteria in selecting the study sample. Artificial milks used at the time the Fels data were collected are no longer available, as manufacturers have improved infant formulas steadily. Thus, wider growth variability may have resulted from responses to "nonoptimal" formulas that subsequently were replaced by others that presumably were improved based upon new knowledge of nutritional needs during infancy.

Conclusions of the Working Group

The Working Group's interpretation of these and other related findings outlined in its report to WHO [5] led it to conclude that new references were necessary and that it was time to consider the production of references that would more closely approximate standards, i.e., to describe how children *should* grow in all settings rather than to limit oneself to a description of *how* children grow in a specific setting and time.

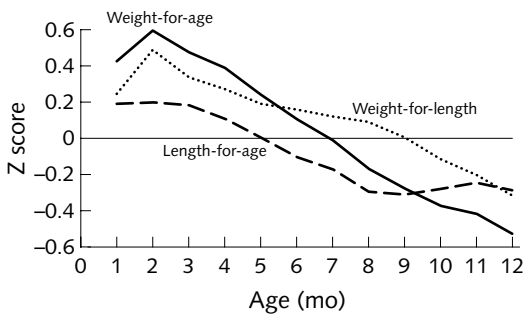


FIG. 1. Mean Z scores of infants in the "pooled breastfed data set" relative to the NCHS/WHO international reference [5]

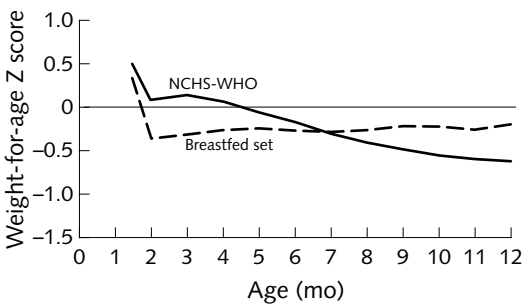


FIG. 2. Mean weight-for-age Z scores of infants enrolled in the WHO Human Reproduction Programme study compared with the NCHS/WHO international reference and the "pooled breastfed data set" [5]

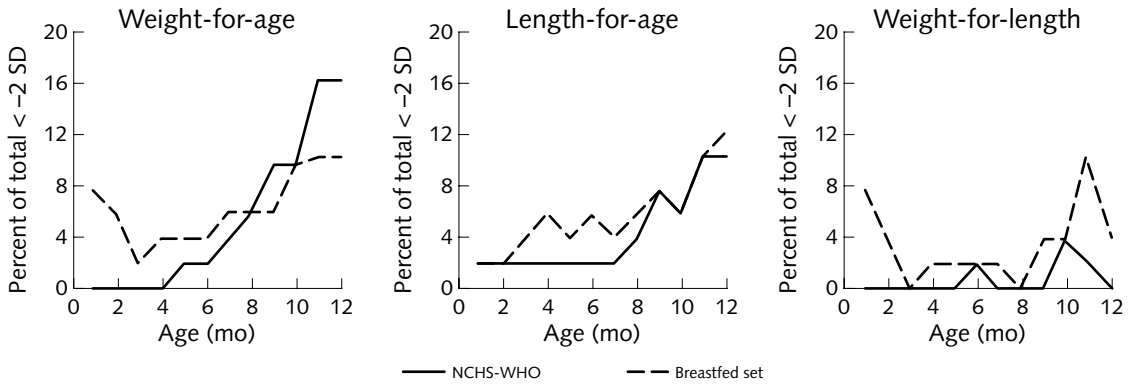


FIG. 3a. Percentages of peri-urban Peruvian infants with weight-for-age, length-for-age, or weight-for-length below the –2 Z score cutoff, according to the NCHS/WHO international reference and the "pooled breastfed data set" [5]

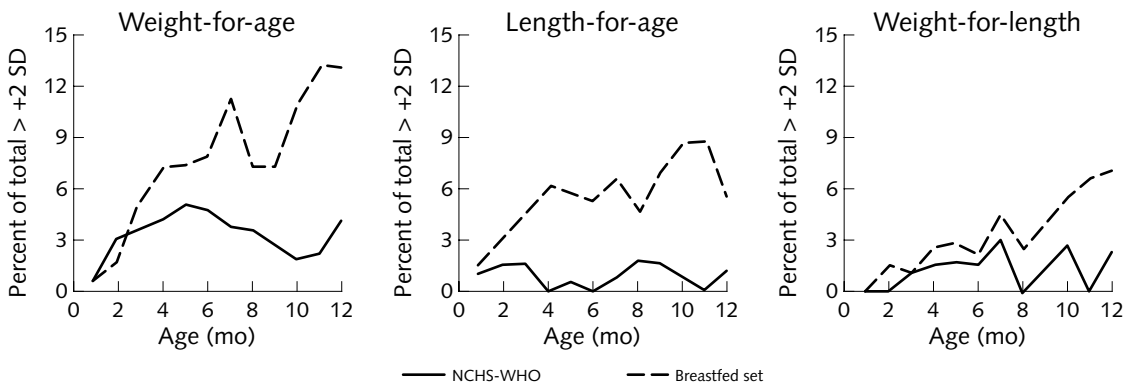


FIG. 3b. Percentages of US and European formula-fed infants with weight-for-age, length-for-age, or weight-for-length above the +2 Z score cutoff, according to the NCHS/WHO international reference and the "pooled breastfed data set" [5]

Three principal lines of thought led to this conclusion. First, the group surmised that at least one key biological assumption inherent in the present international reference is flawed, namely, that infant growth is probably not independent of feeding choices (at least not under conditions that characterized infant feeding choices when the present international reference data were collected). Knowledge of the nutritional, immunological, and reproductive benefits of breastfeeding argues strongly for the breastfed infant as the standard for physiologic growth. The narrower variability estimates derived from the pooled breastfed data set may reflect these biological advantages. The narrow ethnic representation of the pooled breastfed data set is an unsatisfactory explanation for the decreased variability, because of the similarities between the Fels data and the breastfed pooled data set in this regard.

Second, the group recommended that early growth patterns be documented in increments shorter than three months. One possible partial explanation for the deviations between the current international reference and the growth pattern of breastfed infants is that measurements at three-monthly intervals are inadequate to capture the dynamic pattern of growth in the first six months. An accurate depiction of those patterns was viewed as especially important because of the role that growth monitoring plays in lactation management during this period.

Third, the Working Group concluded that limitations inherent to curve-fitting or smoothing techniques available at the time of construction of the present international reference may be an additional explanation for the observed growth discrepancies. Advances in analytical capabilities and approaches have made methods applied to construct the present international reference outmoded.

In response to these findings and recommendations, WHO convened a group in 1995 to develop a protocol for the development of new growth references. Because of the nature of public health programs, WHO asked this second Working Group to consider the inclusion of children through the age of five years.

Ancillary analyses

The deliberations of this second Working Group led to additional analyses that were key to the subsequent

design of the MGRS protocol. The rationale for basing a new reference on breastfed infants was clear and reaffirmed by this group; however, the possibility that other health-related behaviors significantly influenced physiologic growth responses was raised in its discussions. Among the issues of most concern were the timing and nature of complementary feeding, the role of nutrient supplements, and selected parental behaviors, most notably smoking and the use of alcohol and other drugs, and the potential for different growth patterns among breastfed infants of diverse ethnicity. Data from a second HRP data set were used [15, 16] to assess each of these issues.

The results of some of these analyses were published subsequent to their availability for planning the MGRS. Growth patterns of breastfed infants from seven countries (Australia, Chile, China, Guatemala, India, Nigeria, and Sweden) were published in 2000 [17]. Multilevel modeling was used to assess between-site growth differences after adjustments for maternal stature and infant feeding pattern. Approximately 120 infants per site were used for these analyses. Although the study was not restricted to socioeconomically advantaged groups, all women who participated were literate and had educational levels well above the average of their countries of residence. Growth patterns were strikingly similar in all countries except China and India. Maternal education was related to infant growth only in India. All sites were urban except for China. Compared with the arbitrarily selected reference group (Australia), Chinese infants were approximately 3% shorter and Indian infants were approximately 15% lighter at 12 months of age. These analyses demonstrated that breastfed infants from economically privileged families (relative to national norms) were very similar despite the wide ethnic differences and diverse geographic characteristics in this second HRP data set. The findings also underscore the utility of the surveys that were undertaken as a prerequisite to the selection of participating sites. This feature of the MGRS is discussed elsewhere in this supplement [18].

The HRP study was also used to assess associations among growth patterns and different durations of exclusive breastfeeding and the types and frequency of complementary foods introduced between four and six months [19]. Small, statistically significant differences in growth were noted among breastfed infants to whom complementary foods were introduced at different times during that interval; however, the magnitudes of those differences were sufficiently small to be biologically unimportant. The most extreme differences were equivalent to approximately 10 centiles of the weight and height distributions at six months of age. These results provided no compelling evidence of benefit or risk from the timing of complementary feeding between four and six months nor from the frequency or types of complementary foods used during this period by these relatively privileged groups with no major economic constraints and with low rates of infectious illnesses.

In a separate unpublished analysis (report available on request), also based on the second HRP study, associations between the maternal use of alcohol and vitamin or mineral supplements and postnatal infant growth were examined. Alcohol use was examined in the HRP data obtained from Australia, Chile, China, and Sweden. In none of those sites was prenatal or postnatal alcohol use related to postnatal length or weight. The effect of maternal vitamin or mineral supplements was evaluated in the HRP data collected in Australia, Chile, and Sweden. Prenatal or postnatal maternal supplement use was also unrelated to postnatal length or weight in any of those sites.

Maternal use of tobacco was evaluated from the published literature. The second Working Group considered that the effects of smoking on fetal growth [20] and on lactation performance and infant growth [21–23] were important enough to justify inclusion of maternal smoking as an eligibility criterion in the MGRS protocol [18].

Rationale for the MGRS

These analyses, the deliberations of the Working Groups, and extensive peer reviews of the conclusions and recommendations of both Working Groups culminated in the development of a study protocol and operational framework with four salient features: (a) a clearly "prescriptive" approach that included the consideration of infant feeding choices, maternal support for breastfeeding, maternal smoking, and environmental conditions that supported unconstrained physiologic growth; (b) an international sampling frame; (c) heavy reliance on current information technology and its increasing accessibility to document fully the planning and implementation phases of the study, to implement a level of rigor in data management and quality control commensurate with the construction of biological references or standards, and to avoid constraints on the study's selection of analytical methods for curve construction (following a systematic review of contemporary approaches for the analysis of longitudinal and cross-sectional data); and (d) a proposed link between anthropometric assessments and specific functional outcomes of predictive relevance to the well-being of children. This proposed link led to the subsequent addition of the motor development component of the study, which is also described in this supplement [24].

By adopting a "prescriptive" approach, the protocol's design went beyond an update of how children in presumably healthy populations grow at a specific time and place. The MGRS was designed to provide data

that describe "how children *should* grow" by including in the selection criteria of the study specific behaviors that are consistent with current health promotion recommendations (e.g., breastfeeding norms, standard pediatric care, and nonsmoking requirements). Thus, the implemented design advanced beyond the construction of a device for grouping and analyzing data (a *reference*) for the purpose of enabling value-free comparisons, to the explicit recognition of the need for *standards* (or as close to them as possible), i.e., devices that enable value judgments by incorporating norms or targets in their construction.

By including an international sampling frame, the design recognizes the solid evidence that all children grow very similarly for the first five years of life when their physiologic needs are met and their environments support healthy development; nearly all interethnic variability is probably a result of environmental assaults [1, 25–27]. The development of a reference composed of children from all major global regions (in contrast to the present international reference, which is based on children from a single country, the United States) is also likely to be more acceptable for international use. Moreover, it will detract from the perceived need by some to develop country-specific growth norms based on multiple, often inappropriate methods that lead to difficulties in cross-country comparisons and are likely to contribute to faulty national policies. Arguably, the current obesity epidemic in the United States would have been detectable earlier if a prescriptive international reference had been available 20 years ago. An added feature of the design's combined prescriptive and international aspects is the strengthening of advocacy for child health.

Key criteria of reliably robust standards and references are their reproducibility and accessibility to evaluation. Extensive documentation of all stages of development and implementation is indispensable to achieve these characteristics. Among the most important goals of the proposed standards is to remain relevant for as long as possible. This requires that design, implementation, and methodological aspects of sampling strategies, measurements, data management, and analyses be documented as fully as possible. Achieving high standards in the MGRS protocol for each of these features—some of which are reviewed in depth in this supplement [28, 29]—was a key aim. Advances in, and the growing accessibility of, information technology made the task easier to achieve than in the past. The rationale for insisting that these aspects be given scrupulous attention is strengthened by the certainty that knowledge is increasing regarding the functional consequences of early growth patterns and the health behaviors that enable them [30, 31]. Thus, the relevance of MGRS-derived instruments as standards should be amenable to evaluation for the foreseeable future.

Although, as recognized previously, normal growth is necessary to health, it is not sufficient. Interest in growth assessments stems largely from their value as screening tools that signal nonspecific problems when growth is abnormal, or a relative degree of assurance that key physical and emotional needs are being met when growth proceeds as expected. Thus, although normal growth is a necessary enabler of the full complement of functional capacities associated with health, it alone does not assure their attainment. Other resources and conditions, such as educational and physical stimulation within the home, must be accessible to ensure that broader developmental milestones are achieved. This was the basic rationale for the inclusion of motor development assessments in the MGRS. Their broad predictive value and the relative ease with which key motor milestones could be documented in a wide array of field settings supported their inclusion [24]. Linking them closely to anthropometric standards also is expected to be of significant educational value to parents and health-care providers. From a policy perspective, their inclusion is intended to focus attention on growth and broader functional capacities in childhood that are key to normal development in subsequent life stages.

Anticipated results

The MGRS is therefore expected to yield scientifically more robust tools for assessing child growth than are available currently, to strengthen the use of these tools for purposes of child health advocacy, and, because of specific design characteristics discussed in this supplement [18], to provide a wider array of references for expanded uses, e.g., much more appropriate tools for the successful management of early lactation and the monitoring of childhood overweight and obesity.

The current international reference is limited to "attained" measures. This limits the interpretation of anthropometric changes and generally restricts the diagnosis of under- or overnutrition to values that cross a preselected cutoff point assumed to reflect a level of risk for restricted or excessive growth, e.g., the 3rd or 97th centiles, respectively. These are generally interpreted to reflect a level of risk that triggers further evaluation, since only 3% of the target population is expected to be above or below either cutoff; however, for reasons reviewed briefly above, the bases of "value" judgments inherent in such evaluations are problematic, given the "nonprescriptive" nature of sampling schemes upon which the current international reference is based.

International references are currently available only for attained weight-for-age, length/height-for-age, and weight-for-length/height. The MGRS protocol was designed to approximate standards for these and several other attained anthropometric measurements: body mass index (BMI)-for-age, mid-upper-arm

circumference (MUAC)-for-age, head circumference-for-age, subscapular skinfold-for-age, and triceps skinfold-for-age. Additionally, velocity references should be available for a number of growth parameters because of the longitudinal component of the MGRS [18]. Velocity references will most likely be valuable in the early assessment of the risk of overweight in infancy and thus contribute to the early management of this increasingly prevalent and worrisome public health problem. Rather than limiting risk designations to after either state has very likely been achieved, velocity references should enable the identification of children at risk of becoming underweight or overweight. This expanded set of tools is expected to enhance the use and interpretation of anthropometric references, as set out in the 1995 Report of the WHO Expert Committee on this topic [1].

The "prescriptive" sampling scheme described above is also expected to provide improved estimates of the variability of normal growth. These improved estimates should make risk assessments more robust at both the individual and the population levels. If the qualitative differences in variability between the present international reference and the pooled breastfed data set summarized in figure 3 are confirmed by the MGRS, estimates of under- and overnutrition will be impacted, but it is difficult to estimate this quantitatively until analyses of MGRS data are complete.

The selection of breastfed infants as the foundation of new standards also contributes significantly to advocacy in support of current international infant feeding policies [32, 33] and will be much more supportive of lactation management protocols than is the current international reference. The lack of congruence between the feeding histories of infants who contributed to the current international reference and international feeding recommendations unnecessarily sent inconsistent, and potentially confusing, messages. Identifying the breastfed infant as the standard aligns policy with health screening evaluations and potentially provides a goal for manufacturers of infant formula to attain and for national and international regulators to consider in approval processes as new formulations are brought to market.

Broadening the definition of "health" beyond the absence of overt disease to include recommended feeding practices and other health behaviors (e.g., criteria related to maternal smoking behaviors) and selecting infants from populations likely to receive recommended pediatric care should enhance expectations that standards of care and recommended family health-care practices will be accessible to all infants and young children. Tethering such behaviors to the most frequently used health screening tool is thus expected to "raise the bar" substantially in terms of international expectations regarding infant and young child care.

Significance of anticipated results

Upgrading international growth references to tools that more closely resemble standards has substantial significance for other widely accepted international goals. They are expected to make significant contributions to meeting the UN Millennium Development Goals (MDGs) by directly strengthening the framework necessary to achieve them, especially because these new tools are consistent with the human rights approach at the core of the MDGs. The tools will play direct roles at the national, regional, and international levels in monitoring progress toward meeting four of the seven MDGs and, less directly, the remaining three [34]. Although these goals represent a political consensus, and some may question their long-term relevance, the basic aspirations they embody will most likely remain at the core of efforts to narrow social, economic, and health disparities.

Clearly, MDGs such as the eradication of extreme poverty and hunger, achievement of universal primary education, promotion of gender equality and empowerment of women, and reduction in child mortality will each be reflected in improved child well-being. Conversely, improvements in those broad goals will be unattainable unless needs that support normal physical growth are met. Progress in meeting infant and child growth standards will depend significantly on improving maternal health, and so it is likely that these standards also will contribute to the fifth MDG. Similarly, progress in meeting the growth standards will be impossible if we do not succeed in combating HIV/AIDS, malaria, and other diseases, the sixth MDG. Achieving physiologic growth in young children also is linked inextricably to many of the specific aims that comprise the more general MDGs, e.g., ensuring environmental sustainability.

Finally, it is of seminal importance to recognize the basic role that the UN Human Rights Treaty System plays in motivating international aspirations in health and other sectors. The relevance of the MDGs is fully appreciable only within the context of that treaty system. Among the six pillars of the system is the Convention on the Rights of the Child (in force since September 2, 1990).* This convention recognizes duties and obligations to children that cannot be met without attention to normal human development. The use of a growth standard derived from a worldwide sample of children and based on the biological reality that environmental differences rather than genetic endowments are the principal determinants of disparities in physical growth is an important first step in carrying forward our duties and obligations to the human family.

References

1. World Health Organization. Physical status: the use and interpretation of anthropometry. Report of a WHO Expert Committee. Technical Report Series No. 854. Geneva: World Health Organization, 1995.
2. Jelliffe DB. The assessment of the nutritional status of the community. WHO Monograph Series No. 53. Geneva: World Health Organization, 1966.
3. World Health Organization. Methodology of nutritional surveillance. 27th Report of a Joint FAO/UNICEF/WHO Expert Committee. Technical Report Series No. 593. Geneva: World Health Organization, 1976.
4. World Health Organization Working Group. Use and interpretation of anthropometric indicators of nutritional status. Bull WHO 1986;64:929–41.
5. World Health Organization Working Group on Infant Growth. An evaluation of infant growth. Geneva: World Health Organization, 1994.
6. National Center for Health Statistics. Growth curves for children birth–18 years, United States. Vital and Health Statistics, Series 11, No. 165. Department of Health, Education, and Welfare Publication No. 78–1650. Washington, DC: US Government Printing Office, 1977.
7. de Onis M, Yip R. The WHO growth chart: historical considerations and current scientific issues. In: Porrini M, Walter P, eds. Nutrition in pregnancy and growth. Bibl Nutr Dieta 1996;53:74–89.
8. Food and Nutrition Board, National Academy of Sciences–National Research Council. Comparison of body weights and lengths or heights of groups of children. Nutr Rev 1974;32:284–8.
9. Food and Nutrition Board: Committee on Nutrition Advisory to CDC. Comparison of body weights and body heights of groups of children. Atlanta, Ga, USA: US Department of Health, Education, and Welfare, Public Health Service, Centers for Disease Control, 1974.
10. Hamill PVV, Driz TA, Johnson CL, Reed RB, Roche AF, Moore WM. Physical growth: National Center for Health Statistics percentiles. Am J Clin Nutr 1979;32: 607–29.
11. Waterlow JC, Buzina R, Keller W, Lane JM, Nichaman MZ, Tanner JM. The presentation and use of height and weight data for comparing nutritional status of groups of children under the age of 10 years. Bull WHO 1977;55:489–98.
12. World Health Organization. A growth chart for international use in maternal and child health care. Guidelines for primary health care personnel. Geneva: World Health Organization, 1978.
13. World Health Organization Working Group on Infant Growth. An evaluation of infant growth: the use and interpretation of anthropometry in infants. Bull WHO 1995;73:165–74.
14. World Health Organization Task Force for Epidemiological Research on Reproductive Health; Special Programme of Research, Development, and Research Training in Human Reproduction. Progestogen-only contraceptives during lactation: I. Infant growth. Contraception 1994;50:35–53.
15. World Health Organization Task Force on Methods for the Natural Regulation of Fertility. The World Health Organization Multinational Study of Breastfeeding and Lactational Amenorrhea. I. Description of infant feeding patterns and of the return of menses. Fertil Steril 1998;70:448–60.
16. World Health Organization Task Force on Methods for the Natural Regulation of Fertility. The World Health Organization Multinational Study of Breastfeeding and Lactational Amenorrhea. II. Factors associated with the length of amenorrhea. Fertil Steril 1998;70:461–71.
17. World Health Organization Working Group on the Growth Reference Protocol, Task Force on Methods for the Natural Regulation of Fertility. Growth patterns of breastfed infants in seven countries. Acta Paediatr 2000;89:215–22.
18. de Onis M, Garza C, Victora CG, Onyango AW, Frongillo EA, Martines J, for the WHO Multicentre Growth Reference Study Group. The WHO Multicentre Growth Reference Study: planning, study design, and methodology. Food Nutr Bull 2004;25(1)(suppl 1):S15–26.
19. World Health Organization Working Group on the Growth Reference Protocol, Task Force on Methods for the Natural Regulation of Fertility. Growth of healthy infants and the timing, type, and frequency of complementary foods. Am J Clin Nutr 2002;76:620–7.
20. Kramer MS. Determinants of low birth weight: methodological assessment and meta-analysis. Bull WHO 1987;65:663–737.
21. Meyer MB. Breast-feeding and smoking. Lancet 1979;1: 975–6.
22. Vio F, Salazar G, Infante C. Smoking during pregnancy and lactation and its effects on breast-milk volume. Am J Clin Nutr 1991;54:1011–6.
23. Mansbach IK, Greenbaum CW, Sulkes J. Onset and duration of breastfeeding among Israeli mothers: relationships with smoking and type of delivery. Soc Sci Med 1991;33:1391–7.
24. Wijnhoven TMA, de Onis M, Onyango AW, Wang T, Bjoerneboe GEA, Bhandari N, Lartey A, Al Rashidi B, for the WHO Multicentre Growth Reference Study Group. Assessment of gross motor development in the WHO Multicentre Growth Reference Study. Food Nutr Bull 2004;25(1)(suppl 1):S37–45.
25. Habicht JP, Martorell R, Yarbrough C, Malina RM, Klein RE. Height and weight standards for preschool children: How relevant are ethnic differences in growth potential? Lancet 1974;1:611–4.
26. Graitcer PL, Gentry EM. Measuring children: one reference for all. Lancet 1981;2:297–9.

*The six treaties that comprise this system are the Convention on the Elimination of All Forms of Racial Discrimination (in force since January 4, 1969); the International Covenant on Civil and Political Rights (in force since March 23, 1976); the International Covenant on Economic, Social, and Cultural Rights (in force since March 23, 1976); the Convention on the Elimination of Discrimination Against Women (in force since September 3, 1981); the Convention Against Torture (in force since June 26, 1987); and the Convention on the Rights of the Child (in force since September 2, 1990).

27. Martorell R. Child growth retardation: a discussion of its causes and of its relationship to health. In: Baxter KL, Waterlow JC, eds. Nutritional Adaptation in Man. London: John Libbey, 1985:13–30.
28. Onyango AW, Pinol AJ, de Onis M, for the WHO Multicentre Growth Reference Study Group. Managing data for a multicountry longitudinal study: experience from the WHO Multicentre Growth Reference Study. Food Nutr Bull 2004;25(1)(suppl 1):S46–52.
29. de Onis M, Onyango AW, Van den Broeck J, Chumlea WC, Martorell R, for the WHO Multicentre Growth Reference Study Group. Measurement and standardization protocols for anthropometry used in the construction of a new international growth reference. Food Nutr Bull 2004;25(1)(suppl 1):S27–36.
30. Barker, DJ. Programming the baby. In: Barker DJ, ed. Mothers, babies, and disease in later life. London: BMJ Publishing Group, 1994:14–36.
31. Waterland RA, Garza C. Potential mechanisms of metabolic imprinting that lead to chronic disease. Am J Clin Nutr 1999;69:179–97.
32. World Health Organization. The optimal duration of exclusive breastfeeding: a systematic review. Geneva: World Health Organization, 2002.
33. WHO Global Strategy for Infant and Young-Child Feeding. The optimal duration of exclusive breastfeeding, Fifty-fourth WHA A54/INF.DOC/4, May 2001 (Resolution WHA 54.2).
34. UN Millennium Development Goals (MDGs). Available at http://www.un.org/millenniumgoals.

Members of the WHO Multicentre Growth Reference Study Group

Coordinating Team

Mercedes de Onis (Study Coordinator), Adelheid Onyango, Alain Pinol, Trudy Wijnhoven (Department of Nutrition, World Health Organization)

Executive Committee

Cutberto Garza (Chair), Mercedes de Onis, Jose Martines, Reynaldo Martorell, Cesar G. Victora (up to October 2002), Maharaj K. Bhan (from November 2002)

Steering Committee

Coordinating Centre (WHO, Geneva): Mercedes de Onis, Jose Martines, Adelheid Onyango, Alain Pinol

Investigators (by country): Cesar G. Victora, Cora Luiza Araújo (Brazil); Anna Lartey, William B. Owusu (Ghana); Maharaj K. Bhan, Nita Bhandari (India); Kaare R. Norum, Gunn-Elin A. Bjoerneboe (Norway); Ali Jaffer Mohamed (Oman); Kathryn G. Dewey (USA)

Representatives of United Nations Agencies: Cutberto Garza (United Nations University), Krishna Belbase (UNICEF).

Advisory Group

Wm. Cameron Chumlea, Tim Cole, Edward Frongillo, Laurence Grummer-Strawn, Reynaldo Martorell, Roger Shrimpton, Jan Van den Broeck.

Participating countries and investigators

Brazil: Cora Luiza Araújo, Cesar G. Victora, Elaine Albernaz, Elaine Tomasi, Rita de Cássia Fossati da Silveira, Gisele Nader (Departamento de Nutrição and Departamento de Medicina Social, Universidade Federal de Pelotas; and Núcleo de Pediatria and Escola de Psicologia, Universidade Católica de Pelotas)

Ghana: Anna Lartey, William B. Owusu, Isabella Sagoe-Moses, Veronica Gomez, Charles Sagoe-Moses (Department of Nutrition and Food Science, University of Ghana; and Ghana Health Service)

India: Nita Bhandari, Maharaj K. Bhan, Sunita Taneja, Temsunaro Rongsen, Jyotsna Chetia, Pooja Sharma, Rajiv Bahl (All India Institute of Medical Sciences)

Norway: Gunn-Elin A. Bjoerneboe, Anne Baerug, Elisabeth Tufte, Kaare R. Norum (Directorate of Health and Social Affairs; National Breastfeeding Centre, Rikshospitalet University Hospital; and Institute for Nutrition Research, University of Oslo)

Oman: Ali Jaffer Mohamed, Deena Alasfoor, Nitya S. Prakash, Ruth M. Mabry, Hanadi Jamaan Al Rajab, Sahar Abdou Helmi (Ministry of Health)

USA: Kathryn G. Dewey, Laurie A. Nommsen-Rivers, Roberta J. Cohen, M. Jane Heinig (University of California, Davis)

Acknowledgments

The study was supported by the Department of Nutrition of WHO with funding from the Netherlands Minister for Development Cooperation, the Norwegian Royal Ministry of Foreign Affairs, and the United States Department of Agriculture (USDA). Financial support was also provided by the Ministry of Health of Oman, the United States National Institutes of Health, the Brazilian Ministry of Health and Ministry of Science and Technology, the Canadian International Development Agency, the United Nations University,

the Arab Gulf Fund for United Nations Development, the Office of the WHO Representative to India, and the Department of Child and Adolescent Health and Development. The Motor Development Study was partially supported by UNICEF.

First and foremost, we thank the parents and children who participated in the study and the more than 200 field staff on the study teams who made the implementation of this project possible. In particular, we wish to thank the site data management teams, whose diligence has been essential in providing high-quality data throughout the data collection phase. The data managers were Elaine Tomasi (Brazil), Boateng Bannerman (Ghana), Dharmendra Kashyap (India), Karin Rudvin (Norway), Ahmed Al Qasmi and Damodaran Yellapan (Oman), and Laurie Nommsen-Rivers (United States).

We are indebted to the many doctors, nurses, and staff in all the hospitals, clinics, and nursing homes that facilitated the recruitment of newborns to participate in the longitudinal study and the numerous schools and day-care centers that allowed the recruitment of children for the cross-sectional study. Among these are the participating hospitals in Brazil (Santa Casa de Misericórdia, Beneficência Portuguesa and São Francisco de Paula), India (Holy Family Hospital, Mayfair Clinic, Holy Angels Hospital, Sukhda Nursing Home, Sitaram Bhartiya Institute, Moolchand Hospital, Spring Meadows, All India Institute of Medical Sciences, and 66 other hospitals in New Delhi), Ghana (25 hospitals in the Accra/Tema area), Norway (Ullevaal, Aker, and Rikshospitalet University Hospital), Oman (Royal and Khoula hospitals) and United States (Sutter Davis Hospital, Woodland Memorial Hospital, Kaiser Permanente Hospitals of Sacramento, and the University of California at Davis Medical Center).

We thank the lactation counselors for providing support to the mothers who participated in the study. Among these were Gro Nylander, Ragnhild Alquist, Sigfrid Alseth, Anette S. Huitfeldt, Anna-Pia Häggkvist, and Diana C. Tobiesen from the Norwegian team; and Karen Farley, Ann Gorrell, Laura Ortiz, Jeanette Panchula, and Carla Turoff from the US team. We also thank Suman Bhatia, Vinod K. Paul, Ashok K. Deorari, Neena Raina, and the Breastfeeding Promotion Network of India for helping with the training of the lactation counselors and providing ongoing support to the team in India.

We also thank John M. Yeyie, Godlove Addo, Agartha Cofie, and Languon Sebastian from the Ghanaian team; Synnoeve Roedsten, Hilde Nysaether, Ellen Aure, Emmy S. Havdal, Ellinor A. Jensen, Anne-Grethe Krokum, Hilde Loevold, Inger Nevjen, Maren Roedseth, and Sissel Storloes from the Norwegian team; the research assistants from the US team (Swati Deshpande, Cindy Duke, Kathy Harris, Beth Tohill, Nanise Tomlinson, and Tracey Wang); the day-care providers who facilitated the anthropometric standardization sessions at the US site (Leslie Kuss, Julie Cuetara, and the Russell Park Child Development Center); and statistical consultant Janet M. Peerson. We are grateful to Morten Groenn, Erik Boehler, and Jens A. Groegaard for developing the Norwegian criteria on perinatal morbidity; and Knut-Inge Klepp for help in development of the study design in Norway.

We are also grateful to the numerous individuals who with their expertise contributed to the development, review, and refinement of the MGRS protocols. In particular, we thank Ernesto Pollitt from the Department of Pediatrics of the University of California, Davis, USA, for his contribution to the preliminary draft of the motor development protocol. We also thank Noel Cameron, Ray Yip, and Meera Shekar, who contributed to the early stages of the MGRS protocol development. Special thanks also go to Tracey Wang and Maren Rodseth for their assistance in the training of fieldworkers in motor development assessment and to Shaun S'Myth for the milestone drawings.

We thank Eileen Kennedy and Elly Leemhuis, who provided decisive support in the early stages of fundraising for the study. We are also grateful for the support received from the Ministry of Health (Nutrition Division) and the National Programme for Centers of Excellence of Brazil (PRONEX); the Ministry of Health and Family Welfare of the Government of India and the Indian Council of Medical Research; the Norwegian National Nutrition Council; Professors Sefa-Dedeh and Ayernor from the Department of Nutrition and Food Science of the University of Ghana, and the Ghanaian Ministries of Health and Education. In Oman, Ahmed Al Khonji, Director General of the Muscat Region, provided staff for the study and access to Khoula Hospital and health centers in the region, and Ghazi Al Zubaidi also provided staff for the study and gave access to the Royal Hospital. We also gratefully acknowledge the support received from Graeme Clugston, Director of the WHO Department of Nutrition; the WHO country representatives in Brazil, Ghana, India, and Oman; and the WHO Regional Office for South-East Asia.

The WHO Multicentre Growth Reference Study: Planning, study design, and methodology

Mercedes de Onis, Cutberto Garza, Cesar G. Victora, Adelheid W. Onyango, Edward A. Frongillo, and Jose Martines, for the WHO Multicentre Growth Reference Study Group

Abstract

The World Health Organization (WHO) Multicentre Growth Reference Study (MGRS) is a community-based, multicountry project to develop new growth references for infants and young children. The design combines a longitudinal study from birth to 24 months with a cross-sectional study of children aged 18 to 71 months. The pooled sample from the six participating countries (Brazil, Ghana, India, Norway, Oman, and the United States) consists of about 8,500 children. The study subpopulations had socioeconomic conditions favorable to growth, and low mobility, with at least 20% of mothers following feeding recommendations and having access to breastfeeding support. The individual inclusion criteria were absence of health or environmental constraints on growth, adherence to MGRS feeding recommendations, absence of maternal smoking, single term birth, and absence of significant morbidity. In the longitudinal study, mothers and newborns were screened and enrolled at birth and visited at home 21 times: at weeks 1, 2, 4, and 6; monthly from 2 to 12 months; and every 2 months in their second year. In addition to the data collected on anthropometry and motor development, information was gathered on socioeconomic, demographic, and environmental characteristics, perinatal factors, morbidity, and feeding practices. The prescriptive approach taken is expected to provide a single international reference that represents the best description of physiological growth for all children under five years of age and to establish the breastfed infant as the normative model for growth and development.

Key words: Anthropometry, child nutrition, childhood growth, growth curves, growth references, infant feeding practices, infant growth

Introduction

The World Health Organization (WHO), in collaboration with a number of institutions worldwide, is conducting a community-based, multicountry study to develop new growth references for infants and young children, the WHO Multicentre Growth Reference Study (MGRS). The approach taken to develop the new references is fundamentally different from that taken in the past. The new approach describes the growth of children whose care has followed recommended health practices and behaviors associated with healthy outcomes. The new curves may therefore be considered as prescriptive or normative references, as opposed to traditional descriptive references based on geographically representative samples of children, regardless of feeding or other behaviors. The MGRS is taking place in six countries representing the major world regions. This effort involves about 8,500 children and combines a longitudinal study from birth to 24 months with a cross-sectional study of children aged 18 to 71 months. This paper describes the planning, study design, methodology, study organization, and field logistics, and provides an overview of the different phases of the project from its inception in 1990 to its expected completion in 2010.

Mercedes de Onis and Adelheid W. Onyango are affiliated with the Department of Nutrition, World Health Organization, Geneva. Cutberto Garza is affiliated with the United Nations University, Ithaca, New York, USA. Cesar G. Victora is affiliated with the Federal University of Pelotas, Pelotas, Brazil. Edward A. Frongillo is affiliated with the Division of Nutritional Sciences, Cornell University, Ithaca, New York. Jose Martines is affiliated with the Department of Child and Adolescent Health and Development, World Health Organization, Geneva.

Please direct queries to: Mercedes de Onis, Study Coordinator, Department of Nutrition, World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland. Telephone: 41-22-791 3320; fax: 41-22-791 4156; e-mail: deonism@who.int.

Members of the WHO Multicentre Growth Reference Study Group and Acknowledgments are listed at the end of the first paper in this supplement (pp. S13–S14).

Brief history and planning phase of the study

The origins of the MGRS go back to 1990, when the WHO Department of Nutrition established a Working Group to assess the growth patterns of breastfed infants and the relevance of such patterns to the development of growth reference data. The Working Group on Infant Growth was motivated by multiple reports in the literature documenting significant deviations between the growth patterns of healthy breastfed infants and that depicted by the US National Center for Health Statistics (NCHS)/WHO international growth reference. The report of the Working Group was published in 1994 [1, 2]. In its analyses, the Working Group also noted a number of technical problems in the NCHS/WHO international growth reference and concluded that these problems were sufficient to result in potentially harmful decisions in the nutritional management of individual infants and inaccurate population-based assessments.

The group members recommended that a new infant growth reference be developed and that subjects recruited for this purpose should come from populations whose infant-care practices approximated current health recommendations, especially those related to feeding. They further specified that participants in the proposed effort should come from multiple countries, unlike the NCHS/WHO international reference, which is based solely on US children who as infants were predominantly formula-fed [3]. The recommendations of the Working Group were subsequently endorsed by a WHO Expert Committee in 1993 [4, 5] and the World Health Assembly (WHA) in 1994 [6]. The scope and cost of such an ambitious undertaking called for international collaboration. The normative function of WHO placed it in a unique position to provide the leadership required to carry out a project of such complexity and global visibility.

Development of the MGRS protocol

Following the WHA resolution, in 1995 a WHO Working Group on the Growth Reference Protocol was established, formed by pediatricians, nutritionists, human biologists, epidemiologists, and statisticians, to prepare a protocol for the development of a new growth reference based on an international sample of healthy breastfed infants [7–9]. For two years, this group established the framework that resulted in a protocol outlining a fundamentally new approach, prescriptive in nature. Rather than recommending an update of "how children are growing," the group recommended that the reference describe "how children *should* grow." This approach moved past the construction of a device for classifying and analyzing data and allowing the comparison of different populations, to the development of a *standard* (or as close to one as possible), i.e., a device that embodies the concept of a norm or target and thus permits a value judgment. Drafts of the protocol were circulated to numerous external reviewers and presented in scientific meetings and review committees, and an initiative for raising the funds for the study was launched. Reactions from the scientific community as well as from donors were very supportive. However, the high cost of implementation of the study—about 10 million US dollars—represented for some donors too large an investment for a single project. Thus, efforts to raise the necessary funds to support the MGRS have been and continue to be an important aspect of the project's implementation.

Selection of study sites

In 1996, when the main features of the MGRS protocol were settled, we began the process of selecting sites for the implementation of the study. The need to identify sites in each of the six major geographic regions represented a second important challenge in the implementation of the MGRS. The process of selecting the study sites lasted two years and entailed evaluation of specific eligibility criteria for study subpopulations based on the study protocol. Following a presentation of the MGRS at the World Health Assembly, a number of countries expressed an interest in participating in the study. They were requested to send in responses to the checklist of criteria (table 1) documenting the source of the epidemiological data provided.

Since valid epidemiological data were unavailable for some sites to provide information for key items on the checklist, candidate groups were requested to conduct surveys to ascertain the feasibility of carrying out the MGRS. Four surveys were conducted in Asia, one in Africa, and one in the Middle East. The main objective of these surveys was to assess the growth of children living in affluent communities and identify socioeconomic characteristics associated with unconstrained growth in these populations. Information was also gathered on infant feeding patterns, mobility of the population, and other aspects relevant to the protocol. In addition to the survey information and other documentation, candidate sites were visited by members of the Working Group. The final decision about participation was made on the basis of the results of the surveys [10–12] or available epidemiological data from other sources [13], the geographic distribution of the candidate sites, the presence of collaborative institutions able to implement the MGRS protocol, and the availability of national or international funds. The description of the study sites in the six selected countries (Brazil, Ghana, India, Norway, Oman, and the United States) is presented in separate papers in this supplement [14–19] (fig. 1).

TABLE 1. Checklist for the assessment and selection of study sites

Primary criteria	Secondary criteria
Socioeconomic status that does not constrain growth (i.e., epidemiological data showing low infant mortality rate and < 5% prevalence of stunting, wasting, and underweight at 12–23 months of age)	*Rate of hospital deliveries. If home births are frequent, local teams need to prove that obtaining reliable anthropometric measures soon after birth is feasible and that the procedure for identifying newborns in the community does not result in selection biases*
Description of socioeconomic characteristics of study subpopulation	*Sufficient number of eligible births to enroll 300 newborns in 12-month period (at least 7–8 eligible births per week)*
Infant mortality rate in subpopulation	Estimate of the rate of exclusions due to low socioeconomic status, smoking mothers, twins, preterms, etc.
Rates of stunting, wasting, and underweight in subpopulation	Estimated number of monthly births after exclusions
Estimated size of subpopulation	*Mean birthweight in study subpopulation*
Water sources in subpopulation (% with access to safe drinking water)	*Maternal height in study subpopulation*
Low altitude (< 1,500 m)	*Complementary feeding in study subpopulation*
Low mobility of the target population to allow two-year follow-up of children	Energy density of complementary foods
Follow-up rates in previous longitudinal studies	Use of micronutrient supplements (e.g., iron, iodized salt)
Census information on out-migration rates	*Health-related behaviors in study subpopulation*
Minimum of 20% of mothers willing to follow feeding recommendations	Immunization rates
Percentage of mothers in subpopulation who breastfeed for 12 months or more	Pediatric monitoring routines
Percentage of mothers in subpopulation who breastfeed exclusively for 4 months or more	*Environmental hazards*
If these rates are not sufficient, evidence that they could be increased by the study team	Rate of diarrheal diseases
Existence of breastfeeding support system	Presence of significant nonmicrobiological contamination (e.g., exposure to radiation or toxic substances)
Existence of Baby-Friendly Hospitals	*Feasibility of implementing the study protocol*
Description of hospital practices	Sample size calculations
Existence of breastfeeding support groups	Number of hospitals to be surveyed
Presence of experienced lactation consultants	Degree of collaboration from hospitals
Proportion of working mothers and length of maternity leave	Size of geographic area for home visits
Local presence of qualified collaborative institutions	Transportation facilities
Number and qualifications of scientists who will be involved in the study	Location of study headquarters
List of publications of the above scientists	Data entry and management
Description of previous research projects in relevant areas	Estimated costs of study (interviewers, transportation, supervision, lactation support)
Availability of research assistants, interviewers, and data clerks	Rate of refusals among subpopulation in previous studies
Links with other national and international institutes	*Geographic distribution*
Computing facilities	Existence of other candidate sites in the same geographic–ethnic unit
Communications facilities	*Fundability*
	Budget for four-year period
	Likelihood of availability of national or international funds

Preparations for launching the study

During late 1996 and early 1997, the Coordinating Centre, located at the WHO Department of Nutrition in Geneva, prepared the documentation and materials of the study in English, written in great detail, to be used at the study sites for the day-to-day implementation of the study. The documentation included the Manual of Operations, Measurement and Standardization protocols, study questionnaires and interviewer guides, and Data Management protocols (available on request). A training video on anthropometric techniques was prepared for the training and standardization of field staff [20], and a data management system was developed [21]. Study instruments were pretested at the Brazilian site, which served as the pilot site. Study forms and interviewer guides were translated into Arabic, Norwegian, and Portuguese and back-translated into English to ensure that the content of the questions remained unchanged. The only documentation that was developed at a later stage, owing to a shortage of funding, was that related to the Motor Development Study [22]. The late initiation of this study made it impossible for the Brazilian site to participate in this MGRS component. The protocol for the Motor Development Study was pretested at the US site.

While site selection was ongoing, local investigators in confirmed sites proceeded with the recruitment and training of study teams. The planning phase at

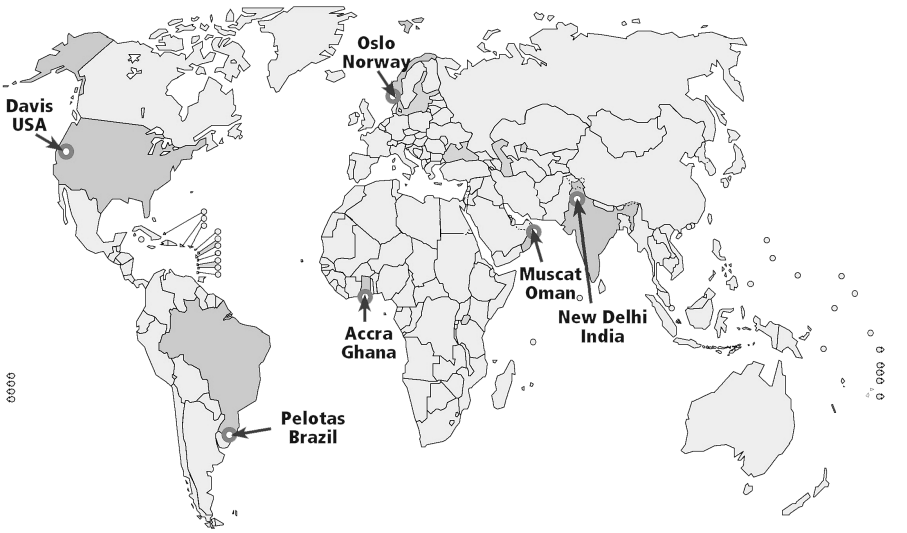


FIG. 1. WHO Multicentre Growth Reference Study map

each study site is described in separate papers in this supplement [14–19]. Intensive exchanges took place between the Coordinating Centre of the MGRS at WHO and the sites to adapt the generic Manual of Operations to local circumstances and to prepare local staff for the launch of the study. Prior to the initiation of data collection, the Coordinating Centre trained and standardized local teams in anthropometric techniques [20], data management [21], and motor development assessment [22].

The planning phase of the MGRS culminated in the enrollment of the first newborn in Pelotas, Brazil, on July 1, 1997. The initiation of data collection elsewhere followed, between 1999 and 2000, according to when sites were identified, local teams were trained and standardized, and funds were identified for the four-year implementation period. Data collection will be completed by November 2003, when the last newborn enrolled in India completes follow-up. The overall project timeline is shown in figure 2. The section that follows describes the study protocol and methods.

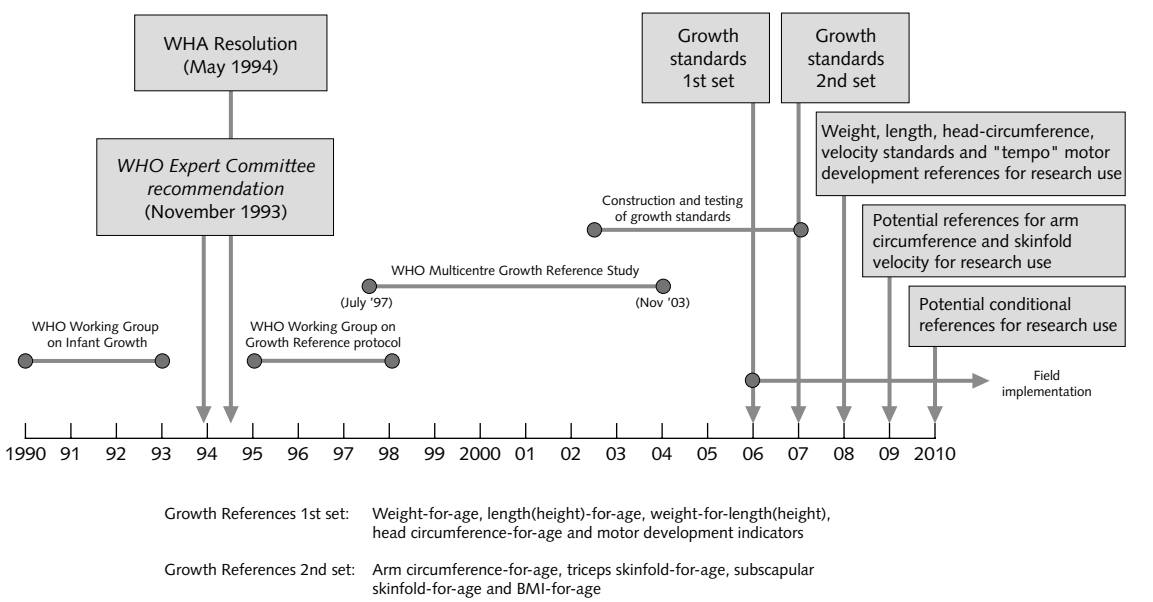


FIG. 2. Timeline of the new international growth references

Methods

Study design

The MGRS design combines a longitudinal study from birth to 24 months with a cross-sectional study of children aged 18 to 71 months. In the longitudinal study, cohorts of newborns were followed for the first two years, with frequent assessments of feeding practices and growth. A longitudinal design for the first two years was needed to provide lactation support to participating mothers, assess selection biases, and provide incremental measurements for the development of growth velocity references. Mothers and children were screened and enrolled at birth and visited at home 21 times: at weeks 1, 2, 4, and 6; monthly from 2 to 12 months; and every two months in the second year. Figure 3 presents the flow chart for the longitudinal study. Mothers enrolled at screening had a two-week period to consider and discuss their participation in the study with their families. Therefore, successful recruitment was determined at the week 2 home visit. Mothers who either refused outright, who posed important restrictions on their participation ("hidden refusals"), or who were found to be ineligible were replaced in the sample. Those who left the study after this point were considered dropouts and were not replaced.

A cross-sectional design was adopted for children aged 18 to 71 months to avoid the time and cost of conducting a longitudinal study in that age range, and also because growth in this age range is more linear than for younger children. Using 18 months as the lower age limit for the cross-sectional study allows an overlap of 6 months with the longitudinal study, which provides information on the transition from supine length to standing height and facilitates the joining of the two data sets. Although the curves will be built for children aged up to 60 months, data collection is extended to 71 months to provide reliable estimates of growth at 60 months (see below). Because of the large number of children required for the cross-sectional study, two sites with small population bases (Brazil and the United States) used a mixed-longitudinal design in which some children were measured two or three times [14, 19].

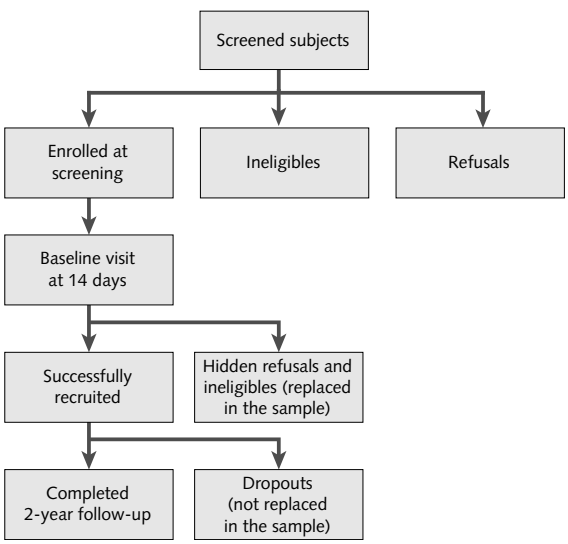


FIG. 3. Flow chart of longitudinal study

The MGRS is a population-based study with well-defined catchment areas from which mother–infant pairs are recruited: the cities of Davis, Muscat, Oslo, and Pelotas and selected affluent neighborhoods of Accra and South Delhi. In all sites, recruitment of infants for the longitudinal study took place in hospitals within 24 hours of birth. The number of participating hospitals was determined to ensure that 80% or more of the population in the designated catchment areas was screened for eligibility. For the cross-sectional study, the sampling strategy was developed according to the circumstances of each site, to provide a sample of children from the same population providing newborns for the longitudinal study [14–19].

A final important feature of the study design is that it pools samples of children who represent a diversity of ethnic backgrounds. The decision to include populations from the major world regions was supported by solid evidence showing that the growth patterns of well-nourished, healthy preschool children across the world are very similar [4, 8]. The surveys conducted as part of the selection process in the developing countries participating in the MGRS demonstrated that this was indeed the case [10–12]. The formulation of a truly international reference is likely to be more acceptable for global use than a reference developed with data obtained from a single country. This procedure averts political concerns that arise from using a single country's child growth pattern as a worldwide standard.

Eligibility criteria for study subpopulations and individual children

The eligibility criteria for study subpopulations were used for selecting the study sites (table 1). It was not necessary for the whole population from the study area to fulfill the criteria, since this restriction would probably have precluded the participation of most sites outside developed countries. These characteristics, however, had to be present among the subpopulations from which study participants were to be drawn. The mean birthweight in the target population was not included as an eligibility criterion; however, it was taken into account when selecting sites.

The eligibility criteria applied to individual mothers and children are listed in table 2. The absence of health, environmental, or economic constraints on

TABLE 2. Eligibility criteria for individual mothers and children

No health, environmental or economic constraints on growth
Mother willing to follow feeding recommendations
Term birth: gestational age ≥ 37 completed weeks (259 days) and < 42 completed weeks (294 days)
Single birth
Absence of significant morbidity
Nonsmoking mother (before and after delivery)

growth was applied as a criterion in the selection of newborns. An objective of the surveys conducted prior to the implementation of the MGRS was to identify socioeconomic factors associated with unconstrained growth in the study subpopulation. Local criteria for screening newborns, based on parental education and/or income levels, were developed accordingly [10–12]. The feeding recommendations with which mothers were required to comply are summarized in table 3. Low-birthweight babies born at term were not excluded, since this restriction would have artificially distorted the lower centiles of the curves in the early months. The list of diagnoses of significant morbidity was developed in consultation with local neonatologists and pediatricians at each site [14–19]. Last, because smoking can affect both lactation performance and infant growth [23–25], as well as birthweight [26], maternal smoking before or after delivery was made an exclusion criterion.

The eligibility criteria were similar for the longitudinal and cross-sectional studies, with the exception of the feeding recommendations, where a minimum duration of three months of any breastfeeding was imposed as an inclusion criterion for the cross-sectional study sample.

Sample size

The precision of growth chart centiles is determined by several factors, of which the most important is sample size. Other relevant factors include study design (cross-sectional versus longitudinal), the timing of measurements, and the method of curve fitting. Four criteria were used to set the sample size for the MGRS: the precision of a given centile at a particular age, the precision of the slope of the median curve over a given age range, the precision of the median curve overall and the influence of data at particular ages, and the precision of the correlation between measurements in the same subjects at different ages. The last criterion is relevant for velocity references. Sample sizes were calculated for each of these four criteria, and it was found that, for each sex, a sample size of 200 for the longitudinal study and 200 per three months for the cross-sectional study would provide adequate precision. These sample sizes were to be obtained by combining data from the six sites.

The sample size calculations yielded the finding that the first few measurements, particularly birthweight, have high variance, whereas between one and four years the variance is low. In addition, limiting the study to children under five years results in increased imprecision during the fifth year. To address the imprecision of the curves at the extremes, birthweight was oversampled and the upper age limit was raised. The sample at birth was increased fourfold, and an upper limit of 71 completed months for the cross-sectional study was implemented to improve the precision of the curves throughout the whole age range of interest.

In the longitudinal study, to obtain 400 children of both sexes, 70 compliant children per site were required to complete the two-year follow-up. The number of newborns to be recruited initially depended on the proportions expected to remain compliant (with feeding recommendations and smoking restrictions) until the age of two years. Based on calculations made from available epidemiological data, the recruitment of a target sample size of 300 newborns per site was set, the only exception being the US site, where the recruitment target was 200 newborns because the expected

TABLE 3. Operational criteria and definitions for compliance to feeding recommendations

Criteria	
Exclusive or predominant breastfeeding for at least 4 months (120 days) Introduction of complementary foods by the age of 6 months (180 days) Partial breastfeeding to be continued for at least 12 months (365 days)	
Definitions	
Exclusive breastfeeding	The infant has received only breastmilk from its mother or a wet nurse, or expressed breastmilk, and no other liquids or solids with the exception of drops or syrups consisting of vitamins, mineral supplements, or medicines
Predominant breastfeeding	The infant's predominant source of nourishment has been breastmilk. However, the infant may also have received water and water-based drinks (e.g., sweetened and flavored water, teas, infusions); fruit juice; oral rehydration salts (ORS) solution; drop and syrup forms of vitamins, minerals and medicines; and ritual fluids (in limited quantities). With the exception of fruit juice and sugar water, no food-based fluid is allowed under this definition

compliance was higher. This total recruitment target fulfilled the requirement that the sample size at birth be at least four times larger than the 400 required at older ages.

To provide similar measurement densities at 18 to 71 months, the cross-sectional study was designed to include the same number of children (70 per three-month period), with each child measured once. The period from 18 to 71 months covers 18 three-month periods, so the nominal sample size required was 70 × 18 = 1,260 children per site. Adding 11% for refusals gave a round sum of 1,400 subjects per site (78 per three-month period). This target sample size was lower for the two sites that used a mixed-longitudinal design, since some children at these sites were measured more than once. Moreover, because the MGRS protocol called for minimizing the number of children participating in both the longitudinal and the cross-sectional samples, the target age interval for the cross-sectional study at the US site was restricted to 27 to 71 months [19]. To fill the gap created by the absence of children in the age range from 24 to 26 months in the US sample, the site in Norway recruited an extra 70 children in this age range.

When the longitudinal cohorts and cross-sectional samples for the six sites are combined, the total MGRS sample size is about 8,500 children. The high compliance and low attrition rates that have been experienced ensure that the new growth curves will be based on a sample size that exceeds the minimum required sample of 200 children for each sex and age group.

Information collected and study questionnaires

The study forms were centrally prepared by the WHO Coordinating Centre accompanied by interviewer guides with detailed instructions for training and field use. The questionnaires included closed questions with precoded answers. In addition to the data collected on anthropometry and motor development, information was gathered on socioeconomic, demographic, and environmental characteristics; perinatal factors; morbidity; and feeding practices. The anthropometric measurements, described in detail in a separate paper [20], included weight, length, height (in the cross-sectional study only), head and arm circumferences, triceps and subscapular skinfold thicknesses, and parental weight and height. Motor development data covered the acquisition of six milestones: sitting without support, hands-and-knees crawling, standing with assistance, walking with assistance, standing alone, and walking alone. The motor development protocol is described in detail in a separate paper in this supplement [22].

All questionnaires were kept as short as possible to improve responsiveness and sample retention. Therefore, all candidate questions were scrutinized initially to ensure that they served at least one of the following purposes: establishing eligibility (e.g., socioeconomic status, intention to breastfeed); describing the sample (e.g., demographic and environmental variables); standardizing findings across centers (e.g., parental height); planning breastfeeding support (e.g., initiation of breastfeeding); assessing continued eligibility (e.g., feeding practices, illnesses); guiding future use of references (e.g., vitamin and mineral supplements); or assessing possible selection biases (e.g., maternal work).

A number of different study forms were used in the longitudinal study:

» A screening form, administered at birth, was used to evaluate eligibility and recruit mothers and newborns. It included data on specific exclusion criteria, such as those related to the family's socioeconomic status, the mother's intention to breastfeed, the newborn's gestational age, and maternal smoking behavior.
» A breastfeeding-in-hospital form, which described breastfeeding initiation, timing, and pattern.
» Four breastfeeding-at-home forms were used at weeks 1 and 2 and months 3 and 6. Information was collected on the establishment of lactation, problems experienced in the first two weeks (such as delayed onset of milk production and breast infections), and practices with potentially adverse influences on continued lactation (such as pacifier use and contraception).
» A baseline form administered at the day 14 visit collected information on socioeconomic, demographic, and environmental factors; pregnancy history; and parental anthropometry.
» The follow-up questionnaire was administered at each of 20 follow-up visits to record detailed information on feeding patterns (including the 24-hour dietary recall for the preceding day); maternal and child morbidity; use of vitamin and mineral supplements; maternal employment, smoking, and weight; and child anthropometry.
» For motor development, as many as 14 forms were completed in months 5 to 24, but children who could walk independently before the age of 24 months required the completion of fewer forms. All six milestones were assessed on each occasion.
» An end-of-participation form specifying the reason for ending participation was completed for all subjects who were recruited at the initial screening. Possible reasons included ineligibility or refusal established at the day 14 visit, reasons for dropping out from the study on a later occasion, and the end of follow-up for those who successfully completed the study.
» A 12-month-visit questionnaire was administered to mothers who, although eligible, did not intend to breastfeed; who refused to participate in the study

at any stage; or who dropped out of the study before 12 months for reasons other than child illness. The form gathered selected anthropometric data and information related to the child's morbidity and feeding history.

The cross-sectional study used three study forms:

» A screening form collected information used to establish eligibility on variables similar to those used in screening for the longitudinal study.
» A survey form covered socioeconomic and demographic factors, child feeding history and morbidity, and parental and child anthropometry.
» In the context of the mixed-longitudinal design, one or two follow-up forms (abbreviated versions of the survey form described above) were used in Brazil and the United States to gather data on anthropometry and child morbidity in the intervals between visits.

Quality control

Rigorous scientific standards have been applied to this complex, multicountry, field-based project. This section summarizes the main measures taken to ensure data quality, most of which are detailed further elsewhere in this supplement [20–22]. Quality control measures included the following:

» Pilot testing of study protocol;
» Use of pretested, standardized data collection forms and detailed interviewer guides;
» Translation into local languages and back-translation of questionnaires and other forms;
» Careful selection, thorough training, and close supervision of staff;
» Regular visits to study sites;
» Training on anthropometric measurements and motor development assessment by international experts with annual site visits by the experts for standardization and/or retraining purposes;
» Regular standardization sessions throughout data collection, with assessment of intra- and interobserver reliability [20, 22];
» Specially designed and highly reliable measuring equipment that was calibrated frequently [20];
» Coordination meetings and staff exchanges among sites;
» Continuous data quality assurance from the point of data collection (independent measurements by two standardized observers [20]), through all stages of data management to their incorporation into the MGRS master files [21];
» Repetition of 10% of all interviews on the telephone;
» Continuous central monitoring of the timing of visits (including delayed, advanced, or missed visits), frequency of repeated measurements, missed measurements, investigation of outliers, terminal digit preference, and results of anthropometric and motor development standardization sessions.

The monitoring of data quality was effective in identifying deviations from MGRS standards, and early, appropriate remedial measures were taken.

Data management and analysis

The MGRS data management system is described in full elsewhere in this supplement [21]. Data were entered concurrently with data collection, verified and validated at the study sites, and sent on a monthly basis to the Coordinating Centre at WHO. MGRS master files were consolidated and ongoing data quality control analyses were carried out at the Coordinating Centre to monitor study implementation and assess adherence to the study protocol.

All data analyses will be conducted at the Coordinating Centre. The Coordinating Centre will be responsible for constructing the new growth references using state-of-the-art statistical techniques. In preparation for the analysis phase, a review of the different methods for the construction of distance, velocity, and conditional growth references was recently conducted by WHO. A full description of the 30 methods reviewed is beyond the scope of this paper. The review document was circulated for external peer review and discussed at a WHO meeting of an ad hoc statistical advisory group. The group identified several criteria for assessing the different methods (e.g., distributional assumptions, curve fitting, age handling, and model simplicity) and, based on these criteria, selected methods to be tested for the growth parameters included in the MGRS. Model diagnostic tools for assessing the appropriateness of the selected methods were also identified. Given the numerous sets of growth reference data that will be produced—including novel references based on circumferences, skinfolds, and growth velocity—the construction and testing of the various references promises to be a complex and challenging task.

Methodological issues

An important concern when proposing a reference based on recommended practices is how such restrictions may affect other characteristics of the reference sample. For example, mothers who choose to breastfeed exclusively or predominantly may also present behaviors other than feeding choices that influence child growth. If a reference population is overly homogeneous, the distribution of values will be too narrow, resulting in statistically based cutoffs that are closer to the mean than would occur in an appropriately heterogeneous reference population.

In response to the concern that the prescriptive approach taken for the development of the new reference might result in an excessive degree of sample selectivity, measures were built into the study protocol

to minimize bias and assess the potential influence of selection bias on the outcomes of interest:

Measures to minimize inappropriate sample selectivity

» *Implementation of the study in sites where at least 20% of the mothers in the study subpopulation were likely to comply with the feeding recommendations of the study (tables 1 and 3).*

» *Application of operational definitions of feeding recommendations that would allow a greater proportion of children to be included in the growth reference data set.* Some flexibility in the operational definitions was expected to reduce selectivity problems with the cohorts to be followed and to lessen economic and logistic constraints. Furthermore, available evidence and analyses conducted during development of the MGRS protocol indicated that there were small, if any, differences between the growth of exclusively and predominantly breastfed infants in the first six months of life [8, 27] and that postnatal growth did not appear to be very sensitive to the differential timing of introduction of complementary foods among healthy infants living in safe environments [9, 28]. It was therefore decided that, for the purpose of constructing the growth curves, the feeding criteria to be used would be those listed in table 3. However, at the field level, mothers participating in the MGRS would be advised to breastfeed their infants exclusively for as close as possible to six months, with introduction of complementary foods by the six-month visit.

» *Provision of intensive breastfeeding support to participating mothers to enhance compliance and reduce selection bias by ensuring a high level of compliance with feeding recommendations.* To allow a high proportion of mothers wishing to breastfeed to actually do so, lactation counseling was made an essential part of the MGRS. At each site, trained counselors visited participating mothers frequently in the first months after delivery to help successful breastfeeding initiation and to advise on subsequent problems. The first visit took place within 24 hours of delivery, and subsequent visits were made at 7, 14, and 30 days, and then monthly thereafter until at least the sixth month. Additional visits were carried out whenever feeding problems occurred. A 24-hour hotline also was made available to mothers for emergency support. Mothers also received advice on complementary foods—with emphasis on energy density, feeding frequency, and micronutrient content—according to locally adapted complementary feeding guidelines. Descriptions of the local lactation counseling teams and complementary feeding guidelines are provided elsewhere in this supplement [14–19].

Compliance with feeding recommendations was monitored centrally throughout the study, and lactation counseling was strengthened as required. Preliminary results strongly suggest that the above measures have been effective and that compliance rates across sites have been high, minimizing concerns about the selectivity of the MGRS sample.

Measures to assess sample biases

Two key measures were included in the study protocol to permit the assessment of possible selection biases affecting the sample:

» *Follow-up of the entire cohort independent of compliance status.* This allows the comparison of the patterns of growth of children whose mothers complied with the feeding and smoking recommendations with those who entered the study but whose mothers subsequently failed to comply with the recommendations of the study.

» *The 12-month study.* This substudy involved visiting a sample of eligible nonparticipating infants on their first birthdays to compare their attained weights and lengths with those of the cohort children. Four categories of children were included in this substudy: those whose mothers refused to participate at screening; those whose mothers did not intend to follow the feeding recommendations at screening; those excluded at the day 14 visit because the mother had started feeding other milks; and those who dropped out of the study before the age of 12 months.

Study organization and field logistics

Study organization

The study organization is presented in figure 4. The study was initiated, coordinated, and managed by the Department of Nutrition of WHO, where the MGRS Coordinating Centre was located. The Steering Committee consisted of WHO staff from the Coordinating Centre, the investigator(s) at each participating site, and representatives from the United Nations University and UNICEF. The Steering Committee met four times throughout implementation of the study to review the progress of the study, ensure uniformity of data collection from the different sites, and discuss substan-

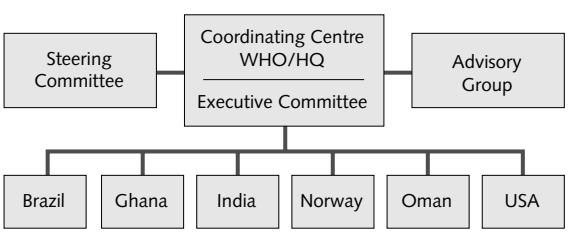


FIG. 4. Study organization

tive issues that arose. The study structure included an Executive Committee, formed by five members of the Steering Committee, which reviewed the progress and problems of the study on a regular basis and resolved substantive issues that arose from the implementation of the study. All local adaptations made to the MGRS protocols or issues related to the technical conduct of the study required review and approval by the Executive Committee. The Executive Committee also decided on the selection of study sites, the continuing participation of selected sites, and issues related to the inclusion or exclusion of data in the pooled international data set. An Advisory Group, formed by internationally recognized experts in anthropometry, epidemiology, statistics, nutrition, and human biology, provided technical advice to the Coordinating Centre, Executive Committee, and Steering Committee. Policies related to the dissemination of results and data ownership were developed prior to initiation of the study.

Field logistics

Fieldwork was undertaken by approximately 200 staff members working in different teams covering the areas of coordination, screening, lactation counseling, follow-up, and cross-sectional study. Data management teams were also present in each site. Further information on the study teams and other aspects of field logistics is presented in the papers describing the implementation of the study at specific sites [14–19]. For those interested in replicating the study elsewhere, the Manual of Operations is available on request from the first author.

Ethical issues

The study complied with the International Ethical Guidelines for Biomedical Research Involving Human Subjects [29] and received ethical approval from international, national, and local ethical review committees. Written informed consent was obtained from the parents of all children enrolled in the study.

Discussion

Growth references for infants and young children are among the most widely used instruments in public health and clinical medicine. In collaboration with a number of institutions worldwide, WHO has undertaken a major initiative to develop new growth references for infants and young children. The approach taken avoids the limitations imposed by descriptive designs that portray growth characteristics of geographically defined samples that are limited in their definition of health by relying only on the absence of overt disease at the time of measurement. Although the absence of disease remains a requirement in the WHO approach, it is no longer a sufficient criterion. The adopted strategy also requires that the reference population be defined on the basis of a number of other parameters centered on normative behaviors and other characteristics strongly associated with healthy outcomes. Furthermore, it requires that an international sample of children be used.

The MGRS is an ambitious undertaking, but the goals established on initiation of the study have been achieved successfully. The rigor with which the protocol was implemented and the data assurance procedures that were put in place have yielded a data set of outstandingly high quality. Factors that contributed to success were modern communication systems that allowed close and frequent contact between the Coordinating Centre and the sites, the continuous monitoring of data quality, the early detection and adoption of remedial measures for identified problems, and ongoing standardization within and between sites. The path to success, however, was not free of challenges.

Initial important challenges were the selection of study sites and the need to raise funding from external donors. The high cost of the study required funding from multiple donors and was largely responsible for the staggered initiation of the study in the six sites, making its management at times difficult. The high level of collaboration and uniformity that was required by a multicenter, multicultural study of this nature also presented major challenges. Close central monitoring was applied to ensure adherence to study procedures to guarantee the collection of comparable data. During the seven years of data collection, the Coordinating Centre maintained almost daily contact with the local investigators and data managers through modern communication systems and conducted frequent site visits to answer queries and assist in the data collection process. Locally, periodic coordination meetings also were conducted. There were also substantial cross-site staff exchanges to assist in lactation support, data management, data quality assurance, or motor development assessments. This created a sense of international teamwork that contributed significantly to the success of the MGRS.

The development and testing of the various growth references promises to be a complex and challenging task. This expectation is borne out by recent national experiences of a similar nature. The wealth of data being collected will allow not only the replacement of the current international references on attained growth (weight-for-age, length/height-for-age, and weight-for-length/height) but also the development of new references for triceps and subscapular skinfolds, head and arm circumferences, and body mass index. The longitudinal nature of the study will also allow the development of growth velocity curves. Health-care providers will not have to wait until children cross an

attained growth threshold to make the diagnosis of under- or overnutrition, because velocity references will enable the early identification of children in the process of becoming under- or overnourished. Similarly, the documentation of the timing of motor milestones in the longitudinal component will further enhance the value of these data by providing a unique link between physical growth and motor development. The main drawback of the new growth curves, however, is that they will cover only children up to five years of age. The need to expand this effort to older children is evident.

Ahead of us lies the implementation of the new growth references at the country level. In preparation for this phase, we recently conducted a worldwide survey of national practices in the use and interpretation of growth charts that highlighted the interest many countries have in adopting the new growth references when they become available [30]. The results from the survey also indicate that the process of replacing existing growth charts and retraining fieldworkers in the uses and interpretation of new ones must go beyond the simple change of charts, to revisiting growth monitoring practices as a whole. Intensive training efforts at all levels will be required to overcome the difficulties health workers experience with the use and interpretation of growth curves and to disseminate knowledge about effective interventions to prevent or treat either excessive or inadequate growth at both the individual and the population levels. Undoubtedly these future efforts will require a number of partnerships for their successful implementation.

Thirteen years have passed since the seed for this effort was planted. It is reasonable to ask whether the preparatory phases could have been shortened. We think that the long preparatory activities, including several Working Groups and review committees, have been decisive for the successful implementation of the study. It is unlikely that the high level of uniformity would have been achieved in such a complex multicultural project without this investment of time and effort.

The completion of weight, length, and head circumference references is anticipated before the end of 2005. The remainder of the references should be ready by 2006 (fig. 2). Of particular concern is a smooth global transition to the new references by field testing and/or use simulation of provisional references that take into account the diverse settings in which individual and population assessments occur in both developed and developing countries. This will be accomplished before the growth references are released.

The MGRS will provide a technically sound set of tools for assessing the growth and development of children worldwide for many years to come. An important characteristic of the new reference is that it makes breastfeeding the biological "norm" and establishes the breastfed infant as the normative model. Health policies and public support for breastfeeding will be strengthened when breastfed infants become the reference for normal growth and development. By prescribing the nature of the sample, the recommended approach will provide a single international reference that represents the best description possible of growth for all children less than five years of age and approximates the closest attainable "standard" of physiologic growth.

Full details about the procedures and methods, such as those contained in this supplement, are often not available in the literature. This study will be an important source of information for years to come about child growth and development and infant nutrition. It is therefore important to have a faithful record of the planning, methodology, and implementation, particularly for the benefit of those who may not have been directly involved with the MGRS but will be using the new growth charts in the near future.

References

1. World Health Organization Working Group on Infant Growth. An evaluation of infant growth. Geneva: World Health Organization, 1994.
2. World Health Organization Working Group on Infant Growth. An evaluation of infant growth: the use and interpretation of anthropometry in infants. Bull WHO 1995;73:165–74.
3. de Onis M, Yip R. The WHO growth chart: historical considerations and current scientific issues. In: Porrini M, Walter P, eds. Nutrition in pregnancy and growth. Bibl Nutr Dieta 1996;53:74–89.
4. World Health Organization. Physical status: the use and interpretation of anthropometry. Report of a WHO Expert Committee. Technical Report Series No. 854. Geneva: World Health Organization, 1995.
5. de Onis M, Habicht JP. Anthropometric reference data for international use: recommendations from a WHO Expert Committee. Am J Clin Nutr 1996;64:650–8.
6. World Health Assembly. Infant and Young Child Nutrition. Resolution WHA47.5 of the 47th World Health Assembly, 11th Plenary Meeting, Geneva, 9 May 1994.
7. de Onis M, Garza C, Habicht JP. Time for a new growth reference. Pediatrics 1997;100(5). Available at http://www.pediatrics.org/cgi/content/full/100/5/e8.
8. WHO Working Group on the Growth Reference Protocol and WHO Task Force on Methods for the Natural Regulation of Fertility. Growth patterns of breastfed infants in seven countries. Acta Paediatr 2000;89:215–22.
9. WHO Working Group on the Growth Reference Protocol and WHO Task Force on Methods for the Natural Regulation of Fertility. Growth of healthy infants and

the timing, type, and frequency of complementary foods. Am J Clin Nutr 2002;76:620–7.
10. Owusu WB, Lartey A, de Onis M, Onyango AW, Frongillo EA. Factors associated with unconstrained growth among affluent Ghanaian children. Acta Paediatr, in press.
11. Bhandari N, Bahl R, Taneja S, de Onis M, Bhan MK. Growth performance of affluent Indian children is similar to that in developed countries. Bull WHO 2002;80:189–95.
12. Mohamed AJ, Onyango AW, de Onis M, Prakash N, Mabry RM, Alasfoor DH. Socioeconomic predictors of unconstrained child growth in Muscat, Oman. East Mediterr Health J, in press.
13. Barros FC, Victora CG, Vaughan JP, Tomasi E, Horta BL, Cesar JA, Menezes MB, Halpern R, Post CL, del Mar Garcia M. The epidemiologic transition in maternal and child health in a Brazilian city, 1982-1993: a comparison of two population-based cohorts. Paediatr Perinat Epidemiol 2001;15:4–11.
14. Araújo CL, Albernaz E, Tomasi E, Victora CG, for the WHO Multicentre Growth Reference Study Group. Implementation of the WHO Multicentre Growth Reference Study in Brazil. Food Nutr Bull 2004;25(1)(suppl 1): S53–9.
15. Lartey A, Owusu WB, Sagoe-Moses I, Gomez V, Sagoe-Moses C, for the WHO Multicentre Growth Reference Study Group. Implementation of the WHO Multicentre Growth Reference Study in Ghana. Food Nutr Bull 2004;25(1)(suppl 1):S60–5.
16. Bhandari N, Taneja S, Rongsen T, Chetia J, Sharma P, Bahl R, Kashyap DK, Bhan MK, for the WHO Multicentre Growth Reference Study Group. Implementation of the WHO Multicentre Growth Reference Study in India. Food Nutr Bull 2004;25(1)(suppl 1):S66–71.
17. Baerug A, Bjoerneboe GEA, Tufte E, Norum KR, for the WHO Multicentre Growth Reference Study Group. Implementation of the WHO Multicentre Growth Reference Study in Norway. Food Nutr Bull 2004;25(1) (suppl 1):S72–7.
18. Prakash NS, Mabry RM, Mohamed AJ, Alasfoor D, for the WHO Multicentre Growth Reference Study Group. Implementation of the WHO Multicentre Growth Reference Study in Oman. Food Nutr Bull 2004;25(1) (suppl 1):S78–83.
19. Dewey KG, Cohen RJ, Nommsen-Rivers LA, Heinig MJ, for the WHO Multicentre Growth Reference Study Group. Implementation of the WHO Multicentre Growth Reference Study in the United States. Food Nutr Bull 2004;25(1)(suppl 1):S84–9.
20. de Onis M, Onyango AW, Van den Broeck J, Chumlea WC, Martorell R, for the WHO Multicentre Growth Reference Study Group. Measurement and standardization protocols for anthropometry used in the construction of a new international growth reference. Food Nutr Bull 2004;25(1)(suppl 1):S27–36.
21. Onyango AW, Pinol AJ, de Onis M, for the WHO Multicentre Growth Reference Study Group. Managing data for a multicountry longitudinal study: experience from the WHO Multicentre Growth Reference Study. Food Nutr Bull 2004;25(1)(suppl 1):S46–52.
22. Wijnhoven TMA, de Onis M, Onyango AW, Wang T, Bjoerneboe GEA, Bhandari N, Lartey A, Al Rashidi B, for the WHO Multicentre Growth Reference Study Group. Assessment of gross motor development in the WHO Multicentre Growth Reference Study. Food Nutr Bull 2004;25(1)(suppl 1):S37–45.
23. Meyer MB. Breast-feeding and smoking. Lancet 1979;1: 975–6.
24. Vio F, Salazar G, Infante C. Smoking during pregnancy and lactation and its effects on breast-milk volume. Am J Clin Nutr 1991;54:1011–6.
25. Mansbach IK, Greenbaum CW, Sulkes J. Onset and duration of breastfeeding among Israeli mothers: relationships with smoking and type of delivery. Soc Sci Med 1991;33:1391–7.
26. Kramer MS. Determinants of low birth weight: methodological assessment and meta-analysis. Bull WHO 1987;65:663–737.
27. Victora CG, Morris SS, Barros FC, Horta BL, Weiderpass E, Tomasi E. Breast-feeding and growth in Brazilian infants. Am J Clin Nutr 1998,67:452–8.
28. Cohen RJ, Brown KH, Canahuati J, Rivera LL, Dewey KG. Effects of age of introduction of complementary foods on infant breast milk intake, total energy intake, and growth: a randomised intervention study in Honduras. Lancet 1994;344:288–93.
29. Council for International Organizations of Medical Sciences/World Health Organization. International Ethical Guidelines for Biomedical Research Involving Human Subjects. Geneva: CIOMS/WHO, 1993.
30. de Onis M, Wijnhoven T, Onyango AW. Worldwide practices in child growth monitoring. J Pediatr, in press.

Measurement and standardization protocols for anthropometry used in the construction of a new international growth reference

Mercedes de Onis, Adelheid W. Onyango, Jan Van den Broeck, Wm. Cameron Chumlea, and Reynaldo Martorell, for the WHO Multicentre Growth Reference Study Group

Abstract

Thorough training, continuous standardization, and close monitoring of the adherence to measurement procedures during data collection are essential for minimizing random error and bias in multicenter studies. Rigorous anthropometry and data collection protocols were used in the WHO Multicentre Growth Reference Study to ensure high data quality. After the initial training and standardization, study teams participated in standardization sessions every two months for a continuous assessment of the precision and accuracy of their measurements. Once a year the teams were restandardized against the WHO lead anthropometrist, who observed their measurement techniques and retrained any deviating observers. Robust and precise equipment was selected and adapted for field use. The anthropometrists worked in pairs, taking measurements independently, and repeating measurements that exceeded preset maximum allowable differences. Ongoing central and local monitoring identified anthropometrists deviating from standard procedures, and immediate corrective action was taken. The procedures described in this paper are a model for research settings.

Mercedes de Onis and Adelheid W. Onyango are affiliated with the Department of Nutrition, World Health Organization, Geneva. Jan Van den Broeck is affiliated with the Africa Centre for Health and Population Studies, Mtubatuba, South Africa. Wm. Cameron Chumlea is affiliated with the Department of Community Health, Lifespan Health Research Center, Wright State University School of Medicine, Kettering, Ohio, USA. Reynaldo Martorell is affiliated with the Department of International Health, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA.

Please direct queries to: Mercedes de Onis, Study Coordinator, Department of Nutrition, World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland. Telephone: 41-22-791 3320; fax: 41-22-791 4156; e-mail: deonism@who.int.

Members of the WHO Multicentre Growth Reference Study Group and Acknowledgments are listed at the end of the first paper in this supplement (pp. S13–S14).

Key words: Anthropometry, growth curves, growth references, height, length, methods, skinfold, weight

Introduction

The World Health Organization (WHO) Multicentre Growth Reference Study (MGRS) was undertaken to generate new growth curves for assessing the growth and development of infants and young children from around the world. The children included in the study came from six countries: Brazil, Ghana, India, Norway, Oman, and the United States. The methodology and eligibility criteria for the study have been described elsewhere in this supplement [1]. Identical, rigorous data collection procedures were followed in all sites in order to minimize measurement error and to avoid systematic differences among sites.

Variability in infant and child measurements can result from a number of influences: the setting in which the measurements are taken; stomach and bladder volume (in the case of weight); diurnal variation (in length/height); the behavior and cooperation of the child being measured; the accuracy and precision of the instruments; the anthropometrist's technical capacity (training, experience, and reliability), fitness, and mood; and the methods of data recording (reading, writing down). Appropriate training and continued standardization, adherence to specified methods and procedures, and monitoring of data quality are essential to reduce measurement error and minimize bias in multisite studies. The purpose of this article is to describe the measurement protocols and routine standardization sessions that were used in the MGRS. The study protocols and quality control procedures can be applied in research settings without substantially increasing costs or complicating logistics.

Selection and training of anthropometrists

The field staff collecting anthropometric data in the

MGRS (referred to herein as observers or anthropometrists) had to have at least secondary school education, be motivated, write legibly, speak the local language, and be able to interact appropriately with the high-socioeconomic-status families that were targeted for the study. All candidates received standardized training, and only those who met the MGRS performance criteria were retained for the study.

The measurement procedures and training guidelines were prepared by the MGRS Coordinating Centre at WHO in Geneva, based on best practices recommended in anthropometry manuals and in the literature [2–8]. The initial training of anthropometrists at each site was carried out by an experienced anthropometrist following the procedures detailed in the MGRS protocol. All anthropometrists were trained to interview mothers, complete the study questionnaires, measure children as described in the protocol, avoid digit preference or transposition of numbers, record measurement values immediately after reading them, and write legibly to reduce mistakes during data transfer. Strict adherence to the measuring techniques and recording procedures was emphasized. Instructions were also provided on handling uncooperative children, taking into account cultural factors and individual mothers' sensitivity to their babies' crying.

Early in the study, four anthropometrists were trained and standardized against an expert designated by WHO as lead anthropometrist for the MGRS (W.C.C.) in a cross-site session held in Rotterdam, the Netherlands. Two of the participants were study supervisors from sites, one was a member of the Coordinating Centre, and the fourth (J.V.dB.) became the second WHO-designated lead anthropometrist for the MGRS.

Following the initial training and before the start of data collection, the anthropometrists in each site were standardized against one of the two WHO lead anthropometrists. The anthropometrist with the best performance at this session was designated "local lead anthropometrist" and was responsible for retraining teammates who deviated from MGRS techniques, and for training newly recruited anthropometrists later in the study. A WHO lead anthropometrist visited each site annually to ensure that identical methods were followed throughout the seven years of the study. The measurement procedures followed in the MGRS were documented on videotape (available to readers on request from the first author) and viewed during training and regularly thereafter to reinforce the key features of the measurement protocols.

Standardization

Given the objectives of the MGRS, standardization within and among sites was a key aspect of the study [1]. An important goal of standardized training is to enable observers to measure accurately, that is, without bias. To achieve this, observers need to be trained to obtain measurements that are on average equal to the values measured by an expert anthropometrist who is considered the "gold standard." The degree of accuracy can be assessed in a test–retest study in which several children are measured by both the expert and the observer, and bias is calculated as the average deviation of the observer's mean measurement values from those of the expert.

It is equally important that the measurements taken be precise, that is, reproducible. High precision is possible only if measurement procedures are highly standardized. Precision is assessed on the basis of differences between replicate measurements taken on several subjects in the test–retest study. The most commonly used parameter for lack of precision is the technical error of measurement (TEM) [9].

Following the initial standardization session and throughout the data collection phase, each site conducted standardization sessions bimonthly (every two months) that coincided once a year with the visit of the WHO lead anthropometrist. The purpose of these sessions was to identify anthropometrists deviating from the MGRS procedures. Corrective actions, such as retraining, were taken whenever deviations in measurement techniques were noted.

The initial standardization session used groups of 20 children for each set of measurements and took five or six days to complete, whereas the bimonthly sessions required only 10 children and could be accomplished in two or three days. At the initial session, the observers were standardized against the WHO lead anthropometrist, who served as the gold standard, whereas the bimonthly sessions used the observers' overall mean of each anthropometric variable as the gold standard. The longitudinal screening and follow-up teams were standardized separately because of the different age groups and settings involved: the screening teams measured newborns in maternity wards, whereas the follow-up teams measured infants and older children during home visits.

Analyses of accuracy and precision were performed soon after the standardization sessions using a centrally prepared Excel spreadsheet program with standard formulas for calculating relevant statistics [9–14]. To illustrate how the observers' performance was assessed, table 1 presents length data from the Rotterdam session, in which 25 children participated. For precision (TEM), the observers' performance compared well with that of the lead anthropometrist and the overall mean. This demonstrated that the participants in the session followed consistent techniques in measuring length and obtained reproducible values. The sign test for precision assesses the "measurement effect," where an observer's retest measurements may be systematically lower or higher than his or her own first measurements

TABLE 1. Precision and accuracy from the standardization session in Rotterdam: length data

Variable	WHO lead anthropometrist	Observer 1	Observer 2	Observer 3	Observer 4	Overall mean
TEM[a] (cm)	0.34	0.48	0.33	0.39	0.35	0.38
F test						
Lead anthropometrist[b]	—	$.10 < p < .25$	$p > .25$	$p > .25$	$p > .25$	
Overall mean[b]	$p > .25$	$.10 < p < .25$	$p > .25$	$p > .25$	$p > .25$	
Sign test[c]	$p > .05$	$p > .05$	$p > .05$	$p > .05$	$p > .05$	
Bias (cm)						
Lead anthropometrist[d]	—	−0.49	−0.21	−0.15	−0.15	
F test[e]	—	$p < .01$	$.01 < p < .05$	$.05 < p < .10$	$p < .01$	
Sign test[f]	—	$p < .05$	$p > .05$	$p > .05$	$p > .05$	
Bias (cm)						
Overall mean[d]	0.21	−0.33	−0.00	0.08	0.07	
F test[e]	$.10 < p < .25$	$p > .25$	$p > .25$	$p > .25$	$p > .25$	
Sign test[f]	$p > .05$	$p < .05$	$p > .05$	$p > .05$	$p > .05$	

a. Technical error of measurement: $\sqrt{(\sum d_i^2/2n)}$; where d_i is the difference between the ith subject's test and retest measurements by the observer and n is the number of measured subjects.

b. F ratio for precision: Observer $\sum d_i^2$/Lead anthropometrist $\sum d_i^2$. When overall mean is the gold standard, d_i in the denominator is the difference between the ith subject's overall mean of test and overall mean of retest measurements.

c. Precision sign test: binomial proportion p, where $p = x/n$, and x is the frequency of the observer's retest scores that are higher (or lower) than the corresponding test scores. Significance is based on exact confidence limits for proportions when $n \leq 75$ (see Table B.11 in Daly and Bourke [10]).

d. Average bias: Observer $\sum \Delta_i /n$; where Δ_i is the difference between the observer's mean and the lead anthropometrist's (or overall) mean measurement for the ith subject.

e. F ratio for bias: Observer $\sum \Delta_i^2$/lead anthropometrist's or overall means') $\sum d_i^2$ (same denominator as the precision F ratio).

f. Bias sign test: binomial proportion p, where $p = x/n$, and x is the frequency of the observer's means that are above (or below) the lead anthropometrist's or overall mean. Significance is based on exact confidence limits for proportions when $n \leq 75$ (see Table B.11 in Daly and Bourke [10]).

[15]. No such measurement effect was evident for any participant in this session. For accuracy, the observers showed a systematic tendency toward negative bias compared with the lead anthropometrist; consequently, the techniques for measuring length were reviewed. As expected, the negative bias was not evident when compared with the overall mean, except for observer 1. Both the F test and the sign test for accuracy are useful. The sign test checks whether poor accuracy results from systematic or occasional bias [10, 15]. For example, the average bias could be low and nonsignificant when a large deviation overwhelms smaller but systematic differences. In this case, the sign test, but not the F test, would indicate bias. For this session, only one observer's bias was systematic, and this was corrected by retraining.

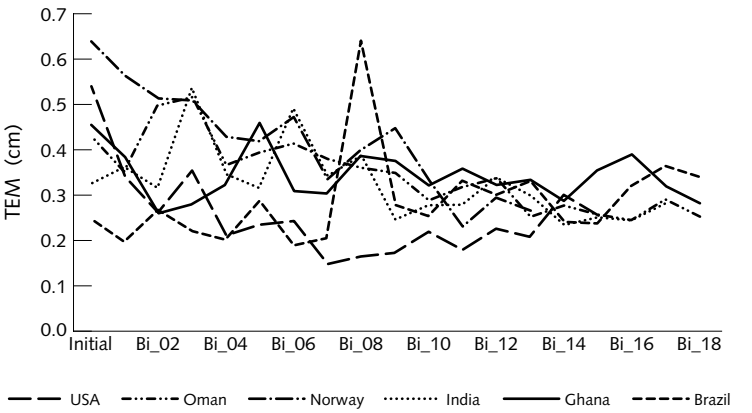
The results of the bimonthly standardization sessions were sent to the Coordinating Centre soon after their completion. The average TEMs for each site were plotted to monitor overall performance over time, as figure 1 illustrates (for length). In general, the TEMs were highest at the start, and following a pattern that is consistent for all the other measurements, precision improved as the observers gained experience. Once stabilized, the average TEMs remained low, reflecting the high precision of the measurements taken by the study teams. When sending the bimonthly results to the Coordinating Centre, sites reported on extraneous circumstances that affected the observed performance. For example, figure 1 shows a peak in TEM for the eighth bimonthly session in Brazil, when the children involved were particularly uncooperative. On the rare occasions when a problem identified in the sites needed external assistance, the WHO lead anthropometrist visited the affected site to retrain the observers. This was the case for triceps skinfold measurements at one site.

Anthropometric procedures

Measuring equipment

All study sites used the same measuring equipment. The instruments needed to be highly accurate and precise, yet sturdy and portable enough to be carried back and forth on home visits.

Length was measured with the Harpenden Infantometer (range, 30–110 cm for portable use, with digit counter readings precise to 1 mm). Because the MGRS protocol specified measuring length in the cross-sectional study for children aged 18 to 30 months (to allow a precise estimation of the systematic difference between length and height), a longer-than-usual infantometer was specially built for the study.



Initial, Bi_01 – Bi_18: Initial and bimonthly standardization sessions

FIG. 1. Technical error of measurement (TEM) for length at initial session and up to 18 bimonthly (every two months) standardization sessions in the study sites

The Harpenden Portable Stadiometer (range, 65–206 cm, digit counter reading) was used to measure both adult and child heights. At the request of WHO, the manufacturer designed a wooden base to replace the heavy carrying case that serves as a mount for the traditional portable stadiometer. This adaptation decreased the weight of the packaged stadiometer by about 7 kg and reduced the time required to assemble it.

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 mm^2; pressure between the jaws, 10 ± 2 g/mm^2; range, 0–40 mm; calibrated to 0.2 mm) was used to measure skinfolds.

To measure weight, we used portable electronic scales that have taring capability and are calibrated to 0.1 kg (UNICEF Electronic Scale 890 or Uniscale). Ideally, newborns should be measured with a scale of higher precision (within 10 g). However, the advantages of the Uniscale greatly outweighed the disadvantage of its lower precision for young babies. The scales were satisfactorily pilot tested in the Brazilian site; they were easy to use and transport, and tared weighing allowed the infants to remain in their mothers' arms where they were more calm and relaxed. This was important for the mothers' positive perception of the study and, thus, participation. The scale's electronic display decreased the observer measurement error. In cold climates, the infants could be wrapped up in a blanket for weighing after the weight of the blanket had been tared. Another advantage of the Uniscale was that it allowed the mother's weight to be recorded at each visit, thus permitting the collection, at no extra cost, of weight data for lactating women.

The equipment was calibrated regularly, usually daily before the home or hospital visits. The scales were calibrated with locally available standard weights over the full weight range, and tared weighing was simulated. The infantometer and stadiometer were calibrated by using metal rods of known lengths. 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.

Anthropometric data collection

Measurements were taken and recorded by two trained and standardized anthropometrists. Both the questionnaire forms and the standard procedures were designed to ensure that each observer read and recorded measurements independently of the other. At each session, the two exchanged roles as "leading" and "assisting" observers. The role of the assisting observer was to help position the child correctly while the leading observer took and recorded measurements. The first observer measured and immediately recorded each of the measurements, and they then exchanged roles so that the second observer could also take the full set of measurements. They then compared their values to ensure that duplicate measurements were within the maximum allowed differences. Any measurements falling outside the maximum allowed differences were repeated by both observers and entered in designated boxes on the data recording sheet. No more than two remeasurements were allowed (i.e., a maximum of three duplicate measurement sets for a given anthropometric parameter at any one visit). All recorded measurements were entered into the computer. The final value to be

used for the construction of the growth curves will be the average of the last pair of measurements. In the rare cases (< 0.1%) when the child was judged to be too agitated for reliable duplicate measurements to be taken, only one set of measurements was recorded. In practice, it was observed that large differences owing to reading or recording errors were resolved by a first repeat measurement. However, when the babies were uncooperative, measurements became increasingly difficult, and hence the decision to discontinue measuring and use unpaired measurements in the few cited cases.

The maximum allowable differences for acceptable precision used in the study for the various anthropometric variables were based on the TEM obtained in the initial standardization session conducted at the Brazilian site. To achieve a rate of remeasurement of around 5%, the maximum allowed differences were set at 2.8 times the TEM achieved during the session, i.e., 7 mm for length, 5 mm for circumferences, and 1.2 mm for skinfolds (table 2). The maximum allowable difference for weight was set at 100 g to allow for rounding off within the smallest calibration unit of the scale. Because skinfold thicknesses were the measurements with which mothers and children were least familiar and felt most uncomfortable, the decision was taken to raise the maximum allowable difference for skinfolds to 2 mm. This was considered to be a more appropriate limit, as a narrower margin might lead to too many repeat measurements, with negative implications for the anthropometrists' morale and the participants' responsiveness.

Measurement schedule

The MGRS anthropometric measurements are weight, recumbent length, standing height, head and arm circumferences, and triceps and subscapular skinfold thickness. For the longitudinal study, newborns were measured at birth (usually within the first 12 hours of life, and never after 24 hours) and visited at home 21 times: at weeks 1, 2, 4, and 6; monthly from 2 to 12 months; and every other month during the second year (table 3). Data collection was more frequent at younger ages so that these early phases of rapid growth could be adequately described. The week 1 visit was done by the lactation counselor, and only weight was measured, following the standard MGRS procedure (using the Uniscale and weighing the baby twice). The mother's weight was recorded at each visit, and the father's weight and both parents' heights were measured once.

In the cross-sectional study, children aged 18 to 71 months were measured once, except in the two sites that used a mixed-longitudinal design [1], in which some children were measured two or three times, at 3-month intervals. All children aged 18 to 30 months had both recumbent length and standing height measured, and parental weights and heights were measured once.

Concerted efforts were made to collect the anthropometric data on scheduled visit dates. Theoretically, the maximum delay or advance of measurements allowed by the protocol was 10% of the child's age (e.g., 3 days at 1 month, 18 days at 6 months), but in practice, teams worked with more restricted tolerable delay or advance targets (0, 1, 2, 4, and 5 days for visits at weeks 1, 2, 4, and 6 and at 2 months, respectively; 7 days for visits taking place at 3 months onwards). Of more than 32,000 home visits completed by April 2003, only 217 (0.7%) were done outside the maximum allowable delay, out of these, 58 (26.7%) exceeded the limit by less than one day.

Measurement techniques

A comprehensive description of the techniques used for the measurements is found in the MGRS Measurement and Standardization Protocols and documented in the anthropometric training video (available on request from the first author). The anthropometrists explained to the mothers all procedures to be undertaken and emphasized that these were harmless. Infants and young children were held by their mothers to foster a sense of

TABLE 2. Maximum allowable differences between the measurements of two observers

Measurement	Brazil TEM[a] from pilot study	Maximum allowable difference
Weight	Not available	100 g
Length	2.5	7.0 mm
Head circumference	1.4	5.0 mm
Arm circumference	1.8	5.0 mm
Triceps skinfold	0.44	2.0 mm
Subscapular skinfold	0.43	2.0 mm

a. TEM, Technical error of measurement (see formula in footnote to table 1).

TABLE 3. Time schedule for the collection of anthropometric measurements in the longitudinal study

Measurement and time frame	Frequency	No. of visits
Weight,[a] length, head circumference		
Birth	Once	1
2–6 wk	Every 2 wk	4
2–12 mo	Monthly	10
14–24 mo	Every 2 mo	6
Arm circumference, skinfold thickness (triceps, subscapular)		
3–12 mo	Monthly	10
14–24 mo	Every 2 mo	6

a. Weight was also measured at week 1 by the lactation counselor.

security for the baby. The anthropometrist's confidence and poise was important for reassuring both mother and child, and included maintaining eye contact and talking to the child in a calm, reassuring voice.

Arm circumference and skinfold measurements were taken on the left side of the body. The choice of which side to measure (right or left) matters little to accuracy and precision [6]; however, the left-hand side is used more often. Length, height, circumferences, and skinfolds were recorded to the last completed unit rather than the nearest unit. To correct for the systematic negative bias introduced by this practice, half of the smallest measurement unit (i.e., half of 0.2 mm for skinfolds and half of 0.1 mm for circumferences) was added to each measurement before analysis. This correction did not apply to weight, which was rounded off to the nearest 100 g.

For measurement of weight, the mother removed all the child's clothes, but as noted earlier, use of a blanket to cover the baby was encouraged in cold weather. The parents took off their shoes, heavy clothing, and other heavy objects before being weighed. They wore light clothing of known weight that was recorded and later subtracted from the subject's weight. This was done by using a list of weights of local clothes. In the longitudinal study, the mother was weighed first, and after her weight was recorded, the scale was tared and the baby was given to her. She was asked to stand still until the baby's weight had been displayed and recorded. When children could not be undressed, they also wore standard light clothing of known weight that was recorded and subtracted to obtain the child's weight. Children aged two years or more in the cross-sectional study were weighed on their own, standing with their feet slightly apart in the center of the platform of the scale.

To measure recumbent length, braids were undone and hair ornaments were removed if they interfered with positioning of the head. Diapers were also removed, because they made it difficult to hold the infant's legs together and straighten them. The leading observer stood on one side of the board to hold down the baby's legs with one hand and move the foot board with the other. The assisting observer stood at the headboard to help position the child's head. The head was positioned so that the crown touched the headboard and a vertical line from the ear canal to the lower border of the eye socket was perpendicular to the horizontal board (i.e., the Frankfort plane positioned vertically). The leading observer positioned the child's shoulders and hips at right angles to the long axis of the body. Gentle pressure was applied to the knees to straighten the legs. To avoid causing injury, minimal but prolonged pressure was applied to the knees of newborns. To take the measurement, the footboard was positioned against the child's feet with the soles flat on the board and the toes pointing upwards. The measurement was recorded to the last completed 1 mm.

To measure standing height, hair ornaments were removed and braids were undone. The child stood on the stadiometer with bare feet placed slightly apart and the back of the head, shoulder blades, buttocks, calves, and heels touching the vertical board. The assisting observer held the child's knees and ankles to keep the legs straight and the feet flat. The leading observer got down to a face-to-face level with the child and positioned the child's head so that a horizontal line drawn from the ear canal to the lower edge of the eye socket ran parallel to the baseboard (i.e., the Frankfort plane positioned horizontally). Because young children have difficulty standing to full stature, a gentle push applied to the tummy was used to help them stand to full height. The headboard was pulled down to rest firmly on top of the head and compress the hair, and the reading was taken to the last completed 1 mm.

To measure head circumference, hairpins or headbands were removed and braids were undone. An infant or child below the age of two years was held on the mother's lap, and older children could stand or sit unassisted. The leading observer stood or sat at the left side of the child, passed the tape around the head, and anchored it just above the eyebrows and over the fullest protuberance of the skull at the back of the head. The assisting observer stood or sat in front of the child and helped by positioning the tape correctly on the side away from the lead observer. Once positioned correctly, the tape was pulled tight to compress the hair and skin, and the reading was recorded to the last completed 1 mm.

The mid-upper-arm point is half the distance between the acromion process (the most lateral bony protuberance of the back of the shoulder) and the olecranon (the bony structure that stands out when the elbow is bent). The midpoint was located and marked for measurement of the mid-upper-arm circumference (MUAC) and triceps skinfold thickness. One observer palpated the shoulder to find the acromion and marked it with a felt-tip pen or cosmetic pencil. The child's forearm was then bent 90° at the elbow, palm facing up, so that the olecranon stood out at the elbow. The observer placed the zero point of the tape on the mark over the acromion process and ran it downward along the back of the arm to the tip of the elbow. The other observer made a small horizontal mark at the midpoint on the posterior aspect of the arm before the tape was removed.

For measurement of the MUAC, the child's arm hung in a relaxed position or was held in the extended position by the assisting observer; care was taken not to flex or tighten the muscles. The tape was then wrapped around the arm over the marked midpoint. The tape had to lie flat around the arm, without compressing the skin or underlying tissue; the assisting observer checked to ensure that there was no gap or compression on the inner part of the arm before the measurement was

recorded to the last completed 1 mm.

A skinfold consists of a double fold of skin and subcutaneous fat, excluding the underlying muscle. The teams were trained to grasp the skinfold gently to avoid causing unnecessary discomfort to the child and compressing the fat. Skinfolds were recorded to the last completed 0.2 mm. For measurement of triceps skinfold thickness, young babies were held by their mothers; older children sat or stood on their own. The left arm hung relaxed at the side or was held down by the mother or assisting observer. The leading observer stood behind the child and picked up the skinfold about 1 cm above the midpoint mark over the triceps muscle, with the fold running downward along the midline of the back upper arm. The caliper jaws were applied at right angles to the "neck" of the fold just below the finger and thumb over the midpoint mark. While maintaining a grip on the skinfold, the observer gently released the caliper handles and allowed the jaws to close on the fat fold for two seconds before taking the reading to the last completed 0.2 mm.

The measurement point for the subscapular skinfold located immediately below the inferior angle of the scapula was identified by palpating and marking the inferior angle of the scapula. The child stood or sat with shoulders relaxed or gently held down to prevent movement of the scapula. The skinfold was picked up 1 cm above and medial to the subscapular mark, and the caliper was applied to the "neck" of the fold over the mark so that the fold ran diagonally down toward the left elbow. The same procedure as described for the triceps skinfold was followed to read and record the measurement.

Quality control during data collection

The observers' performance was monitored in several ways during the study:

The requirement to take and record all measurements independently by the two observers and to compare their measurement values for maximum allowable differences was a key strategy for detecting errors and remeasuring the child on the spot.

The proportion of repeated measurements at each site was closely monitored as an important quality control measure. Low levels of remeasurement signal a possible lack of independence between the observers, whereas high levels might indicate poor measurement techniques on the part of at least one of the observers. The levels of maximum allowable differences selected anticipated repeat rates of about 5%. The observed rates according to site are reported for newborns (fig. 2), young children in the longitudinal study (fig. 3), and older children in the cross-sectional study (fig. 4). For the overall study, the rate of repeated length measurements in the cross-sectional sample was 5%, as expected, but it was double this percentage in the longitudinal study (11%)(table 4). The lowest proportions of repeat measurements were observed for the skinfolds (3% for triceps and 1% for subscapular

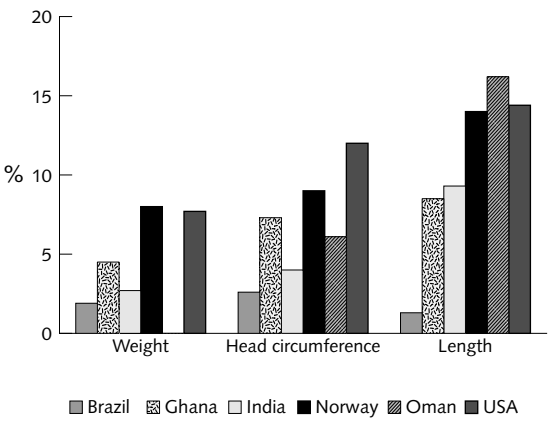


FIG. 2. Percentage of newborn measurements repeated for exceeding the maximum allowable difference between observers

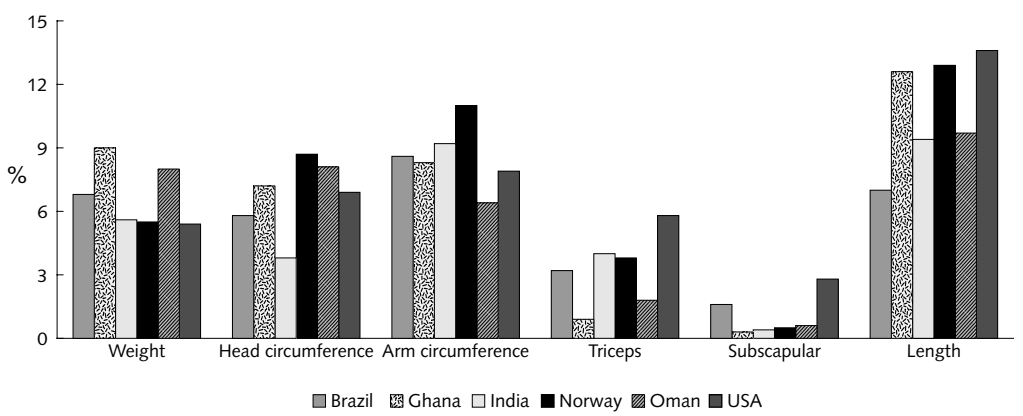


FIG. 3. Percentage of longitudinal follow-up measurements repeated for exceeding the maximum allowable difference between observers

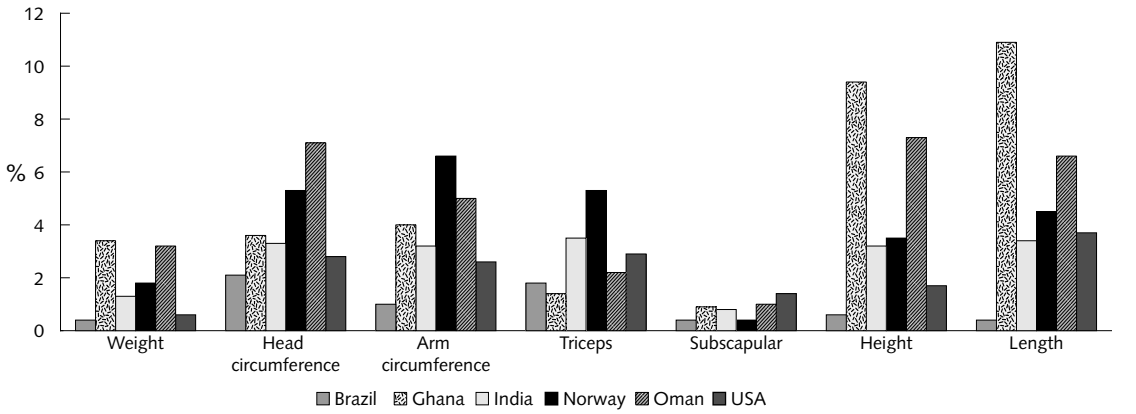


FIG. 4. Percentage of cross-sectional study measurements repeated for exceeding the maximum allowable difference between observers

TABLE 4. Summary of measurements repeated for exceeding the maximum allowable difference between observers

Measurement	No. (%) of measurements		
	Newborns (n = 1,746)	Longitudinal study (n = 31,248)	Cross-sectional study (n = 8,254)
Weight	69 (4.0)	2,194 (6.8)	148 (1.8)
Length	180 (10.8)	3,450 (10.7)	81 (4.9)[a]
Head circumference	114 (6.5)	2,173 (6.8)	335 (4.1)
Arm circumference	N/A[b]	2,761 (8.6)	309 (3.7)
Triceps skinfold	N/A	982 (3.1)	236 (2.9)
Subscapular skinfold	N/A	300 (0.9)	69 (0.8)
Height	N/A	N/A	354 (4.3)

a. n = 1,653: only children aged 18–30 months in the cross-sectional study were measured for length.
b. N/A, Indices not measured.

skinfolds in both longitudinal and cross-sectional components), probably as a result of the adoption of a wider margin of allowable differences than the initial standard set in Brazil (table 4).

The completed questionnaires were delivered soon after the home visit, usually within one or two days, to the local coordination center, where they were checked by the supervisor for completeness and consistency using procedures described elsewhere in this supplement [16]. For anthropometry, the data entry system included built-in range and consistency checks that flagged measurements exceeding ±2 standard deviations of age- and sex-specific reference values for attained size. Flagged values were then checked for consistency between the two observers, consistency with other anthropometric variables measured on the same visit, consistency with previous measurements of the same child, and possible data entry errors.

Periodic computer checks were also done for each observer to detect digit preferences and unusual values. For example, because the skinfold caliper reads to 0.2-mm units, there should be no odd decimal values (e.g., 0.1 mm, 0.3 mm) recorded for skinfolds. Table 5 is a sample digit preference table for triceps skinfold measurements taken by one site team. The output from this analysis was examined for terminal digit preference and avoidance. According to table 5, observer 1 tended to avoid digit 6, but there was no pattern to suggest bias in observed proportions of the other digits. On the other hand, for observer 2, the proportions of digit 0 (8.4%) versus 2 (34.4%) suggested a tendency to overestimate measurements. When the imbalance between two consecutive digits was particularly large, the differences between measurement pairs were analyzed to determine whether the affected observer was biased in relation to others that had been paired with him or her.

Quality control checks were performed by randomly calling approximately 10% of the mothers to repeat a selection of the questions on the study forms and to ask the mother whether the child had been measured twice by the interviewers. These calls also provided the opportunity to monitor participant responsiveness and satisfaction with the study.

Bimonthly (every two months) standardization sessions served to ensure that the observers were not departing from the measuring techniques of the study,

TABLE 5. Sample table of terminal digit preference analysis in longitudinal follow-up study (triceps skinfold data from Oman)

Observer	No. of measurements	Terminal digit % (95% confidence interval)					Probability of equal proportions
		0	2	4	6	8	
1	773	22.6 (19.7, 25.6)	21.6 (18.7, 24.5)	20.6 (17.7, 23.4)	14.2 (11.8, 16.7)	21.0 (18.1, 23.8)	0.002
2	1,051	8.4 (6.7, 10.0)	34.4 (31.6, 37.3)	20.9 (18.5, 23.4)	16.8 (14.6, 19.1)	19.4 (17.0, 21.8)	< 0.0001
3	866	19.7 (17.1, 22.4)	25.1 (22.2, 27.9)	22.4 (19.6, 25.2)	6.6 (4.9, 8.2)	26.2 (23.3, 29.1)	< 0.0001
4	996	23.0 (20.4, 25.6)	20.2 (17.7, 22.7)	23.7 (21.1, 26.3)	15.5 (13.2, 17.7)	17.7 (15.3, 20.0)	< 0.0001
5	839	16.5 (13.9, 19.0)	20.5 (17.8, 23.2)	22.2 (19.4, 25.0)	19.2 (16.5, 21.9)	21.7 (18.9, 24.5)	0.065
6	702	16.1 (13.4, 18.8)	20.2 (17.3, 23.2)	18.0 (15.1, 20.8)	21.9 (18.9, 25.0)	23.8 (20.6, 26.9)	0.01
7	1,123	13.7 (11.7, 15.7)	23.2 (20.7, 25.6)	26.0 (23.4, 28.6)	14.0 (12.0, 16.0)	23.2 (20.7, 25.6)	< 0.0001
8	785	29.9 (26.7, 33.1)	18.6 (15.9, 21.3)	18.5 (15.8, 21.2)	15.8 (13.2, 18.4)	17.2 (14.6, 19.8)	< 0.0001
9	657	19.8 (16.7, 22.8)	20.1 (17.0, 23.2)	21.2 (18.0, 24.3)	16.4 (13.6, 19.3)	22.5 (19.3, 25.7)	0.1514

to monitor precision and accuracy, and to take corrective measures (e.g., retraining) when required.

To maintain a good rapport with the families, each participant in the longitudinal study had one "fixed" fieldworker for the duration of the follow-up. The other fieldworkers were rotated every two months in order to distribute error terms, avoid boredom, and prevent complicity that might undermine the measurement protocol.

Discussion

The rigorous anthropometric protocols described in this paper were set in place to ensure high data quality. These MGRS procedures serve as a model for research settings. The methods and procedures reviewed will be applicable to multi- and single-site studies. It will not be possible to be as rigorous in nonresearch settings, such as child clinics. At the very least, the procedures should be carefully documented in training manuals, staff members collecting anthropometric data should be trained and refresher sessions should be held periodically, weighing scales and any other instruments used should be maintained in good order and calibrated before use, and fieldworkers should be supervised.

The standardization sessions were effective in identifying factors that contribute to low accuracy and precision in anthropometric measurements. Training and retraining opportunities were available to help keep the anthropometrists' skills sharp, as were printed and videotaped reference materials. These were particularly useful when reserve staff were preparing to take part in data collection and when new team members were recruited in the course of the study. In general, new staff began taking anthropometric measurements for the MGRS only after being standardized against the WHO lead anthropometrist.

Factors that affected measurement accuracy and precision included the identification of landmark features when measuring soft tissues (arm circumference and skinfolds). In some sites, the teams experienced difficulties in taking measurements because they did not mark the upper-arm midpoint or the subscapular point. In this respect, the Coordinating Centre's ongoing monitoring of anthropometric data and the regular participation of the WHO lead anthropometrists in site standardization sessions were extremely important for detecting and correcting problems.

For research and programmatic activities, it is relevant to note that that the child's age could affect the precision of some measurements, judging by the differences in repeat rates for arm circumference (9% versus 4%) and head circumference (7% versus 4%) in the longitudinal and cross-sectional studies, respectively. Users who adopt the same limits of maximum allowable differences between independently recorded duplicate measurements could evaluate performance with the MGRS-observed proportions as a reference. Thus, for children below the age of two years, about 11% of length measurement pairs will differ by more than 7 mm. In the same age group, 7% of duplicate head circumferences will differ by more than 5 mm, as will 9% of duplicate arm circumference measurements. Overall, the rates of repeated measurement are expected to be lower in older children, who tend to be

calmer and more cooperative. A team that exceeds these proportions may be in need of further training, and a team that has substantially lower rates may be taking nonindependent measurements.

References

1. de Onis M, Garza C, Victora CG, Onyango AW, Frongillo EA, Martines J, for the WHO Multicentre Growth Reference Study Group. The WHO Multicentre Growth Reference Study: planning, study design, and methodology. Food Nutr Bull 2004;25(1)(suppl 1):S15–26.
2. Lohman TG, Roche AF, Martorell R, eds. Anthropometric standardization reference manual. Champaign, Ill, USA: Human Kinetics Books, 1988.
3. Cameron N. Anthropometric measurements. In: Cameron N, ed. The measurement of human growth. London: Croom Helm, 1984:56–99.
4. Gordon CC, Cameron Chumlea W, Roche AF. Stature, recumbent length, and weight. In: Lohman TG, Roche AF, Martorell R, eds. Anthropometric standardization reference manual. Champaign, Ill, USA: Human Kinetics Books, 1991:3–8.
5. Habicht J-P, Yarbrough C, Martorell R. Anthropometric field methods: criteria for selection. In: Alfin-Slater RB, Kritchevsky D, series eds; Jelliffe DB, Jelliffe EFP, volume eds. Human growth—a comprehensive treatise. Vol 2. Nutrition and growth. New York: Plenum Publishing Company, 1979:365–87.
6. Martorell R, Mendoza F, Mueller WH, Pawson IG. Which side to measure: right or left? In: Lohman TG, Roche AF, Martorell R, eds. Anthropometric standardization reference manual. Champaign, Ill, USA: Human Kinetics Books, 1988:87–91.
7. Tanner JM, Hiernaux J, Jarman S. Growth and physique studies. In: Weiner JS, Lourie JA, eds. Human biology. A guide to field methods. IBP Handbook No. 9. Oxford, UK: Blackwell Scientific Publications, 1969:1–42.
8. World Health Organization. Physical status: the use and interpretation of anthropometry. Report of a WHO Expert Committee. Technical Report Series No. 854. Geneva: World Health Organization, 1995.
9. Malina RM, Hamill PVV, Lemeshow S. Selected measurements of children 6–11 years. Vital Health and Statistics Series 11, No. 123, USDHHS. Washington, DC: US Government Printing Office, 1973.
10. Daly LE, Bourke GJ. Interpretation and uses of medical statistics, 5th ed. Oxford, UK: Blackwell Science, 2000.
11. Johnson TS, Engstrom JL, Gelhar DK. Intra- and interexaminer reliability of anthropometric measurements of term infants. J Pediatr Gastroenterol Nutr 1997;24:497–505.
12. Marks GC, Habicht JP, Mueller WH. Reliability, dependability, and precision of anthropometric measurements. Am J Epidemiol 1989;130:578–87.
13. Martorell R, Habicht JP, Yarbrough C, Guzman G, Klein RE. The identification and evaluation of measurement variability in the anthropometry of preschool children. Am J Phys Anthropol 1975;43:347–52.
14. Villar J, Kestler E, Pareja G. Measurement error in clinical perinatal data. Am J Obstet Gynecol 1989;160:380–2.
15. World Health Organization. Measuring change in nutritional status. Annex 1: Standardization procedures for the collection of weight and height data in the field. Geneva: World Health Organization, 1983.
16. Onyango AW, Pinol AJ, de Onis M, for the WHO Multicentre Growth Reference Study Group. Managing data for a multicountry longitudinal study: experience from the WHO Multicentre Growth Reference Study. Food Nutr Bull 2004;25(1)(suppl 1):S46–52.

Assessment of gross motor development in the WHO Multicentre Growth Reference Study

Trudy M.A. Wijnhoven, Mercedes de Onis, Adelheid W. Onyango, Tracey Wang, Gunn-Elin A. Bjoerneboe, Nita Bhandari, Anna Lartey, and Badriya Al Rashidi, for the WHO Multicentre Growth Reference Study Group

Abstract

The objective of the Motor Development Study was to describe the acquisition of selected gross motor milestones among affluent children growing up in different cultural settings. This study was conducted in Ghana, India, Norway, Oman, and the United States as part of the longitudinal component of the World Health Organization (WHO) Multicentre Growth Reference Study (MGRS). Infants were followed from the age of four months until they could walk independently. Six milestones that are fundamental to acquiring self-sufficient erect locomotion and are simple to evaluate were assessed: sitting without support, hands-and-knees crawling, standing with assistance, walking with assistance, standing alone, and walking alone. The information was collected by both the children's caregivers and trained MGRS fieldworkers. The caregivers assessed and recorded the dates when the milestones were achieved for the first time according to established criteria. Using standardized procedures, the fieldworkers independently assessed the motor performance of the children and checked parental recording at home visits. To ensure standardized data collection, the sites conducted regular standardization sessions. Data collection and data quality control took place simultaneously. Data verification and cleaning were performed until all queries had been satisfactorily resolved.

Key words: Child, child development, infant, longitudinal study, motor development, motor skills

Trudy M.A. Wijnhoven, Mercedes de Onis, and Adelheid W. Onyango are affiliated with the Department of Nutrition, World Health Organization, Geneva, Switzerland. Tracey Wang is affiliated with the Department of Nutrition, University of California, Davis, California, USA. Gunn-Elin A. Bjoerneboe is affiliated with the Norwegian Directorate for Health and Social Affairs, Oslo, Norway. Nita Bhandari is affiliated with the Department of Pediatrics, All India Institute of Medical Sciences, New Delhi, India. Anna Lartey is affiliated with the Department of Nutrition and Food Science, University of Ghana, Legon, Accra, Ghana. Badriya Al Rashidi is affiliated with the Directorate General of Health Services, Muscat Region, Ministry of Health, Muscat, Oman.

Please direct queries to: Mercedes de Onis, Study Coordinator, Department of Nutrition, World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland. Telephone: 41-22-791 3320; fax: 41-22-791 4156; e-mail: deonism@who.int.

Members of the WHO Multicentre Growth Reference Study Group and Acknowledgments are listed at the end of the first paper in this supplement (pp. S13–S14).

Introduction

Motor behavior is an essential aspect of child development. Given the unique opportunity provided by the World Health Organization (WHO) Multicentre Growth Reference Study (MGRS), a component to assess gross motor development was included in the protocol. Motor development is usually assessed in terms of age of achievement of motor milestones [1]. Besides the determination of age at attainment, longitudinal assessment of different types of motor skills has the advantage of providing a profile of their sequence and tempo [2, 3]. However, few studies using a longitudinal design have been done on the age of achievement of certain motor milestones [4–21], and only two of them are of a multicountry nature [10, 21].

To our knowledge, this is the first longitudinal study that has used a standardized protocol to describe gross motor development among groups of children with no health, environmental, or economic constraints on growth, living in different countries. The study sample described by the WHO Task Force for Epidemiological Research on Reproductive Health [21] was stratified into three different socioeconomic-level groups, and the study of Hindley et al. [10] was limited to the assessment of one motor milestone. The MGRS aimed to fill this gap in information by collecting data on six gross motor milestones in five of the countries participating in its longitudinal growth and development study: Ghana, India, Norway, Oman, and the United States. Under the umbrella of the MGRS, the Motor Develop-

ment Study provides a unique opportunity to assess group and individual variability in the acquisition of key motor skills, as well as providing an opportunity to analyze the relationship between physical growth and gross motor development among groups of affluent children growing up in different cultural settings.

This paper outlines the Motor Development Study protocol for collecting information on six motor milestones, the methods and procedures of data collection, and the training and standardization of fieldworkers.

Methods

Study subjects and study design

The motor development assessments were done from the age of four months on all subjects enrolled in the longitudinal component of the MGRS. Details of the enrollment of subjects, the inclusion criteria, and the MGRS study design are explained elsewhere in this supplement [22]. The study took place in five of the six countries participating in the MGRS: Ghana, India, Norway, Oman, and the United States. The implementation of the study protocols in each of these countries is described in separate papers in this supplement [23–27]. The Brazilian site was unable to participate in the Motor Development Study because the site had initiated data collection by the time the decision to assess motor development was taken.

Gross motor milestones: description, criteria, and testing procedure

Six distinct gross motor milestones were selected for study: sitting without support, hands-and-knees crawling, standing with assistance, walking with assistance, standing alone, and walking alone. These milestones were selected because they are considered to be universal, fundamental to the acquisition of self-sufficient erect locomotion, and simple to test and evaluate.

Before the achievement of any of the six motor skills, the child goes through many preceding intermediate stages of development [28, 29]. Evaluation of a milestone performance consists in observing not only what a child does, but also how and with what level of development he or she does it [29]. There is also a need to include in the criteria for testing whether a child can perform a milestone independently or performs it after having been placed into position [30]. Thus, in order to minimize interpersonal interpretation differences, each test item needed to be clearly defined with respect to the method of administration and the interpretation of the child's performance [1].

The descriptions of the six gross motor skills used in this study originated from various existing developmental scales [2, 29, 31–35]. The sequential presentation of the motor milestones followed the pattern generally found in the literature [36–40]. However, occasionally the suggested sequence between two or more milestones might actually be reversed, and observed milestones might be inhibited later [28]. Therefore, no fixed developmental sequence of achievement was assumed.

All milestones were assessed using standardized

TABLE 1. MGRS performance criteria for six gross motor milestones

Gross motor milestone	MGRS performance criteria
Sitting without support	Child sits up straight with the head erect for at least 10 seconds. Child does not use arms or hands to balance body or support position
Hands-and-knees crawling	Child alternately moves forward or backward on hands and knees. The stomach does not touch the supporting surface. There are continuous and consecutive movements, at least three in a row
Standing with assistance	Child stands in upright position on both feet, holding onto a stable object (e.g., furniture) with both hands without leaning on it. The body does not touch the stable object, and the legs support most of the body weight. Child thus stands with assistance for at least 10 seconds
Walking with assistance	Child is in upright position with the back straight. Child makes sideways or forward steps by holding onto a stable object (e.g., furniture) with one or both hands. One leg moves forward while the other supports part of the body weight. Child takes at least five steps in this manner
Standing alone	Child stands in upright position on both feet (not on the toes) with the back straight. The legs support 100% of the child's weight. There is no contact with a person or object. Child stands alone for at least 10 seconds
Walking alone	Child takes at least five steps independently in upright position with the back straight. One leg moves forward while the other supports most of the body weight. There is no contact with a person or object

FIG. 1. Sitting without support

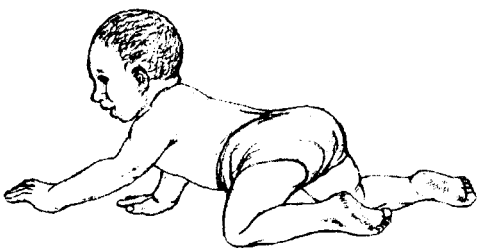


FIG. 2. Hands-and-knees crawling

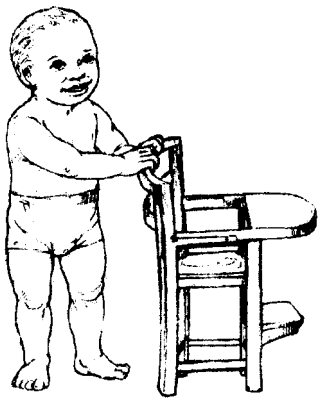


FIG. 3. Standing with assistance

testing procedures. Table 1 describes the MGRS performance criteria for the six milestones. A milestone was considered achieved only if all the given criteria were met.

Sitting without support (fig. 1)

Description. The child is able to balance the weight of the trunk and head without any external support or the use of arms and hands. The child sits up straight with the head erect (that is, not leaning forward). One of the lower limbs is usually flexed.

Criteria. (a) The child's head is erect; (b) the child does not use the arms or hands to balance body or support position; (c) the child sits up straight for at least 10 seconds.

Testing procedure. Facing the child and smiling, the fieldworker places the child in a sitting position. The fieldworker then gives the child a toy to handle with both hands so that he or she is not able to use the arms to support himself or herself.

Hands-and-knees crawling (fig. 2)

Description. This is a phase of a more organized prone movement that refers to the palm-knee position, with alternating movements of the upper and lower limbs: the right arm and left leg move forward or backward synchronously and vice versa in similarly ordered consecutive movements.

Criteria. (a) Alternating movement forward or backward on hands and knees; (b) the child's stomach does not touch the supporting surface; (c) continuous and consecutive movements, at least three in a row.

Testing procedure. The fieldworker places the child in the prone position with the abdomen above the supporting surface. The fieldworker places himself or herself in front of the child, about 120 to 150 cm away. If the child does not crawl spontaneously, the fieldworker shows the child a toy or object that attracts the child's visual attention. The fieldworker (sometimes with the help of the caregiver) then tries to coax the child to crawl toward the toy and grab it.

Standing with assistance (fig. 3)

Description. This is the first direct step toward erect bipedal locomotion, in which the child is for the first time challenged to maintain some balance of the whole body weight so that he or she can move forward. The salient characteristic is whether the child can actually support his or her weight if he or she is holding onto a stable object (e.g., a piece of furniture) with both hands without leaning over or resting the body on the stable object.

Criteria. (a) The child is in an upright position on both feet; (b) the child holds onto a stable object with both hands without leaning on it; (c) the child's body does not touch the stable object; (d) the child's legs support most of the child's body weight; (e) the child thus stands with assistance for at least 10 seconds.

Testing procedure. The fieldworker places the child in a standing position so that the legs support the body weight. The child is placed at a distance from which both hands, but not the body, can reach and hold onto a stable object. Thus, most of the body weight is supported by the child's own feet. The fieldworker should check that the child is not leaning over or resting his or her body on the stable object. The height of the stable object should be at about the same level as the child's stomach.

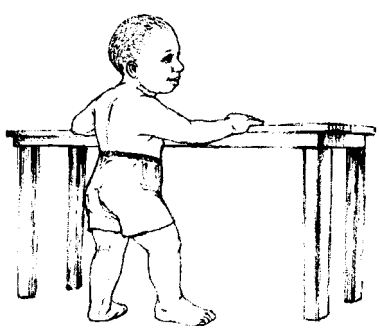


FIG. 4. Walking with assistance

Walking with assistance (fig. 4)

Description. This involves a deliberate attempt to make stepping movements and to make postural adjustments toward this end while holding onto a stable object (e.g., furniture) for support.

Criteria. (a) The child is in an upright position with the back straight; (b) the child makes sideways or forward steps by holding onto a stable object with one or both hands; (c) one leg moves forward while the other supports part of the body weight; (d) the child takes at least five steps in this manner.

Testing procedure. The fieldworker places the child in a standing position so that the legs support most of the body weight. The child is placed at a distance from which he or she can reach and hold onto a stable object with one or both hands. If the child does not move spontaneously, the fieldworker shows the child a toy or object that attracts the child's visual attention. The fieldworker (sometimes with the help of the caregiver) then tries to coax the child to walk toward the toy and grab it. The height of the stable object should be at about the same level as the child's stomach.

Standing alone (fig. 5)

Description. The child shows the capacity for both equilibration and sustaining body weight on the feet. In this position the child's legs show no flexion, and the child is standing on the feet (not on the toes) without leaning over or holding onto an object. The child maintains continuous balance independently.

Criteria. (a) The child is in an upright position on both feet (not on the toes) with the back straight; (b) the child's legs support 100% of the child's weight; (c) there is no contact with a person or an object; (d) the child stands alone for at least 10 seconds.

Testing procedure. The fieldworker places the child with both feet flat on the floor and supports the child to an erect position. Then the fieldworker withdraws the support gradually and temporarily to determine whether the child can modify posture, adjust to the new position, and stand alone for at least 10 seconds.

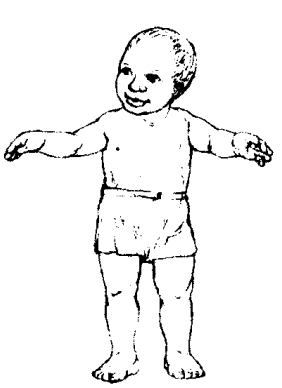


FIG. 5. Standing alone

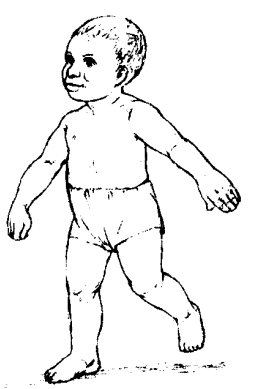


FIG. 6. Walking alone

Walking alone (fig. 6)

Description. The child shows the capacity to balance the body and to control his or her forward stepping movements. There is no need for assistance, because both the postural adjustment and the stepping movements are engaged in independent walking. An important indicator of this phase of erect locomotion is that movement of the entire body does not accompany the child's stepping movements. This phase does not refer to the child's first independent steps when the child is able to take three or four uncertain steps toward the adult's outstretched hands.

Criteria. (a) The child is in an upright position with the back straight; (b) one leg moves forward while the other supports most of the body weight; (c) there is no contact with a person or an object; (d) the child takes at least five steps independently.

Testing procedure. The fieldworker places the child in an erect position out of the reach of any supporting object. Then the fieldworker takes a position about 120 to 150 cm in front of the child and calls the child to move toward the fieldworker. Sometimes, the caregiver needs to encourage the child.

The child's emotional state

Because emotional arousal can either enhance or undermine motor behavior, the fieldworker rated the overall emotional state of the child during the testing of all the six gross motor milestones according to two scales [41]. First, the scale of consciousness was rated either as drowsy or as awake and alert. Second, the child's irritability was rated as being calm, fussy, or upset (crying).

Ideally, the child should be awake, alert, and calm during the assessments of motor skills. Drowsiness, fussiness, and crying were not reasons for not testing if the child was still able to display the milestone under testing. However, if they interfered with assessment, the child was retested when he or she was calm. If a child was asleep, he or she was not woken up to be tested.

In the context of the MGRS, the fieldworkers preferably tested the child on the motor skills after the completion of the anthropometric measurements. However, if the child was known to become upset by the anthropometric measurements, testing on motor milestones occurred prior to these measurements. If the caregiver and/or the child were obviously distraught or if the child was sick during a follow-up visit, testing did not occur.

Training and standardization of fieldworkers

Training

The MGRS fieldworkers selected to carry out the motor development assessments were trained at their own study site by an external expert prior to the initiation of data collection. The training involved lectures, discussions, observations, and assessments of a group of about 30 children (aged 5–13 months). It consisted of two days of initial training, one day of evaluation of the trainees, and two days of guided home visits. During the two-day initial training, the assessments carried out by the trainees were videotaped and reviewed afterwards by the trainer and trainees. The evaluation session (which was recorded as well) assessed the trainees' ability to score the achievement of the six motor milestones. This session involved both trainer and trainees. The trainer tested and scored approximately 10 to 15 children (aged 5–13 months) and did not give any indication of the children's scores to the caregivers or trainees. The trainees observed the child being tested and independently scored the child's performance on each tested skill. After the evaluation, the trainees' scores were compared with the trainer's scores, and in case of disagreements, these were discussed by looking at the videotaped session.

Standardization

The sites conducted regular half-day standardization sessions to determine the interobserver reliability rates of fieldworkers. During each session, one member of the fieldworkers' team tested and scored a group of about 10 children (aged 6–12 months) for the six motor milestones. The assessment and performance of the children were videotaped for subsequent scoring by the other fieldworkers at the same site. At each session, the fieldworker doing the actual testing was rotated so that a different person was the tester. The child's caregiver was present but was requested not to interfere with the assessments. However, when needed, the tester asked for the caregiver's assistance. The tester did not give any indication of the child's scores and wrote them on a standardization record form. Milestone performance could be rated as inability, refusal, ability, or unable to test, according to the established criteria (see below). The other fieldworkers watched the videotaped session and independently scored the performance of the same children on each of the six milestones.

After the conduct of each session, the videotape of the session and the fieldworkers' scores were sent to the Coordinating Centre of the MGRS at WHO in Geneva. The Motor Development Study coordinator on the Coordinating Centre team viewed the tape and scored the performance of the children. The scores given by the coordinator were considered to be the standard (true) scores. Interobserver reliability rates (percentages of agreement) were generated by calculating a correlation between the standard score and the scores obtained by the tester and the observers in a site.

The results of the sessions and comments on the observed disagreement between the standard score and a fieldworker's score, as well as on the tester's performance of the assessments, were sent as feedback to the site. A cutoff point of 90% agreement was set to determine whether further training was required.

Standardization of conditions for testing

It is well documented that child rearing practices [42] and encouragement by training and practice [28, 43] account for part of the variability in the achievement of motor milestones. Data collection in the study took place at the children's homes so that the standardization of the environment was limited. One source of variability, however, that could be controlled for was the social and physical context in which the child was tested and the nature of the objects used for testing. If physically possible and culturally appropriate, the number of persons present during testing was limited to three (fieldworker, caregiver, and child). If limitation of the number of people in the room was not possible, it was imperative that other observers did not move or make verbal comments during testing unless requested. Ideally, the surface of the floor where the assessments took place was clean and free of objects that might interfere with locomotion. Prior to testing, the fieldworker asked the caregiver to select a maximum of three toys or objects that the child liked to play with. It was primarily the fieldworker who carried out motor development assessment during the home visits. However, in some cases it was necessary for the fieldworker to ask for the caregiver's help.

Data collection

The data were recorded by the child's caregivers between follow-up visits and by the trained follow-up team members during these scheduled visits to the children's homes.

Caregiver

At the four-month follow-up visit, the caregiver was informed about the Motor Development Study and asked to start observing and assessing the child's motor developmental level until the child had achieved all six milestones. The caregiver was told to place the child in the appropriate position according to the defined testing procedures as soon as the caregiver observed that the child was making the first movements toward the achievement of a particular milestone. No fixed order of milestone achievement was assumed.

The record form for the caregiver had one page and presented the six drawings of the milestones (figs.

1–6), along with the performance criteria. A date box for each milestone was given, in which the caregiver recorded the date the child met the criteria for this item and thus achieved it for the first time. As soon as the caregiver had recorded the dates of first appearance of all six milestones, the caregiver stopped the motor development assessments.

Fieldworker

The follow-up team member trained in motor development assessments tested and scored all of the six gross motor milestones at each home visit. When both fieldworkers doing the home visits had been trained in motor development assessment, only one of them carried out the assessment and scored the child without the involvement of the other fieldworker. It was not necessary that the same fieldworker carry out all the motor development assessments for a given child. Motor development assessments were carried out monthly during the first year of life, starting from the five-month visit, and then every two months in the second year of life until the child acquired the skill of independent walking. If at the four-month visit, the time point when the fieldworker informed the caregiver about the study, the fieldworker observed that a child had achieved a certain milestone or a caregiver reported its achievement, then the fieldworker started the assessment at that visit. The reasons for examining all the milestones at each home visit were standardization of data collection across study sites, the fact that motor milestones might not occur in a sequential way in all subjects, and the fact that some milestones might be observed and then inhibited later (e.g., after an illness or trauma).

The performance of each milestone was evaluated independently by using four coding possibilities: inability—the child tried but failed to perform the test item because it surpassed his or her developmental level; refusal—the child was calm and alert but just refused to cooperate; ability—the child performed the test item according to the specified criteria; and unable to test—the child could not be tested on this milestone because his or her emotional state (drowsiness, fussiness, or crying) was interfering with testing, the child was sick, or the child's caregiver was distraught. In practice, it was somewhat difficult to differentiate between "refusal" and "unable to test."

The fieldworker took about 10 minutes to test all milestones. Since it was not always possible to get the child's cooperation immediately, the fieldworker was allowed three trials for the assessment of each milestone. The fieldworkers were given no ages at which the infants were expected to achieve each milestone, as this might have influenced their judgment.

For milestones that had not been achieved by the 12-month visit, the fieldworkers called the caregivers in the months with no scheduled follow-up visit during the second year of follow-up (i.e., months 13, 15, and 17). The fieldworker asked whether the child had achieved a specific milestone and reminded the caregiver to fill out the parent's record form. If the child had achieved a specific milestone, the fieldworker verified this by going through the criteria with the caregiver on the phone. Afterwards, at the planned home visit the following month, the fieldworker checked the acquisition of the reported milestone. Figure 7 shows the data collection form used by the fieldworkers for the motor development assessment.

Parental recording

At each visit, the fieldworker asked the caregiver about the milestones achieved since the previous follow-up visit and obtained the date that the caregiver had written down on the record form. If it was found on examination that the milestone(s) reported by the caregiver had not actually been attained by the child, the fieldworker carefully discussed this with the caregiver and explained the criteria again to make sure that the caregiver understood the criteria for the specific milestone. If the caregiver agreed that the child did not fulfill all the criteria, the fieldworker drew a new date line below the recorded date on the parent's record form for the milestone involved and asked the caregiver to record the date when this milestone was achieved according to the established criteria. If, on the other hand, the caregiver was sure that the child had met the criteria for the milestone, the fieldworker transferred to the form (fig. 7) the first written date as the caregiver's recorded date. The fieldworker also verified whether the caregiver had actually tested and recorded the date or simply recalled the date of first achievement. If a child happened to perform the motor skill for the first time at a certain home visit, this date was entered as the caregiver's date. The fieldworker never told the caregiver when a child should be achieving a particular milestone or gave any indication about which milestones the caregiver should be looking for as the child got older.

Data quality control

Data quality assurance started with the fieldworkers carefully filling out the record forms and checking for completeness and accuracy. Additional checks were made by data quality control staff and supervisors at the sites. Extensive quality checking was carried out on the data accumulated at the Coordinating Centre. A printout of the complete set of Motor Development Study records for each child was checked periodically for inconsistencies, such as missing or incorrectly entered caregiver's dates, reported caregiver's date without confirmation of a milestone's achievement by a fieldworker, discontinuation of the Motor Development Study without observed achievement of all

WORLD HEALTH ORGANIZATION

Multicentre Growth Reference Study

Project NUT01 – Motor development assessment

MDS

Identification

(a) Form code: MDS

(b) Study number: NUT01

(c) Site number: ☐

(d) Subject code: ☐☐☐☐☐-☐

(e) Follow-up visit number: ☐☐

(f) Continued testing required? ☐
1 = No 2 = Yes

(g) Date of visit: ☐☐/☐☐/☐☐
Day Month Year

(h) Examiner's code: ☐☐☐

Remarks

Test items	Examiner report: (a) Observed *1 = No (inability)* *2 = No (refusal)* *3 = Yes* *9 = Unable to test*	Caretaker report: (b) Precise date of first achievement *(Taken from the parent's record form). Only enter date(s) for milestone(s) achieved for the first time between the previous follow-up and present visit.*	Caretaker report: (c) Type of record *1 = Tested and recorded* *2 = Recalled*
1. Sitting without support	☐	☐☐/☐☐/☐☐ *Day Month Year*	☐
2. Hands-and-knees crawling	☐	☐☐/☐☐/☐☐ *Day Month Year*	☐
3. Standing with assistance	☐	☐☐/☐☐/☐☐ *Day Month Year*	☐
4. Walking with assistance	☐	☐☐/☐☐/☐☐ *Day Month Year*	☐
5. Standing alone	☐	☐☐/☐☐/☐☐ *Day Month Year*	☐
6. Walking alone	☐	☐☐/☐☐/☐☐ *Day Month Year*	☐

7. Child's emotional state *Rate the child's emotional state during the testing of all the milestones.* *Enter a code for each of the two scales.*	(a) First scale *1= Drowsy* *2= Awake and alert* ☐	(b) Second scale *1= Calm* *2= Fussy* *3= Crying* ☐

FIG. 7. Fieldworker's data collection form for motor development assessment

six milestones, or order of milestone achievement (e.g., walking alone before walking with assistance). The reported inhibition of a milestone was queried as well as differences between a Motor Development Study home visit date and a follow-up visit date. The inconsistencies were sent to the sites for investigation, verification, and correction at the source. This process of data verification and cleaning between the sites and the Coordinating Centre was continued until all data queries had been satisfactorily resolved. Detailed descriptions of the MGRS data management procedures are given elsewhere in this supplement [44].

Conclusions

The Motor Development Study aimed to describe the acquisition of six universal gross motor milestones in the first two years of life among affluent children growing up in different cultural settings, and thereby fill an existing gap in knowledge. The uniqueness of this study includes the opportunity to link growth and motor development in one international reference. The same protocol was used in the five countries that participated in the study, and the motor development assessments were performed by standardized fieldworkers. This is expected to minimize the influence of respondent biases on the outcome. At the same time, having caregiver records of the exact dates of milestone achievement facilitates internal cross-validation with fieldworkers' records and comparison of the MGRS data with previous studies that relied on parental reporting alone. Achievement of the six milestones was assessed repeatedly between 4 and 24 months of age, which will make it possible to describe their sequence and tempo in addition to the ages when milestones were acquired. The availability in the MGRS of information on breastfeeding and complementary feeding will also permit studies of associations between child feeding and motor development.

Although the study was conducted in a standardized manner, it also had limitations. We did not collect information on stimulation and child rearing practices that might influence milestone acquisition [28, 42, 43]. Thus, although it will be possible to examine associations between motor development and child feeding, morbidity, and overall physical growth, assessment of the possible influence of psychosocial stimulation on the reported outcomes will be limited to the examination of their ecological associations with the socioeconomic and demographic profiles found in the MGRS. Despite this limitation, this study provides an important addition to the literature on gross motor development in different cultural settings and should serve as a baseline for more focused studies of both motor and cognitive development.

References

1. Phatak A, Phatak P. Development of Indian babies and its assessment. In: Udani PM, ed. Textbook of pediatrics. With special reference to problems of child health in developing countries. New Delhi: Jaypee Brothers, 1991: 126–36.
2. Hellbrugge T, Pohl P. Munich functional diagnostic tests and early behavioural diagnosis. In: Udani PM, ed. Textbook of pediatrics. With special reference to problems of child health in developing countries. New Delhi: Jaypee Brothers, 1991:136–43.
3. Thelen E. Motor development. A new synthesis. Am Psychol 1995;50:79–95.
4. Bhandari A, Ghosh BN. A longitudinal study on child development in relation to socio-economic factors. Indian J Med Res 1980;72:677–84.
5. Capute AJ, Shapiro BK, Palmer FB, Ross A, Wachtel RC. Normal gross motor development: the influences of race, sex and socio-economic status. Dev Med Child Neurol 1985;27:635–43.
6. Carruth BR, Skinner JD. Feeding behaviors and other motor development in healthy children (2–24 months). J Am Coll Nutr 2002;21:88–96.
7. Cheung YB, Yip PSF, Karlberg JPE. Fetal growth, early postnatal growth and motor development in Pakistani infants. Int J Epidemiol 2001;30:66–74.
8. Das VK, Sharma NL. Developmental milestones in a selective sample of Lucknow children. A longitudinal study. Indian J Pediatr 1973;40:1–7.
9. Grantham-McGregor SM, Back EH. Gross motor development in Jamaican infants. Dev Med Child Neurol 1971; 13:79–87.
10. Hindley CB, Filliozat AM, Klackenberg G, Nicolet-Meister D, Sand EA. Differences in age of walking in five European longitudinal samples. Hum Biol 1966;38: 364–79.
11. Iloeje SO, Obiekwe VU, Kaine WN. Gross motor development of Nigerian children. Ann Trop Paediatr 1991;11:33–9.
12. Jahari AB, Saco-Pollitt C, Husaini MA, Pollitt E. Effects of an energy and micronutrient supplement on motor development and motor activity in undernourished children in Indonesia. Eur J Clin Nutr 2000;54(suppl 2):S60–8.
13. Largo RH, Molinari L, Weber M, Comenale Pinto L, Duc G. Early development of locomotion: significance of prematurity, cerebral palsy and sex. Dev Med Child Neurol 1985;27:183–91.
14. Neligan G, Prudham D. Norms for four standard developmental milestones by sex, social class and place in family. Dev Med Child Neurol 1969;11:413–22.
15. Palisano RJ. Use of chronological and adjusted ages to compare motor development of healthy preterm and fullterm infants. Dev Med Child Neurol 1986;28: 180–7.
16. Persson K, Strömberg B. Structured observation of motor performance (SOMP-I) applied to neonatally healthy fullterm infants at the ages of 0–10 months. Early Hum Dev 1995;40:127–43.
17. Phatak P. Motor and mental development of Indian babies from 1 month to 30 months. Indian Pediatr 1969; 6:18–23.
18. Yaqoob M, Ferngren H, Jalil F, Nazir R, Karlberg J. Early

child health in Lahore, Pakistan: XII. Milestones. Acta Paediatr Suppl 1993;390:151–7.
19. Touwen B. Neurological development in infancy. Clinics in Developmental Medicine, No. 58. London: William Heinemann Medical Books, 1976.
20. Vaivre-Douret L, Burnod Y. Development of a global motor rating scale for young children (0–4 years) including eye-hand grip coordination. Child Care Health Dev 2001;27:515–34.
21. World Health Organization Task Force for Epidemiological Research on Reproductive Health, Special Programme of Research, Development and Research Training in Human Reproduction. Progestogen-only contraceptives during lactation: II. Infant development. Contraception 1994;50:55–68.
22. de Onis M, Garza C, Victora CG, Onyango AW, Frongillo EA, Martines J, for the WHO Multicentre Growth Reference Study Group. The WHO Multicentre Growth Reference Study: planning, study design, and methodology. Food Nutr Bull 2004;25(1)(suppl 1):S15–26.
23. Lartey A, Owusu WB, Sagoe-Moses I, Gomez V, Sagoe-Moses C, for the WHO Multicentre Growth Reference Study Group. Implementation of the WHO Multicentre Growth Reference Study in Ghana. Food Nutr Bull 2004;25(1)(suppl 1):S60–5.
24. Bhandari N, Taneja S, Rongsen T, Chetia J, Sharma P, Bahl R, Kashyap DK, Bhan MK, for the WHO Multicentre Growth Reference Study Group. Implementation of the WHO Multicentre Growth Reference Study in India. Food Nutr Bull 2004;25(1)(suppl 1):S66–71.
25. Baerug A, Bjoerneboe G-E A, Tufte E, Norum KR, for the WHO Multicentre Growth Reference Study Group. Implementation of the WHO Multicentre Growth Reference Study in Norway. Food Nutr Bull 2004;25(1)(suppl 1):S72–7.
26. Prakash NS, Mabry RM, Mohamed AJ, Alasfoor D, for the WHO Multicentre Growth Reference Study Group. Implementation of the WHO Multicentre Growth Reference Study in Oman. Food Nutr Bull 2004;25(1)(suppl 1):S78–83.
27. Dewey KG, Cohen RJ, Nommsen-Rivers LA, Heinig MJ, for the WHO Multicentre Growth Reference Study Group. Implementation of the WHO Multicentre Growth Reference Study in the United States. Food Nutr Bull 2004;25(1)(suppl 1):S84–9.
28. Bergenn VW, Dalton TC, Lipsitt LP. Myrtle B. McGraw: a growth scientist. Dev Psychol 1992;28:381–95.
29. Illingworth RS. The normal child: Some problems of the early years and their treatment, 10th ed. Edinburgh, Scotland: Churchill Livingstone, 1991.
30. Zuckerman BS, Frank DA. Infancy. In: Levine MD, Carey WB, Crocker AC, Gross RT, eds. Developmental-behavioral pediatrics. Philadelphia, Pa, USA: W.B. Saunders, 1983:81–96.
31. Appendix on neurological development in infancy: Griffith's Development Scale for the first two years of life. In: Udani PM, ed. Textbook of pediatrics. With special reference to problems of child health in developing countries. New Delhi: Jaypee Brothers, 1991:148–63.
32. Bayley Scales of Infant Development. Motor scale record form, 2nd ed. San Antonio, Tex, USA: Psychological Corporation, 1993.
33. Frankenburg WK, Dodds JB. Denver II training manual. Denver, Col, USA: Denver Developmental Materials, 1990.
34. Knobloch H, Pasamanick B, Sherard ES. A developmental screening inventory for infants. Pediatrics 1966;38: 1095–1108.
35. Zdanska-Brincken M, Wolanski N. A graphic method for the evaluation of motor development in infants. Dev Med Child Neurol 1969;11:228–41.
36. Capute AJ, Accardo PJ. The infant neurodevelopmental assessment: a clinical interpretive manual for CAT-CLAMS in the first two years of life, part 2. Curr Probl Pediatr 1996;26:279–306.
37. Gesell A, Amatruda CS. Developmental diagnosis, 2nd ed. New York: Paul B. Hoeber, 1947.
38. Kenny TJ, Clemmens RL. Behavioral pediatrics and child development: a clinical handbook, 2nd ed. Baltimore, Md, USA: Williams & Wilkins, 1980.
39. Needlman RD. Growth and development. In: Nelson WE, Behrman RE, Kliegman RM, Arvin AM, eds. Nelson textbook of pediatrics, 15th ed. Philadelphia, Pa, USA: W.B. Saunders, 1996:30–45.
40. Nelms BC, Mullins RG. Growth and development. A primary health care approach. Englewood Cliffs, NJ, USA: Prentice-Hall, 1982.
41. Brazelton TB. Echelle d'évaluation du comportement néonatal. Neuropsychiatr Enfance Adolesc 1983;31:61–96.
42. Werner EE. Infants around the world: cross-cultural studies of psychomotor development from birth to two years. J Cross Cult Psychol 1972;3:111–34.
43. Super CM. Environmental effects on motor development: the case of 'African infant precocity.' Dev Med Child Neurol 1976;18:561–7.
44. Onyango AW, Pinol AJ, de Onis M, for the WHO Multicentre Growth Reference Study Group. Managing data for a multicountry longitudinal study: experience from the WHO Multicentre Growth Reference Study. Food Nutr Bull 2004;25(1)(suppl 1):S46–52.