October 2015

Monitoring human growth and development: a continuum from the womb to the classroom

José Villar
University of Oxford

Aris T. Papageorghiou
University of Oxford

Ruyan Pang
Peking University

Laurent J. Salomon
Université Paris Descartes

Ana Langer
Harvard School of Public Health

See next page for additional authors

Authors

José Villar, Aris T. Papageorghiou, Ruyan Pang, Laurent J. Salomon, Ana Langer, Cesar Victora, Manorama Purwar, Cameron Chumlea, Wu Qingqing, Sicco A. Scherjon, Fernando C. Barros, Maria Carvalho, Douglas G. Altman, Francesca Giuliani, Enrico Bertino, Yasmin A. Jaffer, Leila Cheikh Ismail, Eric O. Ohuma, Ann Lambert, J. Alison Noble, Michael G. Gravett, Zulfiqar A. Bhutta, and Stephen H. Kennedy

OBSTETRICS

Monitoring human growth and development: a continuum from the womb to the classroom

José Villar, MD; Aris T. Papageorghiou, MD; Ruyan Pang, MD; Laurent J. Salomon, PhD; Ana Langer, MD;
Cesar Victora, MD; Manorama Purwar, MD; Cameron Chumlea, PhD; Wu Qingqing, MD; Sicco A. Scherjon, MD;
Fernando C. Barros, MD; Maria Carvalho, MD; Douglas G. Altman, DSc; Francesca Giuliani, MD; Enrico Bertino, MD;
Yasmin A. Jaffer, MD; Leila Cheikh Ismail, PhD; Eric O. Ohuma, MSc; Ann Lambert, PhD; J. Alison Noble, DPhil;
Michael G. Gravett, MD; Zulfiqar A. Bhutta, PhD; Stephen H. Kennedy, MD

A comprehensive set of fully integrated anthropometric measures is needed to evaluate human growth from conception to infancy so that consistent judgments can be made about the appropriateness of fetal and infant growth. At present, there are 2 barriers to this strategy. First, descriptive reference charts, which are derived from local, unselected samples with inadequate methods and poor characterization of their putatively healthy populations, commonly are used rather than prescriptive standards. The use of prescriptive standards is justified by the extensive biologic, genetic, and epidemiologic evidence that skeletal growth is similar from conception to childhood across geographic populations, when health, nutrition, environmental, and health care needs are met. Second, clinicians currently screen fetuses, newborn infants, and infants at all levels of care with a wide range of charts and cutoff points, often with limited appreciation of the underlying population or quality of the study that generated the charts. Adding to the confusion, infants are evaluated after birth with a single prescriptive tool: the World Health Organization Child Growth Standards, which were derived from healthy, breastfed newborn infants, infants, and young children from populations that have been exposed to few growth-restricting factors. The International Fetal and Newborn Growth Consortium for the 21st Century Project addressed these issues by providing international standards for gestational age estimation, first-trimester fetal size, fetal growth, newborn size for gestational age, and postnatal growth of preterm infants, all of which complement the World Health Organization Child Growth Standards conceptually, methodologically, and analytically. Hence, growth and development can now, for the first time, be monitored globally across the vital first 1000 days and all the way to 5 years of age. It is clear that an integrative approach to monitoring growth and development from pregnancy to school age is desirable, scientifically supported, and likely to improve care, referral patterns, and reporting systems. Such integration can be achieved only through the use of international growth standards, especially in increasingly diverse, mixed ancestry populations. Resistance to new scientific developments has been hugely problematic in medicine; however, we are confident that the obstetric and neonatal communities will join their pediatric colleagues worldwide in the adoption of this integrative strategy.

Key words: continuity of care, growth monitoring, prescriptive standards

Optimizing growth and development from conception to childhood through good nutrition, a clean environment, and adequate holistic health care is essential for the improvement of health and economic development of populations.[1,2] Throughout this critical period, growth, which is a continuous process, must be monitored routinely with the use of congruent screening tools and criteria. Focusing conceptually on only 1 specific phase or a single summary value (eg, late fetal growth or estimated fetal weight) has limited biologic basis and lessens the chances for timely and appropriate interventions.

Growth monitoring (GM) is an integral and undisputed component of evidence-based antenatal and newborn care worldwide, as it is for infants and children. However, to be effective, GM requires a comprehensive set of anthropometric standards that enable skeletal growth (eg, fetal head circumference or postnatal length) and fat-related markers (eg, fetal abdominal circumference or postnatal weight) to be assessed longitudinally so that judgments can be made about the appropriateness of growth patterns and deviations (eg, whether the fetus/newborn/infant is wasted, stunted, or overweight/obese).[3] These tools have been available to evaluate term infants'

postnatal growth, but not fetal growth, newborn size, or the postnatal growth of preterm newborn infants.

Two barriers thwart the implementation of the missing perinatal component of the GM strategy. First, unlike in most other fields of medicine, descriptive reference charts, rather than prescriptive standards, are used in obstetric and neonatal practice. Standards are preferable because they describe aspirational, biologic norms that are achieved by healthy populations and individuals throughout the world. References, on the other hand, describe the distribution of variables that are observed in unselected samples at a given time and place, often decades ago.[4] They provide information that is of limited value today for clinicians, parents, and families because the criteria used to select subjects and define their health were often ill-defined. This applies, for example, to the Hadlock charts of estimated fetal weight, which are presently used worldwide, evaluated in 109 fetuses from a hospital in Texas in the 1980s.[5]

The second barrier is the large number and limited methodologic quality of the charts that are available to obstetricians and neonatologists. In a series of systematic reviews, we showed that there are (1) 29 published charts for estimating gestational age that use crown rump length, of which only 4 satisfied minimum quality criteria[6]; (2) 83 published fetal size charts for monitoring growth by ultrasound scanning, of which only 12 used a reliable dating method[7]; (3) 102 published charts of birthweight for gestational age, of which only 8 satisfied minimum quality criteria,[8] and

(4) 61 published postnatal growth charts for preterm infants with considerable shortcomings in the quality of anthropometric evaluation, gestational age estimation, length of follow-up evaluation, and reportage of postnatal care, feeding regimes, and morbidities. In addition, the choice of a particular reference chart is too often based on clinicians' preferences or on the default chart offered in the ultrasound machine's software, which can lead to different references being used even within the same medical practice. Finally, there is considerable variability in the definition of intrauterine growth restriction (IUGR) with the use of combinations of ultrasound measures.

Consequently, clinicians currently monitor fetuses and newborn infants at different levels of care and institutions with a wide range of charts. Variable cutoff points (3rd, 5th, 10th, 90th, 95th, or 97th percentiles) are used to define "normality," macrosomia, and IUGR. Often there is little appreciation of the underlying population or quality of the study that generated the chart being used. Consequently, fetuses can be classified as growth-restricted or overgrown in 1 part of a city or country and as healthy in another. For example, Salomon et al[9] showed that the proportion of fetuses who were classified as having a biparietal diameter below the 5th percentile at 20-24 weeks of gestation can range from 6.6—23.7% with the use of 3 different popular ultrasound reference charts. More confusing still is that, only 1 month after birth, infants are evaluated by pediatricians with the use of a single, prescriptive tool, the

World Health Organization (WHO) Child Growth Standards, that was derived from healthy newborn infants from populations with few growth-restricting factors whose mothers followed breastfeeding recomendations.[10] This situation must result inevitably in diagnostic inaccuracies, confusion for parents, and inappropriate interventions.

In an era of evidence-based medicine, supporting an approach in which different tools and criteria are used depending on whether the babies are inside the womb, newly born, or aged ≥1 month is difficult to justify scientifically and to parents. The fact that the WHO Child Growth Standards have been adopted in >125 countries, including the United States, United Kingdom, and Norway,[11] demonstrates that standardization of care is feasible and acceptable across vastly different countries and medical systems.

The International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st) Project has produced an integrated set of standards for gestational age estimation,[12] first-trimester fetal size,[12] fetal growth[13] (to be supported by fetal growth velocity and estimated fetal weight standards in 2016), newborn size for gestational age,[14] postnatal growth of preterm infants,[15] and, in 2016, infant development at 2 years old.[16] The percentiles of the INTERGROWTH-21st newborn standards are remarkably similar at term to those of the WHO Child Growth Standards (Figure 1), which is not surprising because the 2 studies adopted the same conceptual, methodologic, and analytic approaches. This means that growth and

Programa de Pós-Graduação em Saúde e Comportamento, Universidade Católica de Pelotas, Pelotas, RS, Brazil (Dr Barros); Faculty of Health Sciences, Aga Khan University, Nairobi, Kenya (Dr Carvalho); Centre for Statistics in Medicine, Botnar Research Centre, University of Oxford, Oxford, UK (Drs Altman and Ohuma); Dipartimento di Scienze della Sanita Pubblica e Pediatriche, Struttura Complessa di Neonatologia Universitaria, Università degli Studi di Torino, Torino, Italy (Drs Giuliani and Bertino); Department of Family & Community Health, Ministry of Health, Muscat, Sultanate of Oman (Dr Jaffer); Global Alliance to Prevent Prematurity and Stillbirth (GAPPS), Seattle Children's, Seattle, WA (Dr Gravett); Center of Excellence in Women & Child Health, The Aga Khan University, Karachi, Pakistan and Center for Global Child Health, Hospital for Sick Children, Toronto, Canada (Dr Bhutta).
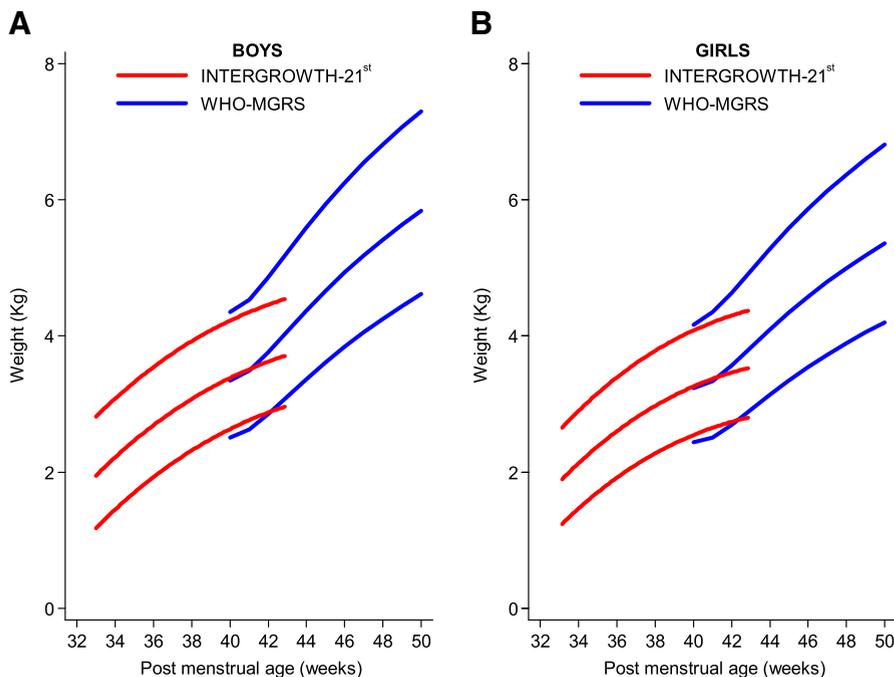
FIGURE 1

INTERGROWTH-21st Newborn Standards complementing the WHO Child Growth Standards

The 3rd, 50th, and 97th percentile curves for birthweight according to gestational age and sex (**A,** boys; **B,** girls) from the INTERGROWTH-21st Newborn Size Standards[14] (*red lines*) followed by the corresponding 3rd, 50th, and 97th percentile curves from the World Health Organization Child Growth Standards for term (40 weeks of gestation) newborn infants according to sex.[9]

WHO-MGRS, World Health Organization Child Growth Standards.

Villar. Monitoring human growth and development. Am J Obstet Gynecol 2015.

development, for the first time, can be monitored with the same high-quality tools across the vital first 1000 days and up to the age of 5 years.

The INTERGROWTH-21st standards were developed from a prospective, population-based project that selected 8 urban areas across 5 continents. Most inhabitants were healthy, adequately nourished, and educated with minimal environmental constraints on growth.[17] Within each of these study sites, in a

second sampling step, pregnant women at low risk of IUGR were recruited.[18] All of them had a reliable estimate of gestational age confirmed by ultrasound scanning. The same, specially adapted, ultrasound equipment was used at all sites to allow blinding of measurements. A novel quality-control strategy, which included centralized image storage, enabled the independent review of images and measurements.[19,20] Newborn infants were measured with the same methods and quality control protocols as in the WHO Child Growth Standards Study[21] with a follow-up study of their growth, diet, morbidity, and cognitive development[16] until 2 years of age from which a functional classification of IUGR and macrosomia will be generated. Reassuringly, the growth and development at 1 year of age of term and preterm newborn infants closely match the WHO Child Growth Standards.[13,15] Although we have presented our curves with commonly used statistical cutoffs (eg, 3rd and 97th percentiles), as is the case for infant and child growth standards,[10] we presently are conducting analyses to identify which cutoff points best predict perinatal and postnatal outcomes[14] for incorporation in an evidence-based triage for global perinatal care.

It could be argued that the observed geographic and ethnic variations in fetal growth across populations preclude the use of international standards. These arguments are difficult to support, given the widespread use in medicine of universal definitions based on data that are

TABLE

Variance component analysis for fetal newborn infant and child skeletal growth and size

Variable	Fetal ultrasound scan, %		Birth: newborn[22]	Infancy: preterm infant length[15]	Infancy and childhood: infant length[10]	Childhood: child height[23]
	Fetal crown-rump length[22]	Fetal head circumference[22]				
Variance between study sites, %	1.9	2.6	3.5	0.2	3.4	3.0
Variance between individuals within a site, %	—	18.6	—	57.1	70.0	—
Residual variance, %	98.1	78.8	96.5	42.7	26.6	—

Variance between individuals for these measures cannot be estimated because they were taken from cross-sectional data.

Villar. Monitoring human growth and development. Am J Obstet Gynecol 2015.

obtained from healthy subjects and their use for comparisons across populations. As an example, no one suggests that the definition of anemia should be altered locally for malnourished populations; in fact, international comparisons regarding the prevalence of anemia would be impossible without universally accepted hemoglobin cutoff points.

GM of the fetus, neonate, and child up to 5 years old should be done in the same way. International standards are justified by the evidence that skeletal and linear growth in humans is similar from conception to childhood across geographic populations, when health, nutrition, and health care needs are met. These data demonstrate that only <4% of the total variability in growth and size during pregnancy, at birth, during infancy, and childhood can be attributed to differences between populations, as opposed to variability among individuals within the same population (Table).[10,22,23]

The alternative (ie, the use of local charts based on unselected samples) involves major clinical and public health risks. First, if fetuses and newborn infants are evaluated in such a manner, the rate of IUGR or small-for-gestational age <10th percentile, by definition, will be close to 10% in all populations, regardless of the local prevalence of malnutrition or morbidity, which is

---

FIGURE 2
Sensitivity analyses of the populations in the INTERGROWTH-21st Project
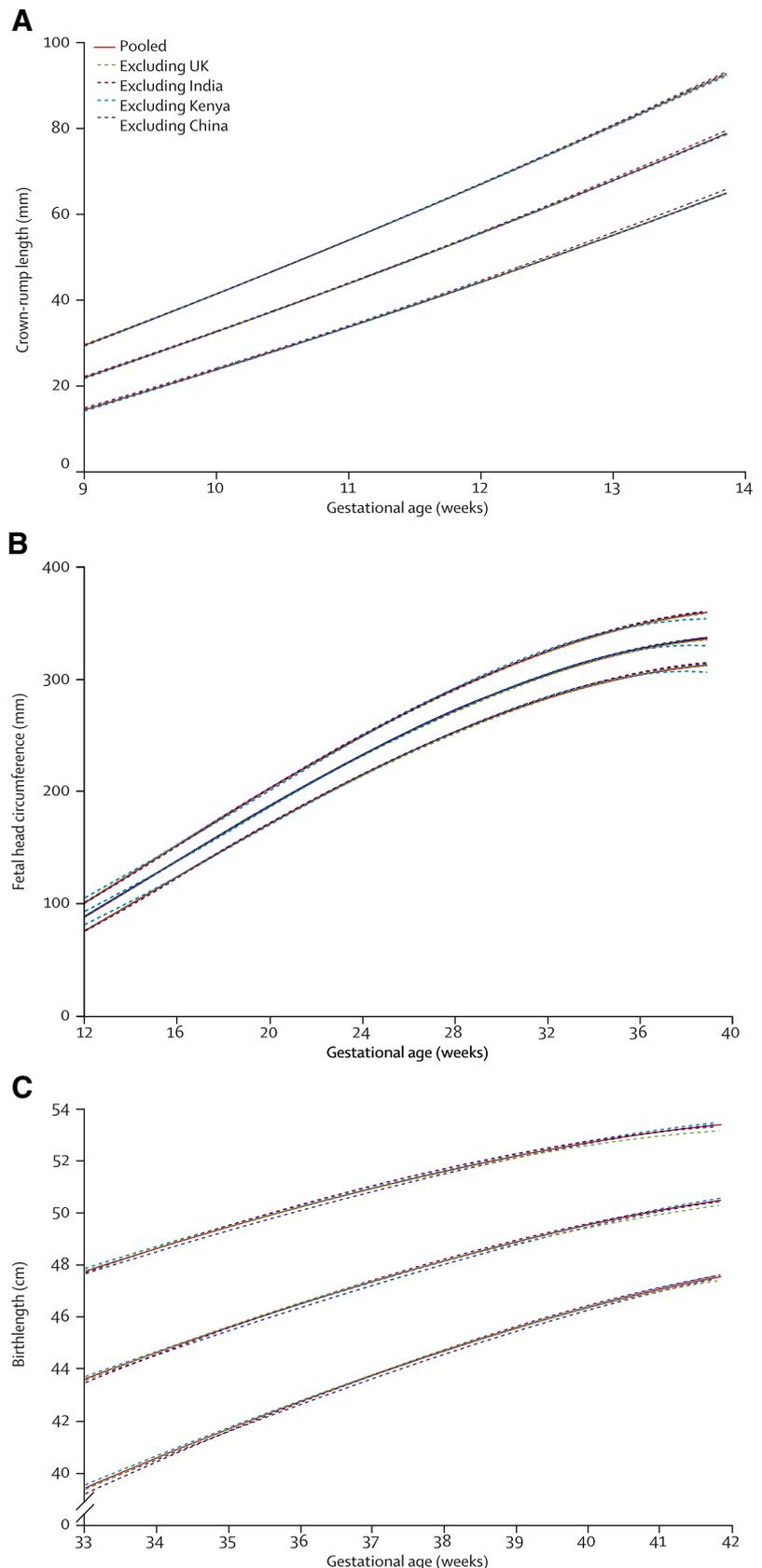
A and **B,** Crown rump length and fetal head circumference 3rd, 50th, and 97th percentiles for the total Fetal Growth Longitudinal Study population (*solid line*); the remaining sample after data from the sites in China, India, Kenya, and the UK were excluded one at a time. **C,** Birth length at 3rd, 50th, and 97th percentiles for the low-risk proportion of the total Newborn Cross-Sectional Study population (*solid line*), and the remaining sample after data from the sites in China, India, Kenya, and the UK were excluded one at a time. A detailed description of the studies and analytic strategy is available in the article by Villar et al.[22]

epidemiologically implausible. For example, recently it was estimated that the prevalence of small-for-gestational age in low-to-middle income countries is at least 24%[24] and the prevalence of stunting in these countries ranges from 10–40%,[2] which is in keeping with the corresponding rates of stillbirth and neonatal, infant, and <5-year morbidity and mortality.[25]

At the other end of the spectrum, whenever local unselected populations in developed regions are used as their own reference, overweight newborn infants are considered "normal" (ie, their early obesity is not diagnosed). This occurs because many countries and subpopulations, in the midst of the childhood obesity and diabetes mellitus epidemics,[26] have shifted their percentiles upwards to include bigger babies. For example, in England in 2011-2012 the rate of birthweight >90th percentile was 11% using local references but 19% using international standards.[27] Only by using a single set of international newborn size standards[14] can the 2 situations described earlier be avoided.

Another argument against the use of the international newborn size standards is the variability that is observed between the highest and lowest birthweights in the populations that are used for their construction. Clearly, some variability exists even in the healthy populations included in the INTERGROWTH-21st Project, but such a comparison is not relevant to the question of whether to use international growth standards. It was never suggested nor recommended that data from a single country should be used to evaluate another population or vice versa. The question is not how a single population of fetuses or newborn babies compares in length or weight with another population such as the United Kingdom's. The question, instead, is how each of these compares with the international standards that were constructed with all populations combined. Sensitivity analyses demonstrated that the exclusion of any of the populations in the WHO Child Growth Standards Study or INTERGROWTH-21st Project had negligible effect on the pooled values from all the study sites and that the

percentiles were indistinguishable when superimposed (Figure 2).[22]

Furthermore, recommending that growth charts should be ethnic/racial-specific, based on the concept that certain ethnic or racial groups have a genetic predisposition to small or large size at birth, has no scientific basis in nonisolated populations. First, ethnicity and race are social, not biologic, constructs. As a *Nature Biotechnology* 2005 editorial concluded: "Scientifically, race is a meaningless marker of anything. Pooling people in race silos is akin to zoologists grouping raccoons, tigers and okapis on the basis that they are all stripey."[28] These are not even well-defined terms: a systematic review identified 116 definitions of self-reported race or ethnicity in the medical literature.[29] Second, most populations have been subjected to marked genetic admixture, including the United States,[30] such that >6 million North Americans, who self-identify as European, probably carry African ancestry.[31] Third, nearly 700 genetic variants have been implicated in adult height (compared with the few associated with skin pigmentation), but even these only explain one-fifth of the heritability of height.[32] The concept is also contradicted by the overwhelming literature that links physical growth in humans to socioeconomic, environmental, health, and nutrition conditions worldwide.[33] Hence, there is no scientific basis for the adaption of fetal or newborn growth charts based on the self-reported ethnicity of the mother, which is a procedure that makes even less practical sense when we consider the multiple permutations of mixed-ethnicity couples and the extent of present-day mass migration.

In summary, integrated monitoring of growth and development from pregnancy to school age is desirable, scientifically supported, and likely both to standardize and improve health care and referral patterns. Such integration can be achieved only through the use of international growth standards, especially in increasingly diverse mixed ancestry populations. These standards allow, for the first time, comparisons to be made across populations from

conception onwards. They also greatly facilitate our ability to address concerns about fetal development and associated health problems in later life by producing standardized global data instead of the present paltry mishmash.

However, the implementation of the standards requires giving consideration to the resource implications, which include the capacity of referral systems to manage the number and severity of identified cases (which depends on the cutoff points selected) and the prevalence of risk factors. Operating within a uniform system requires professional adaptation and coordination across levels of care. As never occurred when the large number of currently used fetal and neonatal charts were introduced, it will be important to monitor how the new standards are adopted and how deviations from optimal growth are interpreted. We are confident that, in time, the INTERGROWTH-21st standards will be incorporated into care and research[34,35] just as the WHO Child Growth Standards are used currently by pediatricians worldwide.[36] ∎

REFERENCES

1. Bhutta ZA, Das JK, Rizvi A, et al. Evidence-based interventions for improvement of maternal and child nutrition: what can be done and at what cost? Lancet 2013;382:452-77.
2. Black RE, Victora CG, Walker SP, et al. Maternal and child undernutrition and overweight in low-income and middle-income countries. Lancet 2013;382:427-51.
3. Victora CG, Villar J, Barros FC, et al. Anthropometric characterization of impaired fetal growth: risk factors for and prognosis of newborns with stunting or wasting. JAMA Pediatr 2015;169:e151431.
4. Bertino E, Milani S, Fabris C, De Curtis M. Neonatal anthropometric charts: what they are, what they are not. Arch Dis Child Fetal Neonatal Ed 2007;92:F7-10.
5. Hadlock FP, Harrist RB, Sharman RS, Deter RL, Park SK. Estimation of fetal weight with the use of head, body, and femur measurements-a prospective study. Am J Obstet Gynecol 1985;151:333-7.
6. Napolitano R, Dhami J, Ohuma EO, et al. Pregnancy dating by fetal crown-rump length: a systematic review of charts. BJOG 2014;121:556-65.
7. Ioannou C, Talbot K, Ohuma E, et al. Systematic review of methodology used in ultrasound studies aimed at creating charts of fetal size. BJOG 2012;119:1425-39.

8. Giuliani F, Spada E, Ohuma E, et al. Systematic review of the methodological quality of studies designed to create neonatal anthropometric charts. Acta Paediatr 2015 [Epub ahead of print] DOI: http://dx.doi.org/10.1111/apa.13112.

9. Salomon LJ, Bernard JP, Duyme M, Buvat I, Ville Y. The impact of choice of reference charts and equations on the assessment of fetal biometry. Ultrasound Obstet Gynecol 2005;25:559-65.

10. De Onis M, Garza C, Onyango AW, Martorell R. WHO child growth standards. Acta Paediatr Suppl 2006;450:1-101.

11. De Onis M, Onyango A, Borghi E, et al. Worldwide implementation of the WHO child growth standards. Public Health Nutr 2012;15:1603-10.

12. Papageorghiou AT, Kennedy SH, Salomon LJ, et al. International standards for early fetal size and pregnancy dating based on ultrasound measurement of crown-rump length in the first trimester of pregnancy. Ultrasound Obstet Gynecol 2014;44:641-8.

13. Papageorghiou AT, Ohuma EO, Altman DG, et al. International standards for fetal growth based on serial ultrasound measurements: the fetal growth longitudinal study of the INTERGROWTH-21st Project. Lancet 2014;384:869-79.

14. Villar J, Cheikh Ismail L, Victora CG, et al. International standards for newborn weight, length, and head circumference by gestational age and sex: the newborn cross-sectional study of the INTERGROWTH-21st Project. Lancet 2014;384:857-68.

15. Villar J, Giuliani F, Bhutta ZA, et al. Postnatal growth standards for infants born preterm: the preterm postnatal follow-up study of the INTERGROWTH-21st Project. Lancet Global Health 2015 [Accepted for publication].

16. Fernandes M, Stein A, Newton CR, et al. The INTERGROWTH-21st Project neurodevelopment package: a novel method for the multi-dimensional assessment of neurodevelopment in pre-school age children. PLoS One 2014;9:e113360.

17. Eskenazi B, Bradman A, Finkton D, et al. A rapid questionnaire assessment of environmental exposures to pregnant women in the INTERGROWTH-21st Project. BJOG 2013;120(suppl 2):129-38.

18. Villar J, Altman DG, Purwar M, et al. The objectives, design and implementation of the INTERGROWTH-21st Project. BJOG 2013;120(suppl 2):9-26.

19. Papageorghiou A, Sarris I, Ioannou C, et al. Ultrasound methodology used to construct the fetal growth standards in the INTERGROWTH-21st Project. BJOG 2013;120(suppl 2):27-32.

20. Sarris I, Ioannou C, Ohuma E, et al. Standardisation and quality control of ultrasound measurements taken in the INTERGROWTH-21st Project. BJOG 2013;120(suppl 2):33-7.

21. Cheikh Ismail L, Knight H, Ohuma E, et al. Anthropometric standardisation and quality control protocols for the construction of new, international, fetal and newborn growth standards: the INTERGROWTH-21st Project. BJOG 2013;120(suppl 2):48-55.

22. Villar J, Papageorghiou AT, Pang R, et al. The likeness of fetal growth and newborn size across non-isolated populations in the INTERGROWTH-21st Project: the fetal growth longitudinal study and newborn cross-sectional study. Lancet Diabetes Endocrinol 2014;2:781-92.

23. 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.

24. Kozuki N, Katz J, Christian P; Child Health Epidemiology Reference Group Preterm Birth-SGA Working Group. Comparison of US birth weight references and the International Fetal and Newborn Growth Consortium for the 21st Century standard. JAMA Pediatr 2015;169:e151438.

25. Liu L, Oza S, Hogan D, et al. Global, regional, and national causes of child mortality in 2000-13, with projections to inform post-2015 priorities: an updated systematic analysis. Lancet 2015;385:430-40.

26. Lobstein T, Jackson-Leach R, Moodie ML, et al. Child and adolescent obesity: part of a bigger picture. Lancet 2015;385:2510-20.

27. Hirst JE, Villar J, Papageorghiou AT, Ohuma EO, Kennedy SH. Action to prevent childhood obesity should start before pregnancy. Lancet 2015 [Accepted for publication].

28. Illuminating BiDil. Nat Biotechnol 2005;23:903.

29. Ma IW, Khan NA, Kang A, Zalunardo N, Palepu A. Systematic review identified suboptimal reporting and use of race/ethnicity in general medical journals. J Clin Epidemiol 2007;60:572-8.

30. Montinaro F, Busby GB, Pascali VL, Myers S, Hellenthal G, Capelli C. Unravelling the hidden ancestry of American admixed populations. Nat Commun 2015;6:6596.

31. Bryc K, Durand EY, Macpherson JM, Reich D, Mountain JL. The genetic ancestry of African Americans, Latinos, and European Americans across the United States. Am J Hum Genet 2015;96:37-53.

32. Wood AR, Esko T, Yang J, et al. Defining the role of common variation in the genomic and biological architecture of adult human height. Nat Genet 2014;46:1173-86.

33. Coffey D, Deaton A, Dreze J, Spears D, Tarozzi A. Stunting among children: facts and implications. Economic and Political Weekly 2013;48:68-70.

34. McCarthy EA, Walker SP. International fetal growth standards: one size fits all. Lancet 2014;384:835-6.

35. Dickinson JE. Fetal growth: too big, too little, just right. Aust N Z J Obstet Gynaecol 2015;55:1-2.

36. Stevens GA, Finucane MM, Paciorek CJ, et al. Trends in mild, moderate, and severe stunting and underweight, and progress towards MDG 1 in 141 developing countries: a systematic analysis of population representative data. Lancet 2012;380:824-34.