May 2013

The objectives, design and implementation of the INTERGROWTH-21st Project

Villar J
University of Oxford

Altman D. G
University of Oxford

Purwar M
Ketkar Nursing Home, Nagpur, India

Noble J. A
University of Oxford

Knight H. E
University of Oxford

See next page for additional authors

Authors
Villar J, Altman D. G, Purwar M, Noble J. A, Knight H. E, Ruyan P, Ismail L. Cheikh, Barros F. C, Lambert A, Papageorghiou A. T, Carvalho M, Jaffer Y. A, Bertino E, Gravett M. G, Bhutta Z. A, and Kennedy S. H

DOI: 10.1111/1471-0528.12047
www.bjog.org

Original article

The objectives, design and implementation of the INTERGROWTH-21st Project

J Villar,[a] DG Altman,[b] M Purwar,[c] JA Noble,[d] HE Knight,[a] P Ruyan,[e] L Cheikh Ismail,[a] FC Barros,[f,g] A Lambert,[a] AT Papageorghiou,[a] M Carvalho,[h] YA Jaffer,[i] E Bertino,[j] MG Gravett,[k] ZA Bhutta,[l,*] SH Kennedy,[a,*] for the International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st)

[a] Nuffield Department of Obstetrics & Gynaecology, and Oxford Maternal & Perinatal Health Institute, Green Templeton College, University of Oxford, Oxford, UK [b] Centre for Statistics in Medicine, University of Oxford, Oxford, UK [c] Ketkar Nursing Home, Nagpur, India [d] Department of Engineering Science, University of Oxford, Oxford, UK [e] School of Public Health, Peking University, Beijing, China [f] Programa de Pós-Graduação em Saúde e Comportamento, Universidade Católica de Pelotas, Pelotas, RS, Brazil [g] Programa de Pós-Graduação em Epidemiologia, Universidade Federal de Pelotas, Pelotas, RS, Brazil [h] Faculty of Health Sciences, Aga Khan University, Nairobi, Kenya [i] Department of Family & Community Health, Ministry of Health, Muscat, Sultanate of Oman [j] Dipartimento di Scienze Pediatriche e dell'Adolescenza, Cattedra di Neonatologia, Università degli Studi di Torino, Torino, Italy [k] University of Washington School of Medicine, Seattle, WA, USA [l] Division of Women & Child Health, The Aga Khan University, Karachi, Pakistan
Correspondence: Prof J Villar, Nuffield Department of Obstetrics & Gynaecology, University of Oxford, Women's Centre, Level 3, John Radcliffe Hospital, Headington, Oxford OX3 9DU, UK. Email jose.villar@obs-gyn.ox.ac.uk
[*] Joint senior authors.

Accepted 16 September 2012.

INTERGROWTH-21st is a multicentre, multiethnic, population-based project, being conducted in eight geographical areas (Brazil, China, India, Italy, Kenya, Oman, UK and USA), with technical support from four global specialised units, to study growth, health and nutrition from early pregnancy to infancy. It aims to produce prescriptive growth standards, which conceptually extend the World Health Organization (WHO) Multicentre Growth Reference Study (MGRS) to cover fetal and newborn life. The new international standards will describe: (1) fetal growth assessed by clinical and ultrasound measures; (2) postnatal growth of term and preterm infants up to 2 years of age; and (3) the relationship between birthweight, length and head circumference, gestational age and perinatal outcomes. As the project has selected healthy cohorts with no obvious risk factors for intrauterine growth restriction, these standards will describe how all fetuses and newborns *should* grow, as opposed to traditional charts that describe how some have grown at a given place and time. These growth patterns will be related to morbidity and mortality to identify levels of perinatal risk. Additional aims include phenotypic characterisation of the preterm and impaired fetal growth syndromes and development of a prediction model, based on multiple ultrasound measurements, to estimate gestational age for use in pregnant women without access to early/frequent antenatal care.

Keywords Fetal, growth, INTERGROWTH-21st, newborn, nutrition, preterm, standards.

Please cite this paper as: Villar J, Altman D, Purwar M, Noble J, Knight H, Ruyan P, Cheikh Ismail L, Barros F, Lambert A, Papageorghiou A, Carvalho M, Jaffer Y, Bertino E, Gravett M, Bhutta Z, Kennedy S, for the International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st). The objectives, design and implementation of the INTERGROWTH-21st Project. BJOG 2013;120 (Suppl. 2): 9–26.

Introduction

Antenatal and postnatal care consist mostly of a series of screening tests of varied complexity, implemented at different levels of care, which together contribute to evaluating the overall health and nutritional status of each pregnant woman and newborn baby.[1,2] Objective assessments of fetal and neonatal growth deviations can play a major role in routine clinical care, as well as maternal and neonatal health research. The usefulness and limitations of such screening methods have been evaluated in randomised controlled trials over the last decade.[3,4] In some pregnancies and newborns, especially those that are preterm, there is a need to monitor growth more closely to decide if clinical interventions are required. However, international growth charts analogous to the World Health Organization

(WHO) Multicentre Growth Reference Study (MGRS) standards for infants and young children,[5,6] developed using a prescriptive approach, are not available.

This paper summarises the research strategy, methodology and implementation processes across eight geographically defined study sites (Pelotas, Brazil; Beijing, China; Nagpur, India; Turin, Italy; Nairobi, Kenya; Muscat, Oman; Oxford, UK and Seattle, USA) and the Project Coordinating Unit, (Nuffield Department of Obstetrics & Gynaecology, University of Oxford, UK), for the studies that comprise the International Fetal & Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st) Project. It complements the detailed protocol, data collection forms and operation manuals being used during the project itself that have been available on our website (www.intergrowth21.org.uk) from the outset.

Background

In 2006, WHO released its Child Growth Standards for children aged 0–5 years, which were generated by the WHO MGRS.[5,6] Two characteristics made MGRS unique and unprecedented in its field: (1) the study included populations from several countries (Brazil, Ghana, India, Norway, Oman and the USA) and (2) a prescriptive approach was used to select the study populations, i.e. only children from populations with minimal environmental constraints on growth were included. This was achieved by recruiting children of affluent and educated parents, because high education and family income have been identified as the environmental variables most likely to be associated with optimal child growth. In addition, chronic illness, unwillingness to adhere to MGRS feeding recommendations and maternal smoking were used at study entry as exclusion criteria.

By virtue of these characteristics, MGRS provided the strong scientific foundations for developing standards that indicate how children should grow, as opposed to previous studies that simply described actual patterns of growth at a particular time and place. Consequently, the WHO Child Growth Standards[5,6] are now being used worldwide to judge children's growth because they demonstrate how healthy children grow in an environment that allows them to achieve their full growth potential.

We aimed to extend the concepts promoted by WHO and the MGRS investigators into fetal and neonatal life. Our project is, therefore, based on the same prescriptive approach with international representation. The design, implementation and conduct of our project, and dissemination of the results, as well as their incorporation into clinical practice guidelines and healthcare policies, build on what has been achieved by the international MGRS team, offering to countries a conceptual continuity between the development and implementation of prenatal and postnatal growth standards.

Conceptual issues guiding the INTERGROWTH-21st Project

Figure 1 presents the conceptual framework behind the three primary objectives: the production of three prescriptive, international, multiethnic, growth standards that are described in the next section. There is considerable evidence to justify using such international standards in the field of perinatal medicine. For example, most, if not all, reference values for disease screening or diagnosis are not ethnicity specific. Epidemiological and clinical studies have also consistently demonstrated similar growth patterns in children from high socio-economic backgrounds across populations,[5,6] and growth patterns for both infants and children are more affected by health, socio-economic status and environmental conditions than genetic differences. Furthermore, human growth is a complex genetic trait involving many genes; it is unlikely to be affected by genetic variations responsible for characteristics such as skin colour. Finally, ethnicity-specific standards are impractical tools for use in most multiethnic populations served by healthcare systems, especially as admixture is increasing.

The concept of prescriptive growth requires that the populations used to construct the standards live in environments with no socio-economic constraints on growth, and receive up-to-date, evidence-based, medical care and appropriate nutrition. We have followed such principles as described below. The project design (described under 'Study characteristics' in Figure 1) will make the new standards consistent, conceptually, with those recently produced by WHO for infants and young children. We, have therefore, conducted prospective, population-based, multiethnic studies using standardised methodology in geographical areas where there is high quality maternal and neonatal care.

INTERGROWTH-21st objectives

The INTERGROWTH-21st Project has five objectives, involving separate studies. The first three are to produce, for practical clinical applications and to monitor population trends, three sets of international growth standards describing:

- Fetal growth from early pregnancy (Fetal Growth Longitudinal Study—FGLS).
- Postnatal growth of preterm infants (Preterm Postnatal Follow-up Study—PPFS).
- Birthweight, length and head circumference for gestational age (Newborn Cross-sectional Study—NCSS).

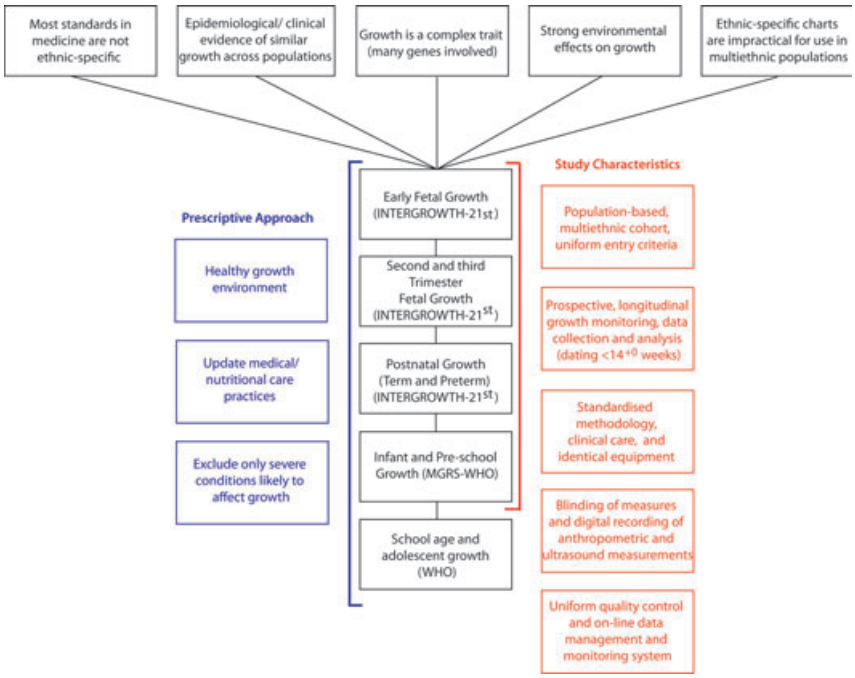

Figure 1. Conceptual framework for the construction of international growth standards based on the prescriptive approach.

These standards will be related to perinatal morbidity and mortality to identify levels of risk.

The remaining two objectives are:

(1) To investigate the determinants of preterm delivery and impaired fetal growth in this sample (Preterm and Impaired Fetal Growth Syndromes Study—PIFGSS).

(2) To develop a prediction model, based on multiple two-dimensional (2D) ultrasound measurements, for estimating gestational age during mid–late pregnancy for use in women without access to early/frequent antenatal care (Mid-late Pregnancy Gestational Age Prediction Study—MPGAPS).

Selection of INTERGROWTH-21st Project sites

The selection of the populations contributing participants to the project occurred on two levels, cluster and individual. The cluster level involved selecting a geographical area (e.g. city or part of a city with clear political or geographical limits) followed by the selection, within each area, of health institutions where women at low–medium risk for impaired fetal growth attend for antenatal/delivery care and infant follow-up. At the level of the individual, it involved selecting, within these populations and hospitals, women or newborns with specific characteristics required for the project's different components.

A list of nine potential study sites was created with the intention of achieving geographical distribution across continents. Previous participation in multicentre observational studies, including MGRS,[7–9] and maternal/perinatal health randomised controlled trials was considered beneficial. For logistical reasons, one site was unable to participate leaving a total of eight, two of which had participated in MGRS (Figure 2).

Within each geographical area, we identified all institutions that could provide the required participants. Locally adapted definitions were used for socio-economic characteristics associated with unconstrained growth in these populations, including measures of household income, housing tenure, education, occupation and employment status using locally selected cut-off points. These variables have been recently identified as mediating factors in the relationship between birthweight and ethnicity.[10]

To select institutions in developing countries, we first conducted a census of all hospitals where deliveries take place in each geographical area and identified those classified locally as 'private' or 'corporation' hospitals and/or those serving the middle to upper socio-economic population. However, we have included all institutions in a region if most deliveries take place in a few central institutions (e.g. Brazil and China).

The institutions selected in each geographical area delivered >80% of the eligible women in the target population. We concentrated on institutions with >1000 deliveries per

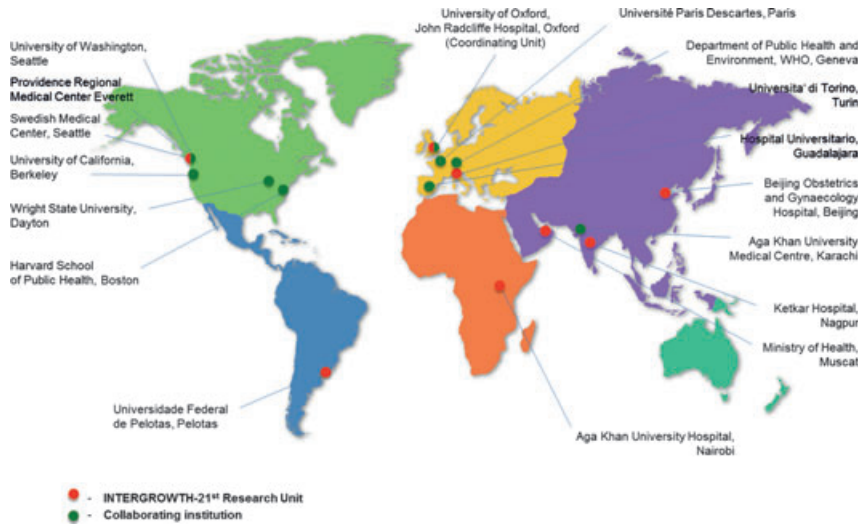

Figure 2. Collaborating institutions participating in the INTERGROWTH-21st Project.

year, as it was important to involve as few hospitals as possible for logistical reasons (Figure 3).

The institutions in developing countries serve low-risk populations as defined by: (1) low birthweight rate <10% and mean birthweight >3100 g; (2) located at an altitude <1600 m; (3) perinatal mortality <20 per 1000 live births; (4) mothers attending antenatal care in these institutions should plan to deliver there or in a similar hospital located in the same geographical area; (5) >75% of mothers have attained an educational level greater than the locally defined cut-off point, and (6) absence of known non-microbiological contamination such as pollution, domestic smoke, radiation or any other toxic substances, evaluated at the cluster level using a data collection form specifically developed for the project.

We present below the description of the different studies and the research strategies comprising the INTERGROWTH-21st Project.

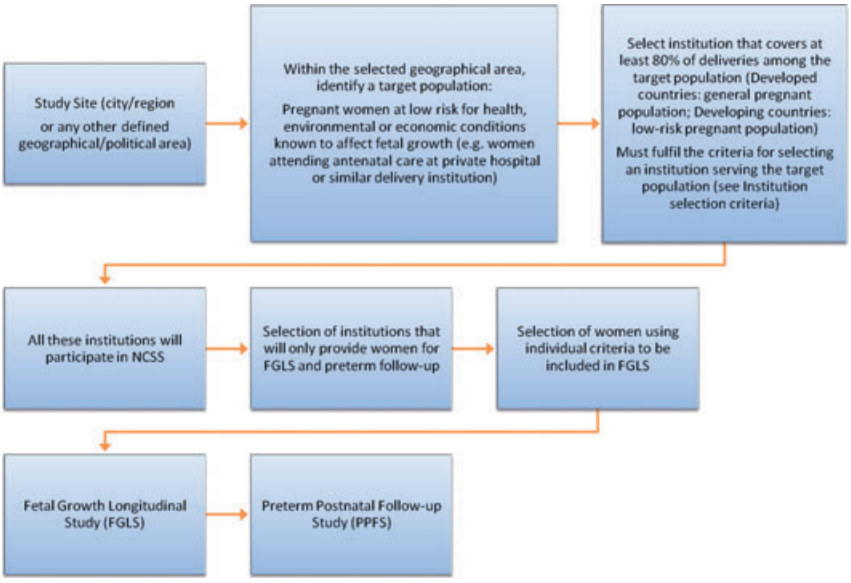

Figure 3. Selection of populations for the INTERGROWTH-21st Project.

1. The Fetal Growth Longitudinal Study (FGLS)

This study aims to develop: (1) new, international, fetal growth standards using 2D ultrasound to measure the most commonly acquired dimensions of fetal size, and (2) a new, international, symphyseal–fundal height standard.

FGLS individual entry criteria

The methods used to select either a 'very healthy' or a 'healthy' study population, and the risk factors associated with intrauterine growth restriction (IUGR) to be excluded are debatable. The most important considerations for selecting a 'healthy' population with no obvious risk factors for IUGR or over-growth at the first antenatal care visit early in pregnancy are: (1) achieving a balance between strict criteria for risk and external validity of the study population, and (2) the logistics of screening for factors that are not part of routine care or for which consensus is lacking about their effect on fetal growth.

There is an extensive literature on risk factors for preterm delivery and IUGR, particularly in low-income and middle-income countries, as well as data from our own large-scale studies that have systematically addressed the issue.[11–19] Hence, in defining the FGLS eligibility criteria, we could have excluded women with every possible risk factor for poor pregnancy outcomes. However, we believed that it was preferable, in the initial screening process, to select specific factors commonly used to identify women who would benefit from low-risk, routine antenatal care[1] (Box 1).

All women $<14^{+0}$ weeks of gestation by menstrual dates, attending their first antenatal care visit in the selected institutions were screened at study entry using the above criteria. This resulted in a study population likely to need only routine antenatal care, i.e. a group of apparently healthy women who could follow basic antenatal care models. Some variables have clear thresholds (e.g. urine or previous fetal death); for others, the thresholds are less clear-cut (e.g. maternal height or education)[20] and, inevitably, we had to make some arbitrary decisions.

Estimating gestational age at study entry

Clearly, establishing a very precise determination of gestational age is vitally important for constructing growth standards (in addition to clinical management) and we were prepared to screen large numbers of women, if necessary, to obtain the ideal study population.

There are three ways to estimate gestational age early in pregnancy: (1) the first day of the last menstrual period (LMP) alone; (2) early ($<14^{+0}$ weeks of gestation) ultrasound alone assuming that the early growth of the fetus has been normal, or (3) LMP and ultrasound combined. The implications of these different methods on research findings have recently been discussed.[21] Dating by LMP and ultrasound is clearly an assessment of different parameters and both methods have their limitations. Although it has been suggested that ultrasound before 14^{+0} weeks of gestation may be better by an average of 2–3 days in predicting the date of delivery, in clinical practice, both methods are often used in combination and it is now generally recommended that both should be retained for fetal growth

Box 1. Women must have the following characteristics at booking (<14 weeks of gestation)

a) Aged ≥18 and <35 years.

b) Body mass index ≥18.5 and <30 kg/m^2.

c) Height ≥153 cm.

d) Singleton pregnancy.

e) A known last menstrual period with regular cycles (defined as 28 ± 4 days) without hormonal contraceptive use, or breastfeeding in the 2 months before pregnancy.

f) Natural conception.

g) No relevant past medical history, with no need for long-term medication (including fertility treatment and over-the-counter medicines, but excluding routine iron, folate, calcium, iodine or multivitamin supplements).

h) No evidence of socio-economic constraints likely to impede fetal growth identified using local definitions of social risk.

i) No use of tobacco or recreational drugs such as cannabis in the 3 months before or after becoming pregnant.

j) No heavy alcohol use (defined as >5 units (50 ml pure alcohol) per week) since becoming pregnant.

k) No more than one miscarriage in the two previous consecutive pregnancies.

l) No previous baby delivered preterm ($<37^{+0}$ weeks of gestation) or with a birthweight <2500 g or >4500 g.

m) No previous neonatal or fetal death, previous baby with any congenital malformations, and no evidence in present pregnancy of congenital disease or fetal anomaly.

n) No previous pregnancy affected by pre-eclampsia/eclampsia, HELLP syndrome or a related pregnancy-associated condition.

o) No clinically significant atypical red cell alloantibodies.

p) Negative urinalysis.

q) Systolic blood pressure <140 mmHg and diastolic blood pressure <90 mmHg.

r) No diagnosis or treatment for anaemia during this pregnancy (haemoglobin levels will be monitored throughout pregnancy).

s) No clinical evidence of any other sexually transmitted diseases, including syphilis and clinical trichomoniasis.

t) Not in an occupation with risk of exposure to chemicals or toxic substances, or very physically demanding activity to be evaluated by local standards. Also women should not be conducting vigorous or contact sports, such as scuba diving or similar activities.

monitoring. Based on the 2004 Birth Cohort from the Brazilian Study Centre, we anticipated that ~10% of women in these selected subpopulations would have an unreliable LMP, although this figure may be as high as 30% in the general population.

Taking all factors into consideration, we decided not to base gestational age solely on LMP or early ultrasound; rather, we chose a combined, two-stage process to determine gestational age. In women with a reliable LMP and regular periods who were 9^{+0} to 13^{+6} weeks pregnant, their gestational age was confirmed with a standardised, ultrasound crown–rump length (CRL) measurement, using an internationally recognised chart.[22] If the difference between the CRL and LMP estimates was ≤7 days, the LMP was considered valid and taken as the true biological date. Those women with differences >7 days were not considered for the study because either the LMP estimation may not be reliable or the discrepancy indicated an early fetal growth alteration.

Hence, we did not include participants if fetal size on ultrasound at the first scan was discrepant from the LMP. If ultrasound alone was used to estimate gestational age and the measurement was erroneous, that error would be incorporated into the growth chart. This is a circular argument that we tried to avoid. Furthermore, the early ultrasound estimation of gestational age has the limitation that all fetuses with a given CRL value will have the same gestational age, which is clearly not possible, i.e. using an ultrasound-based gestational age, such as CRL alone, excludes any biological variability, which is a major limitation of previous data evaluating fetal growth.[23,24]

Nutritional adequacy during pregnancy

The study population was expected to have an adequate nutritional status, based on the same conceptual approach used in MGRS. In that study, it was assumed from the outset that children recruited from affluent populations consumed adequate complementary foods. This proved to be the case when their complementary diets were analysed.[25] In the present study, we adopted the same concept for women but, in addition to selecting those with adequate nutritional status before pregnancy, we developed general nutritional guidelines for pregnant/lactating women, suitable for local use, based on the best available evidence, for promotion among the participating mothers and care providers.

Routine nutritional supplements, e.g. protein or energy, were not given because: (1) they are not components of the recommended antenatal care package in these populations,[1] and (2) we recruited women with adequate nutritional status. Iron-folate supplementation was prescribed if necessary for anaemia during/after pregnancy, but routinely given only if such a policy was in place in the institution. A similar position was taken with calcium supplementation for the prevention of pre-eclampsia and preterm delivery.

Women were asked if they were taking nutritional supplements; this information was recorded at each antenatal care visit. It is not practical in a study of this size to measure adherence in any other way or to obtain individual intakes, e.g. 24-hour recall, considering the poor reliability of such instruments for individual assessment. Finally, as an objective outcome of the nutritional status during pregnancy, we monitored weight gain at each visit using standardised methodology.

Pregnancy follow-up

Women in FGLS received standardised antenatal care based on the recommended WHO new antenatal care package (modified or upgraded according to local practices).[1] All participants were followed throughout pregnancy from their first clinic visit, irrespective of the pregnancy outcome. As a general principle, the number of exclusions from the analysis for the creation of the final fetal growth standards will be as small as possible. They will be confined to fetuses with congenital abnormalities (based on a final evaluation at birth); multiple pregnancies that were not identified at recruitment; mothers diagnosed with catastrophic fetal death or very severe medical conditions not evident at recruitment (e.g. cancer, HIV); those with severe pregnancy-related conditions requiring hospital admission (e.g. eclampsia or severe pre-eclampsia), and those found later in pregnancy to fulfil one of the exclusion criteria (e.g. women who started to smoke during pregnancy). Hospital admission per se is not a reason for exclusion: women admitted simply for 'observation' still contribute data to the fetal growth standards unless they developed one of the conditions listed above. All sites are malaria-free eco-zones. Nevertheless, we adhered to local protocols and excluded from incorporation in the fetal growth standards any woman with evidence of malaria infection during the pregnancy. The final definitions of these conditions were approved by the Project Steering Committee before analysing any data.

To date, FGLS has had fewer follow-up problems than many of our previous randomised controlled trials. This was expected because we are studying well-educated women who are enthusiastic about the aims of the study. Nevertheless, to ensure that the loss to follow-up remains as low as possible, we maintain very close contact with participants, reminding them about imminent visits and conducting home visits if necessary.

FGLS 2D ultrasound measurements

Detailed descriptions of the ultrasound protocol and the extensive quality control measures being employed in FGLS are presented elsewhere.[26,27] In preparing the protocol, it

became clear that performing more than six ultrasound examinations after the dating scan would present undesirable logistical problems and possibly cause inconvenience to mothers. It was also evident that a minimum growth change has to occur between visits to be reliably measured by ultrasound, considering the errors resulting from the equipment and observers. Hence, even if growth velocity by unit of time (e.g. 1 week) for some measurements is high during certain gestational periods, the actual change may not be reliably measured.

We considered alternative spacing between measurements, such as longer intervals early (e.g. 8 weeks), and shorter ones (e.g. 4 weeks) later, in pregnancy. However, there is evidence that measures such as biparietal diameter and femur length continue to increase almost linearly until 34 weeks. The need to coordinate multiple antenatal visits, at different time intervals for a large number of women, was an argument against adopting variable timings. Finally, as women were recruited more or less spread across the 9^{+0} to 13^{+6} weeks of gestation range, we expected a distribution of visits throughout pregnancy in most centres. Therefore, after the first scan (9^{+0} to 13^{+6} weeks), we performed scans targeted at 5-weekly (±1 week) intervals, i.e. 14–18, 19–23, 24–28, 29–33, 34–38 and 39–42 weeks. So, a woman could have scans at 10, 15, 20, 25, 30, 35 and 40 weeks or 11, 16, 21, 26, 31, 36 and 41 weeks depending upon the gestational age at recruitment and duration of pregnancy. This has the advantage of providing greater coverage of the whole pregnancy and easier scheduling of scans, especially as we allowed a leeway of up to a week either side of each planned scan visit. A standard ultrasound examination includes five measurements, in addition to biparietal diameter and femur length, at each visit from 14 weeks onwards: occipito-frontal diameter; head circumference; transverse abdominal diameter; anterior–posterior abdominal diameter, and abdominal circumference.

All clinical data in FGLS are obtained as part of routine practice and available to care providers at all times. Our policy with regard to the ultrasound findings has been as follows: (1) the gestational age estimate is incorporated into the medical records; (2) an ultrasound examination for structural, congenital malformations is performed at the time of the third scan before 24^{+0} weeks of gestation, if this conforms with local practice, and the results are incorporated into the medical records, and (3) as far as 2D ultrasound measurements are concerned, after the blinded values have been submitted electronically to the dataset, the mean measurement is made available for clinical use.

Finally, as there are concerns about the misuse of ultrasound for sex selection, especially if the practice is illegal, we were extremely vigilant in regions where this practice is known to occur and we continuously monitored sex ratios in the sample.

FGLS symphyseal–fundal height measures

We also aim to produce a new, international, symphyseal–fundal height standard from FGLS to update the one we first produced 30 years ago based on a single site sample.[28] The present study is the first of its type ever conducted and the tool will aid antenatal care at the primary-care level. Symphyseal–fundal height measurements are being taken at the same time as the ultrasound scans, using similarly rigorous protocol-driven methods, i.e. blinded, duplicate measurements, followed by the same clinician (see www.intergrowth21.org.uk). They will be used to produce a standardised, validated, multiethnic chart, in a selected healthy population, to replace those used in primary healthcare units and hospitals around the world.

To summarise, Figure 4 presents the patient flow in FGLS, highlighting the study forms being used. This flow chart is continually populated by the actual numbers enrolled using our centrally coordinated, online, data management system.

2. The Preterm Postnatal Follow-up Study (PPFS)

The aim of this study is to develop postnatal growth charts for preterm newborns based on the conceptual principles presented recently by our group.[29,30]

PPFS individual entry criteria and follow-up

All preterm newborns ($\geq 26^{+0}$ but $<37^{+0}$ weeks of gestation) from the FGLS cohort are being followed for 8 months after delivery to evaluate postnatal growth. Studying all preterm infants born to mothers selected using the prescriptive approach (i.e. pregnancies at low risk of fetal growth alterations) allows evaluation of the rate of preterm delivery to be expected from such a low-risk population.

Our strategy makes it possible to study mostly preterm newborns born after 32^{+0} weeks of gestation until they reach 'term', with at least 5 months of 'true' (after term) postnatal life. Although we are following all preterm newborns, we are not including those born $\leq 30^{+0}$ weeks of gestation in the growth standards because of the associated severe morbidity and mortality, and consequent need therefore for intensive care. Indeed, we are observing that in this selected pregnant population even the rate of preterm birth $<34^{+0}$ weeks of gestation is <1%. We are, however, following these subgroups of very preterm infants for exploratory analysis of their growth pattern across these populations. In addition, using the same protocol, maternal entry criteria and standardisation procedures, we have studied (for separate analysis) all preterm births in the study institutions within the 30^{+0} to 34^{+0}-weeks of gestation range, even if the mothers were not part of FGLS.

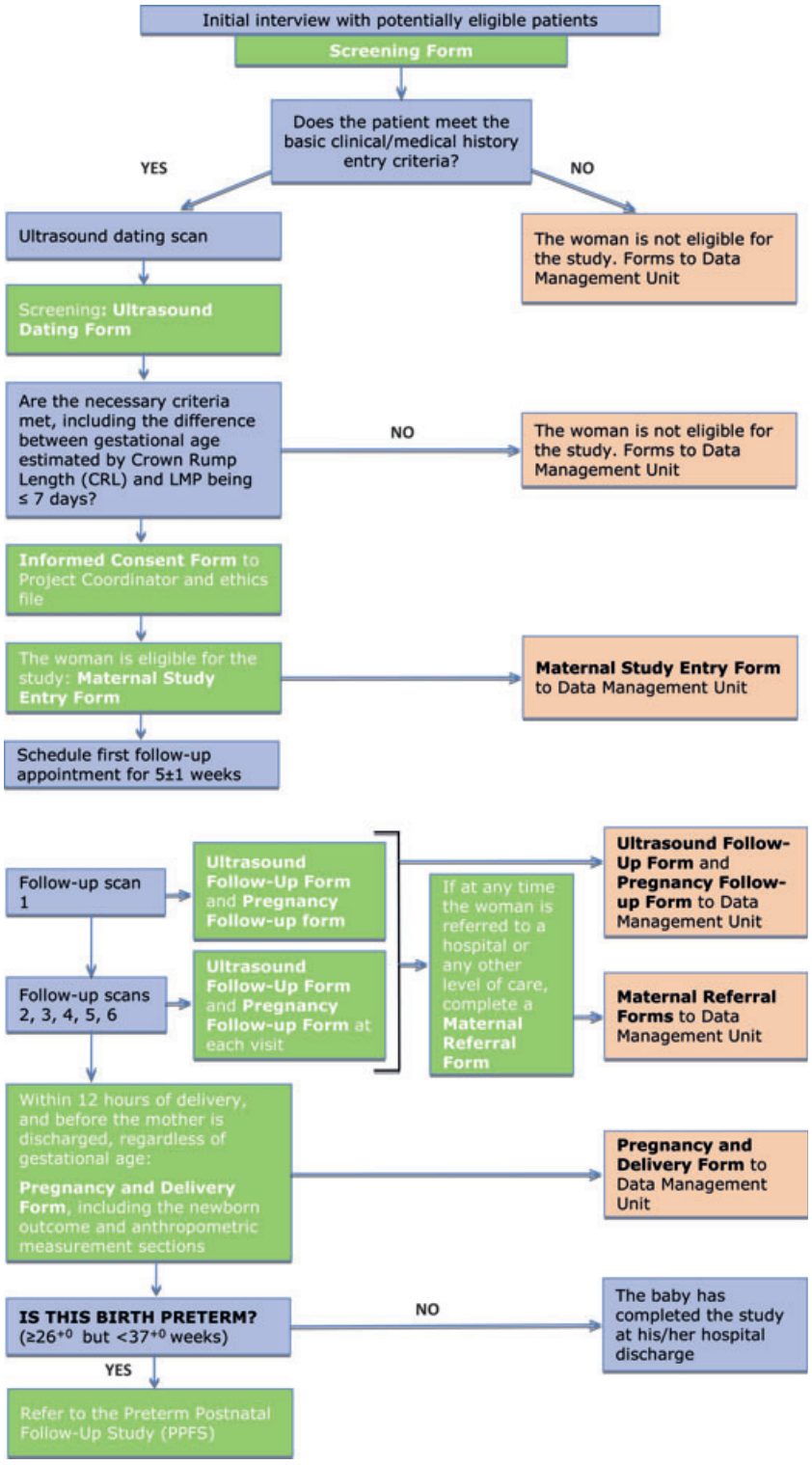

Figure 4. The patient flow during the Fetal Growth Longitudinal Study.

A fixed follow-up period (rather than a 'postnatal age limit') was chosen to simplify running the study and reduce loss to follow-up (Figure 5). Nevertheless, an analysis based on the time from conception (corrected age) will be performed to compare preterm babies with their in utero counterparts conceived around the same

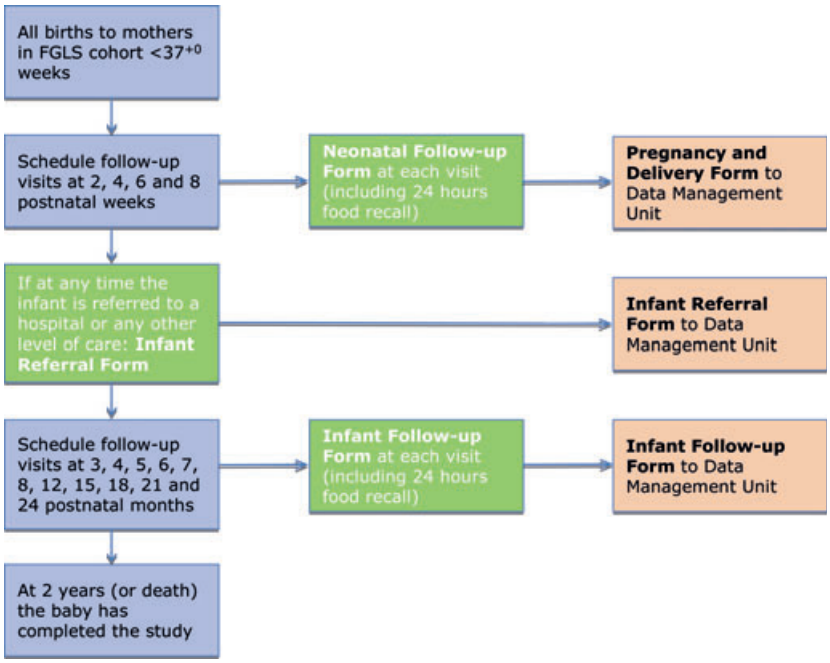

Figure 5. The patient flow during the preterm postnatal follow-up study.

time. This relatively short follow-up period should reduce: (1) inconvenience to the mothers, (2) the need for home visits, and (3) the study's cost and complexity. Finally, although we wish to produce new growth charts for the first 6 months of life, we have extended the follow-up period to 8 months for these infants to avoid the so-called 'right-edge' effect in the construction of the growth standards.[31]

From the total preterm population in FGLS, we shall select newborns to be included in the PPFS standards who have met the criteria of 'healthy or stable' preterm decided a priori. We recognise that this is a difficult definition and have discussed its concept and implications in detail in a recent publication.[30] In short, preterm infants contributing to the new standards should be free of congenital malformations and major neonatal conditions associated with impaired postnatal growth. They should have received standardised, evidence-based care and been breastfed exclusively/predominantly for as long as possible based on current recommendations. A detailed description of these criteria and the definitions to be used are presented in Box 2 and Table 1. This strategy should provide a population that is conceptually as close as possible (adapted to the level of maturation of these preterm infants) to the prescriptive approach used to construct the MGRS standards.

The INTERGROWTH-21st Neonatal Group, in collaboration with international advisors, have standardised clinical management across sites and described the clinical conditions that are exclusion criteria for the growth standards to facilitate uniform diagnosis across sites. This information and our operations manual (see www.intergrowth21.org.uk), with standardised data collection forms,

Box 2: Characteristics of prescriptive standards for monitoring preterm postnatal growth

Characteristics

Multiethnic, population-based, prospective data collected under recent medical care.

Healthy, well-nourished maternal population.

Early evaluation of gestational age confirmed by ultrasound examination before 14 weeks.

Inclusion only of preterm deliveries (not using low birthweight as a proxy).

Prospective ultrasound measures of fetal growth to exclude fetuses with evidence of impaired fetal growth.

Preterm infants included only if they do not have major neonatal complications, neonatal surgery, congenital malformations or death in the complete follow-up period.

Standardisation of feeding practices and newborn care among study centres.

Standardisation of anthropometric measurements including use of the same equipment and techniques.

High frequency of anthropometric measurements during periods of fast growth (e.g. every 2 weeks during the first *2* months).

Follow-up period during infancy to allow interface with WHO child growth standards.

Adequate sample size for each range of gestational ages to allow presentation by *Z* scores and centiles.

Table 1. Criteria for evaluating preterm newborns according to complications associated with postnatal growth

Organ or system	Physiological immaturity, leading to increased risk of	Conditions associated with a lack of organ maturation	Pathological conditions
Brain	IVH, Papile's grades I–IV PVL, de Vries' grades I–IV Neurological impairments	IVH, grades I–II PVL, grades I–II	IVH, grades III–IV PVL, grades III–IV Cerebral palsy Neurological impairments Hydrocephalus requiring ventriculoperitoneal shunt
Lung	Respiratory distress	Respiratory distress syndrome, not leading to bronchopulmonary dysplasia	Respiratory distress syndrome, leading to bronchopulmonary dysplasia or chronic lung disease
	Apnoeas	Apnoea of prematurity Transient tachypnoea of the newborn	
Heart	Patent ductus arteriosus	Not requiring surgery	Requiring surgery
Gastrointestinal	NEC, Bell's stages I–III Parenteral feeding Gastro-oesophageal reflux	NEC, stage I Parenteral feeding for ≤7 days Gastro-oesophageal reflux	NEC, stage II–III Exclusive parenteral feeding >7 days Complicated gastro-oesophageal reflux/short bowel syndrome Focal intestinal perforation
Liver	Hyperbilirubinaemia	Requiring or not requiring blood exchange transfusion	Kernicterus
Kidneys	Renal failure	Transitory renal failure not requiring dialysis	Any renal failure requiring dialysis
		Low urinary output and/or creatinine >1.0 for <3 months	Low urinary output and increased creatinine levels for >3 months
Eyes	Retinopathy	Retinopathy stages I and II	Retinopathy stages, III, IV, V Complete blindness
Immune system	Infections	No sepsis	Sepsis
Other	Temperature instability	Low temperature Continuous stoppage of enteral feeding for ≤3 days	Stoppage of enteral feeding >3 days

IVH, intraventricular haemorrhage; NEC, necrotising enterocolitis; PVL, periventricular leukomalacia.

have harmonised the implementation of the study across the participating neonatal units. Adherence to these protocols is being monitored by the Project Coordinating Unit and the Neonatal Group (in collaboration with the international advisors).

Clearly, for some subgroups of very preterm infants, as discussed above, this study will allow only exploratory analyses; however, it is still worthwhile because of the unique opportunity to study fetal and postnatal preterm growth longitudinally. Some babies, especially those born before 34^{+0} weeks of gestation, are being managed in a neonatal intensive care unit (NICU) for several days. Measurements of these babies are being taken as described in the study protocol, but in accordance with the clinical status of the infant and the unit's protocols. To ensure that 'clinically stable' babies are comparable across different NICUs, the Neonatal Team is conducting regular standardisation and monitoring activities based on the study manual.[32] Ultimately, in such an acute clinical care setting, final judgements are made by the attending staff; realistically, this is how the growth standards will be used in practice anyway.

The postnatal anthropometric measurements are weight, length and head circumference. Abdominal circumference is not included because it is not used in routine neonatal practice and respiratory movements in these tiny newborns make its measurement unreliable. The three measurements (plus a standard clinical evaluation and records of morbidity and food intake) are being taken every 2 weeks during the first 8 weeks, and then every 4 weeks until 8 postnatal months, using essentially the same methodology employed in MGRS.[6] The only logistical difference is that, in the present study, measures, interviews and clinical evaluations are conducted at a special follow-up clinic in the corresponding hospital, provided that the clinical condition of the infant and routine local care permit this. Routine home visits (as conducted in MGRS) are not taking place, except

for those mothers who do not comply with the protocol's scheduled visits or those newborns that, because of local customs, may only be evaluated at home during the first month of life. The preterm babies have a maximum of 11 follow-up visits over 8 months.

3. The Newborn Cross-Sectional Study (NCSS)

The aims of NCSS are to: (1) produce birthweight, length and head circumference for gestational age standards describing fetal size at birth, and (2) provide data for epidemiological studies of the different phenotypes of the impaired fetal growth and preterm delivery syndromes. The relationship of size at birth with neonatal morbidity and mortality is also being examined. The complete NCSS population includes two subpopulations: one selected using the FGLS entry criteria (the same questions were included in the data collection forms for both studies) to construct the prescriptive standards of size at birth. The second, composed of the remaining newborns from higher-risk pregnancies, will contribute to the epidemiological analyses and the evaluation of the new standards as related to neonatal morbidity and mortality. Hence, NCSS is a descriptive, population-based study that aims to include all babies born within the INTERGROWTH-21st geographical areas during a fixed period of ~12 months (the actual period varies across sites to enable the target of 7000 deliveries per site to be attained). Consequently, NCSS includes an 'FGLS-like' population, as well as the infants born to mothers enrolled in FGLS itself. Having data on both groups will allow us to understand the similarities in terms of maternal and fetal characteristics, which will provide external validity for the fetal growth standards.

The overwhelming majority of babies in NCSS will have had their gestational age confirmed by ultrasound at their first antenatal visit and close to 80% of the mothers will have received care early ($<24^{+0}$ weeks of gestation) in pregnancy. This has been achieved because all the hospitals in the INTERGROWTH-21st Project adopted a policy, during the preparatory phase of the study, of confirming gestational age with a dating ultrasound examination at the first visit. The strategy of including the complete newborn population allows us not only to study a large number of preterm and growth-restricted infants, as well as their subgroups, under a wide range of conditions affecting these outcomes, but also to understand the epidemiological distribution of these groups and their associated risk factors in populations. We are also able to compare the total underlying populations across sites in terms of sociodemographics and pregnancy outcomes.

All newborns during the study period, including those admitted to NICU, special care or referred to another level of care, are assessed on a daily basis until hospital discharge to document severe morbidities and detect neonatal deaths. We have also made strenuous efforts to coordinate and promote evidence-based care and promoted breastfeeding practices for all neonates using specific protocols agreed among the institutions' lead neonatologists.[2] We recognise that differences in practice will persist despite our best efforts. However, we believe this is unavoidable in a pragmatic study such as this, which is trying to reflect what happens in routine clinical practice. Furthermore, we are similarly making strenuous efforts to standardise the main protocols for feeding practices in each NICU.[32] Implementation of these protocols is monitored during the routine site visits by members of the Project Coordinating Unit, and the Neonatal and Anthropometric Teams.

Birth size for gestational age standards, developed using the subpopulation of NCSS selected using FGLS criteria (i.e. the 'FGLS-like' population), will be related to indicators of perinatal outcome to establish risk levels associated with different growth patterns. The 'ideal' outcome is perinatal mortality, but its anticipated infrequent occurrence (<1.5%) in this low–medium risk population makes it unrealistic to have a sample large enough for the necessary number of events across the gestational age distribution.

We have therefore decided to use an unweighted composite outcome, including at least one of the following conditions: stillbirth, neonatal death occurring up to hospital discharge of the newborn, and newborn stay in NICU for ≥7 days. This index requires no standardisation of clinical diagnoses across hospitals and is accepted as a marker in large, international, population-based studies of newborns that died or were severely ill. Other groups have successfully used these composite indices.[33–35] It could be argued that intrapartum stillbirth may not be completely related to fetal growth and should not be included in this index. We believe this is a valid point but as it will not be possible to separate those intrapartum deaths that are related to IUGR from those that are unrelated, we are keeping it in the index.

The index is a good proxy for adverse perinatal outcomes, including severe morbidity, and obviates the need for uniform neonatal diagnoses across NICUs. It has the disadvantage that it only includes events up until hospital discharge. Hence, it risks excluding from the total number of early neonatal deaths some cases of healthy, mostly term babies delivered vaginally who die after early hospital discharge, or those who develop severe complications ≤7 days postpartum and do not return to the same hospital for care. However, missing these isolated cases was considered logistically preferable to establishing a follow-up regimen requiring universal home visits up to 28 days postpartum. Nevertheless, in some sites with registration systems, it will be possible to explore how many later neonatal deaths were missed.

In summary, for NCSS, we are including all newborns delivered at the institutions that cover >80% of all deliver-

ies in the geographical areas selected for the project. These deliveries occur over a 12 month study period or up to the point when the target sample of 7000 newborns is reached. The FGLS population includes women from these underlining populations who have agreed to participate and meet the individual selection criteria until the target sample number is reached. For PPFS, we are enrolling all preterm babies from the FGLS cohort.

4. The Preterm and Impaired Fetal Growth Syndromes Study (PIFGSS)

The complex interactions between risk factors, clinical presentations and underlying biological processes for these syndromes in relation to adverse perinatal outcomes are poorly understood. Moreover, the failure to understand the heterogeneity of preterm birth and IUGR is now accepted as a major limitation in understanding and so preventing these syndromes.[36–38] The objective of this study is, therefore, to explore and validate phenotypic subgroups of the preterm and IUGR syndromes in the population-based NCSS worldwide. Hence, we will be able to estimate both relative and attributable risks for aetiological and risk factors.

To identify risk factors, data collected during routine and specialist antenatal care from all women delivering in the institutions over a fixed period are being included. A summary antenatal and delivery form was introduced into these institutions, which conforms with: (1) the new WHO model of antenatal care for basic routine care,[1] and (2) local protocols for special cases, standardised by us from previous trials in pre-eclampsia, hypertension, urinary tract conditions, and intrapartum and postpartum care. An important conceptual issue is that we do not aim to detect any new, unexplored risk factors. Rather, we plan to determine how known risk factors and clinical conditions, which are routinely recorded during standard antenatal care, are distributed or cluster in the main, preterm and IUGR, phenotypic subgroups across these populations.

Specifically, we plan to investigate the determinants of preterm delivery and small for gestational age newborns from clinical and pregnancy-related, as well as routine laboratory, demographic and socio-economic, variables obtained from all women attending the study sites without exclusion (as opposed to FGLS, which aims to produce fetal growth standards from a sample of selected, healthy women). It would, of course, be very interesting to collect more detailed information about other variables such as placental pathology and cervical length, or test biomarkers, e.g. of inflammation or infection. However, the question, as always, is when to stop adding more variables to an already complex population-based study, particularly as some of these screening tests are not clinically relevant at present. Finally, as we have the complete birth population, we shall explore the gestational age cut-off points for the definition of preterm delivery to validate the standard (clearly arbitrary) $<37^{+0}$ weeks of gestation cut-off point.

We will apply several statistical strategies to understand the different components of the ~3500 preterm newborns ($<37^{+0}$ weeks of gestation) or close to 5000 IUGR newborns (<10th centile) in the total NCSS population. Initially, we will consider the classic preterm subgroups based on the mode of delivery, e.g. induced versus spontaneous deliveries; premature rupture of membranes; pre-eclampsia-related versus smoking-related or unexplained IUGR, as well other pathological and physiological conditions.[13] However, the most innovative part of this analysis will be the investigation of new phenotypic classifications based on associations among aetiological factors, clinical events, initiation of parturition and pathway to delivery. Cluster analysis and similar data reduction strategies to evaluate how sets of variables are grouped to determine subgroups of preterm birth will be explored. Once phenotypic subgroups are defined, their associations with risk factors and perinatal outcomes will also be examined. Lastly, we intend to explore several factors that might explain variability in fetal growth across populations.[39,40]

5. The Mid–late Pregnancy Gestational Age Prediction Study (MPGAPS)

We aim to develop a prediction model, using multiple, prospectively collected, 2D ultrasound measurements, to estimate gestational age at a single visit during mid–late pregnancy ($>24^{+0}$ weeks of gestation). Current ultrasound-based gestational age estimation in infrequent attendees or women with limited access to care is usually determined by a single measurement, e.g. biparietal diameter, but the estimates have large errors, wide confidence intervals, and assume that fetal growth has been similar to that of the population from which the equations were derived. We are planning to develop equations using several ultrasound measures taken at a single visit which, if achievable, would be a major contribution to the care of high-risk women attending hospitals only once or twice late in pregnancy. We are also planning to produce new international equations to estimate gestational age before 14^{+0} weeks of gestation from CRL measures to replace the large number of equations that are presently used.

General sample size considerations

FGLS

Sample size calculations for growth standards have not been extensively developed and therefore are supported by a limited literature. We reviewed several strategies and estimated that the target sample of 500 pregnant women per population group should be adequate, after excluding

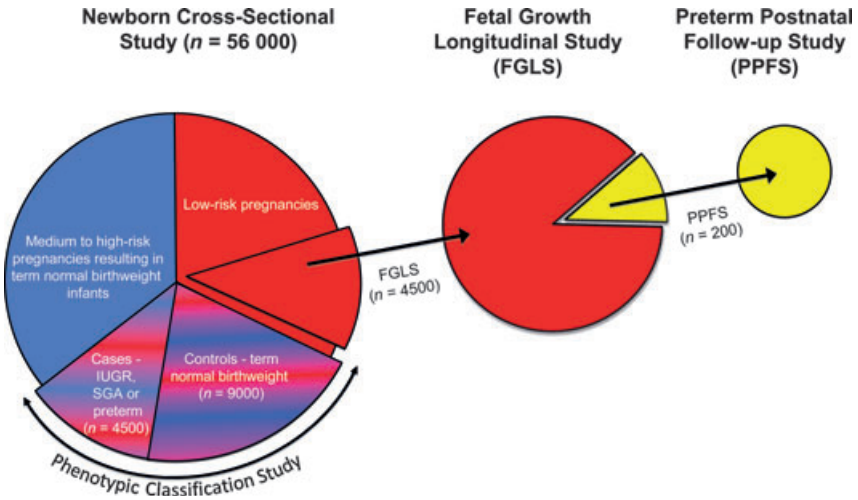

Figure 6. A population flow diagram for the INTERGROWTH-21st Project.

women with complicated pregnancies and those lost to follow-up. This sample size is larger than most previous longitudinal studies, even if each study population is considered separately, and adequate to explore site-specific differences. We estimated that fewer than 5% of women would be lost to follow-up (based on an average figure of ~3% in our previous large trials). We also adjusted for the fact that ~3% of women will be excluded from the study population to be used in the preparation of the fetal growth standards because of developing problems severe enough to affect fetal growth.

PPFS

A cohort of ~250 preterm babies will be recruited from FGLS for PPFS. It is expected that the gestational age distribution will provide a total of 200 babies $>34^{+0}$ weeks of gestation (those between 26^{+0} and 30^{+0} weeks of gestation will be followed, but will not be part of the standards). Only the subgroup of babies free of major clinical complications, as previously described, will be used to create the standards. We recognise that the sample size calculations of FGLS are influenced by logistical issues and the need to provide preterm newborns to be part of PPFS rather than statistical calculations alone. However, it is still a large sample by preterm study standards and we shall have very detailed follow-up data, increasing the power of the sample for creating charts. We consider that having a full set of fetal and newborn growth patterns from a cohort of preterm newborns is important, even without the power to explore gestational age subgroups or early postnatal morbidity.

NCSS

We are aiming for a total of ~56 000 newborns with very detailed information, which will provide a sample of ~40% eligible babies for the newborn standards, using the same criteria as for the FGLS population (the 'FGLS-like subpopulation'). We expect to include overall ~2000 newborns with a positive, severe perinatal morbidity and mortality index, our established primary outcome.

Summary of the study populations and their flow

Figure 6 shows the population-based nature of the project and the relationships between subpopulations. We believe this comprehensive, yet integrated, series of study populations makes the INTERGROWTH-21st Project unique. The populations of mostly low- to medium-risk women were selected using institutional level criteria; these ~56 000 women are included in NCSS. From this population, using criteria at the level of the individual, we have identified women at low risk of IUGR ('FGLS-like population'), from which women in FGLS itself are recruited. This cohort is being followed from $<14^{+0}$ weeks of gestation to delivery to produce the fetal growth standards. Those babies born prematurely ($\geq 26^{+0}$ but $<37^{+0}$ weeks of gestation) to women in FGLS are being followed for 8 months after birth in PPFS. Lastly, data from NCSS and FGLS are contributing to two other separate studies, PIFGSS and MPGAPS, respectively.

Data collection instruments

All documentation and forms used in the INTERGROWTH-21st studies were prepared by the Project Coordinating Unit, tested at the local level and introduced into the specially developed electronic data management system. All forms are integrated and linked to reduce duplication in the data collection process and facilitate data quality control mechanisms. We are only collecting data that are

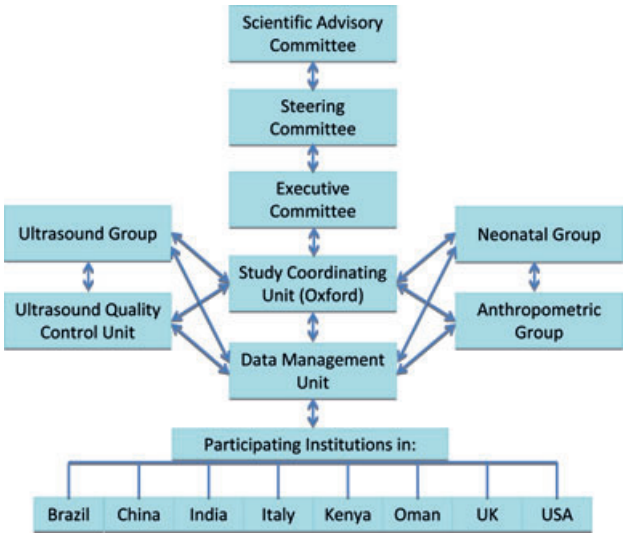

Figure 7. The structure and management of the INTERGROWTH-21st Project.

relevant to the specific aims of the studies, i.e. we have tried hard to eliminate unnecessary or duplicate data collection, as collecting too much inevitably reduces data quality in large field studies. One example is our decision not to collect detailed individual nutritional intake from this well-nourished, healthy population. A more detailed description of the data collection system, including the actual forms, and the quality control strategies is presented in another paper of this supplement.[41]

Assessing similarities among the populations of fetuses and newborns whose mothers' health, nutritional and care needs are met

To assess any differences in the populations of fetuses and newborns (whose mothers have optimal health, nutritional status and obstetric care), we will follow the basic principles adopted in MGRS.[5] As in that study, we consider that the main question to ask is whether or not the variability across populations for any given growth parameter is larger than the variability of the same parameter within populations. However, this is not the same as saying that the populations do or do not have identical growth.

To compare growth across populations, the recommended measures in the nutritional field are fat-free mass indicators, e.g. infant length, as they are: (1) resistant to skewing in response to 'excessive nutrition'; (2) normally distributed (unlike fat-related indicators); (3) more precise than fat-related indicators; and (4) they were used to construct the MGRS standards, which provides a conceptual link between the projects.

We consider a number of fat-free indicators in the fetus to be the counterparts to length or height in infants and children. These are: (1) CRL, which is generally accepted as the best ultrasound marker for fetal size at $<14^{+0}$ weeks of gestation, and (2) head circumference in later pregnancy, as it is widely used in clinical practice and more resistant to fetal environmental insults (the so-called 'brain-sparing effect') than abdominal circumference. Importantly, head circumference is also the only measure available from early gestation ($>14^{+0}$ weeks of gestation) to childhood. Femur length, the other potential linear measure, is routinely recorded in pregnancy but it represents only a small part of newborn total length and does not have a newborn counterpart.

Therefore, we decided that the main fetal indicators to be used to make judgements on the similarity of fetal growth patterns across these low-risk populations will be CRL and head circumference, complemented as a secondary parameter by femur length. For all newborns and preterm infants we will use, as in MGRS, length at birth and in the first 8 months as the main indicator for such comparisons. The conceptual basis for the population comparisons in the INTERGROWTH-21st Project is, therefore, entirely consistent with MGRS.

Structure and management of the project

This comprehensive project was planned as a large-scale, worldwide, collaborative effort (Figure 2). It primarily involves eight participating institutions in widely spread geographical areas and their corresponding research units that, in coordination with the local health authorities, interact with the Project Coordinating Unit at the University of Oxford (Figure 7). This structure facilitates integration of the large number of researchers and clinicians needed to implement this multifaceted project, as well as expert advisers and groups. They are vital to the success of a project of this complexity, because not all the expertise required was available in the primary research institutions and external scrutiny of the study procedures is also necessary to reduce bias. Furthermore, we recognised that changing medical practice, especially in disciplines such as obstetric ultrasound, requires strong political support and credibility.

Therefore, we enrolled specific research support units in data management and analysis; ultrasound quality control; neonatal care; anthropometric quality control and standardisation, and perinatal epidemiology (Figure 7). The standardisation process for ultrasound measures is, to the best of our knowledge, unique and complemented by anthropometric standardisation similar to that used in MGRS. Each of these separate units is independent of the daily running of the project to provide, as far as possible, unbiased assessments. Each is led by a recognised world

expert in the field and has four or five members. In total, we estimate that this large-scale, international collaboration involves over 200 scientists, clinicians, researchers and collaborators supporting close to 60 000 families with the aim of improving maternal and infant health worldwide.

We believe that this model promotes high-quality research and helps the studies to be completed in a relatively short period of time. Importantly, we have created a dynamic, independent, international network of channels reaching the local level, for implementation of effective interventions. In our opinion, it should replace the classic unilateral model of collaboration.

Discussion

In this paper, we have described the structure and organisation of a comprehensive research project designed to elucidate the normal pattern of fetal, neonatal and preterm growth across different geographical areas from conception to infancy, as recommended 20 years ago by WHO's corresponding Expert Committee.[42]

We also aim to explore the characteristics of subgroups of premature newborns and/or those with impaired fetal growth to gain a better understanding of these conditions. This project, we believe, is unique because it combines the following characteristics: (1) it uses a prescriptive approach to construct growth standards for fetuses and newborns along the lines recommended by WHO advisory committees and following MGRS;[6] (2) unlike most of its obstetric predecessors, it is population-based rather than hospital-based; (3) it covers different ethnic populations across continents; (4) it has a comprehensive, standardised, quality control strategy for all fetal ultrasound measures (seldom used in previous studies) and newborn anthropometric measures, run by a centralised quality control unit; (5) it has a large sample size that allows stratification and exploration of a wide range of biological questions; (6) it relates, for the first time, longitudinal fetal growth patterns to perinatal outcomes to establish levels of risk associated with different growth trajectories; (7) the epidemiological evaluation of the preterm delivery and IUGR syndromes is based, for the first time, on new phenotypic classifications that depart from the classical, descriptive subgroups based only on delivery information, e.g. premature rupture of membranes or caesarean section; (8) for these epidemiological analyses, gestational age is estimated accurately using both LMP and early ultrasound examination, standardised across all countries; and (9) it will provide specific clinical products (e.g. a symphyseal–fundal height standard) for use at different levels of care as primary screening tools.

Of course, implementing a project of this nature, which includes studies being performed across several continents, presents major challenges for any research team. Recruiting ~56 000 pregnant women from eight centres is a huge task, specifically as the number of eligible participants for some studies has varied across sites and maternity units. We anticipated an average eligibility rate of 60% for low-risk women. However, the actual rate was lower (~40%) in some populations, especially in developed counties, paradoxically because of, for example, higher than expected smoking rates and the very high proportion of overweight and older women. Fortunately, each centre devised local strategies to maximise recruitment rates for FGLS, which are described in the country-specific papers of this supplement. Importantly, so far, the participants have been enthusiastic about continuing in follow-up studies and we expect that, overall, the loss to follow-up rate will be <3%.

In addition, standardisation of a large number of health professionals operating in varied parts of the world and monitoring the quality of data produced by different centres are always major challenges, especially in obstetric ultrasound, a field that does not have a long tradition of standardising practices. Considerable organisation was required to plan and coordinate the implementation of a single ultrasound protocol with the same ultrasound machine in each centre. In one country, difficulty in importing the new model, which had not yet been approved for clinical use, delayed the start of FGLS at that site. Nevertheless, using the same machine, specially prepared for FGLS, at each site is a major strength of the study and should result in greater consistency of the ultrasound measurements.

Similarly, follow-up of the preterm newborns presents considerable challenges: (1) difficulties measuring very small infants, especially those in an incubator; (2) morbidities associated with prematurity, which may affect the measures to be taken even after discharge; (3) mortality affecting the final sample size and the follow-up; and (4) standardisation of feeding practices across sites, including the promotion of breastfeeding. The overall proportion of preterm births in this healthy, low-risk cohort is expected to be very low (close to 5%) with a small number of births $<33^{+0}$ weeks of gestation. Nevertheless, we expect that adherence to the protocol for the preterm babies will be very high given the parents' concerns for their health. This increases power and reduces the problem of the small number of preterm infants. Interestingly, the final rate of preterm birth ($<37^{+0}$ weeks of gestation) in this cohort can be used as the target, lowest level for comparative purposes in the future.

Selecting an adequately powered sample size for studies such as this is always a compromise between statistical estimates, logistics, cost and the different objectives of the studies—while being aware of the need not to overextend the research team by attempting to answer too many questions. Our first challenge was to have a sample

size in FGLS large enough to create the cohort for the preterm postnatal growth standards (the smallest of our populations) and, second, to have stable outer centiles for the birthweight for gestational age standards. We also needed to have sufficient power to explore population-specific growth in FGLS, in the event that population differences did emerge from the data on the main fetal growth indicators. Having said that, we are confident of the power of this study, which will be analysed using statistical methods appropriate for longitudinal (repeated) measurements.[31,43]

Based on the findings of MGRS,[44–46] we think it unlikely that each study population will have its own distinct growth pattern. Nevertheless, it is possible that some of the fat-dependent ultrasound measures such as abdominal circumference will differ in certain populations, although the challenge here, as in all similar studies, is to define what is a biologically relevant 'distinct growth pattern'. We shall explore this question in sensitivity analyses, having decided a priori, following MGRS principles, that fat-free mass indicators, e.g. CRL and head circumference, are the markers of fetal growth to assess variability across and within populations. Constructing growth curves is a sophisticated statistical task but the experience gained in MGRS will help considerably.[31]

However, perhaps the most important challenge to overcome is maintaining the motivation of the local investigators and data collection teams throughout the project. They are the people who will guarantee the quality of the project, the achieved sample size and eventually the implementation of its results. As we are close to the final sample needed, our experience to date has confirmed that the teams we selected are undertaking research of the highest quality.

The specific aims of this project are to produce new, scientifically valid tools for use at different levels of perinatal care to complement existing clinical tools. We shall produce a new, international, symphyseal–fundal height chart for use in routine antenatal care worldwide and ultrasound standards to evaluate fetal growth. However, we have debated the obvious question: what to do if IUGR is diagnosed using these charts? To contribute to the debate, we have published a series of review papers critically evaluating interventions for fetal growth restriction described in several randomised controlled trials[17,47] and another group's recent update.[48] Unfortunately, there are too few interventions for treating IUGR, other than planned elective delivery. This explains why referral to an adequate level of perinatal care is so important an option. Finally, the results from the epidemiological studies of the phenotypes of preterm and IUGR syndromes could contribute to a better understanding, and therefore identification of, more preventive and clinical interventions.

Our study focuses on helping clinicians to detect impaired fetal growth accurately, which should avoid incorrect diagnoses and, thereby, iatrogenic preterm births. We strongly believe that the misuse of technology, including ultrasound, is one of the factors responsible for unnecessary medical interventions. We think that interventions can only be effective if they focus on the factors responsible for the growth restriction. Small for gestational age is a very heterogeneous condition and it is unlikely that a 'silver bullet' will resolve all cases even if correctly detected. Overall, we believe that accurate early identification (avoiding false positives) and appropriate referral to adequate levels of obstetric and newborn care will have a great impact on neonatal morbidity and mortality, as well as resource allocation in developing countries.

In summary, the new fetal international standard will facilitate the correct interpretation of uterine height values as a first level of screening for fetal growth alterations. It will be complemented by ultrasound scanning at levels of care where it is required by the risk profile of the population served. It is anticipated, therefore, that the new set of complementary charts will have a major clinical impact on overall antenatal and postnatal care. For example, in the case of ultrasound screening, use of the charts is likely to result in fewer unnecessary interventions, such as caesarean sections, because of an incorrect diagnosis of IUGR. We anticipate the newborn standards being used in all institutions where perinatal care is delivered, including rural areas, and the preterm standards being used in NICUs and at other levels of secondary care. We believe this is the first, comprehensive approach – from early pregnancy to infancy – to evaluate growth and its consequences and that it is the most effective way to improve growth monitoring and clinical care for pregnant women and newborns.

Disclosure of interests

None.

Contribution to authorship

JV and SHK wrote the manuscript and all authors read and approved the final version.

Details of ethics approval

The INTERGROWTH-21st Project was approved by the Oxfordshire Research Ethics Committee 'C' (reference: 08/H0606/139), and the research ethics committees of the individual participating institutions and corresponding health authorities where the Project was implemented.

Funding

This Project was supported by the INTERGROWTH-21st Grant ID# 49038 from the Bill & Melinda Gates Founda-

tion to the University of Oxford, for which we are very grateful.

Acknowledgements

A full list of Members of the International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st) and its Committees appears in the preliminary pages of this supplement. ■

References

1 Villar J, Ba'aqeel H, Piaggio G, Lumbiganon P, Miguel Belizan J, Farnot U, et al. WHO antenatal care randomised trial for the evaluation of a new model of routine antenatal care. *Lancet* 2001;357:1551–64.
2 Carroli G, Rooney C, Villar J. How effective is antenatal care in preventing maternal mortality and serious morbidity? An overview of the evidence. *Paediatr Perinat Epidemiol* 2001;15(Suppl. 1):1–42.
3 Whitworth M, Bricker L, Neilson JP, Dowswell T. Ultrasound for fetal assessment in early pregnancy. *Cochrane Database Syst Rev* 2010;4:CD007058.
4 Neilson JP. Symphysis–fundal height measurement in pregnancy. *Cochrane Database Syst Rev* 2007;4:CD000944.
5 de Onis M, Garza C, Onyango AW, Martorell R. WHO Child Growth Standards. *Acta Paediatr* 2006;450:1–101.
6 de Onis M, Onyango AW, Van den Broeck J, Chumlea WC, Martorell R. Measurement and standardization protocols for anthropometry used in the construction of a new international growth reference. *Food Nutr Bull* 2004;25(Suppl. 1):S27–36.
7 Villar J, Valladares E, Wojdyla D, Zavaleta N, Carroli G, Velazco A, et al. Caesarean delivery rates and pregnancy outcomes: the 2005 WHO global survey on maternal and perinatal health in Latin America. *Lancet* 2006;367:1819–29.
8 Villar J, Abdel-Aleem H, Merialdi M, Mathai M, Ali MM, Zavaleta N, et al. World Health Organization randomized trial of calcium supplementation among low calcium intake pregnant women. *Am J Obstet Gynecol* 2006;194:639–49.
9 Gulmezoglu AM, Villar J, Ngoc NT, Piaggio G, Carroli G, Adetoro L, et al. WHO multicentre randomised trial of misoprostol in the management of the third stage of labour. *Lancet* 2001;358:689–95.
10 Kelly Y, Panico L, Bartley M, Marmot M, Nazroo J, Sacker A. Why does birthweight vary among ethnic groups in the UK? Findings from the Millennium Cohort Study. *J Public Health* 2009;31:131–7.
11 Kramer M, Victora CG. Low birth weight and perinatal mortality. In: *Nutrition and Health in Developing Countries*. Totowa, NJ: Humana Press; 2001. pp. 57–69.
12 Black RE, Allen LH, Bhutta ZA, Caulfield LE, de Onis M, Ezzati M, et al. Maternal and child undernutrition: global and regional exposures and health consequences. *Lancet* 2008;371:243–60.
13 Villar J, Abalos E, Carroli G, Giordano D, Wojdyla D, Piaggio G, et al. Heterogeneity of perinatal outcomes in the preterm delivery syndrome. *Obstet Gynecol* 2004;104:78–87.
14 Villar J, Carroli G, Wojdyla D, Abalos E, Giordano D, Ba'aqeel H, et al. Preeclampsia, gestational hypertension and intrauterine growth restriction, related or independent conditions? *Am J Obstet Gynecol* 2006;194:921–31.
15 Villar J, Gulmezoglu AM, de Onis M. Nutritional and antimicrobial interventions to prevent preterm birth: an overview of randomized controlled trials. *Obstet Gynecol Surv* 1998;53:575–85.
16 Gulmezoglu M, de Onis M, Villar J. Effectiveness of interventions to prevent or treat impaired fetal growth. *Obstet Gynecol Surv* 1997;52:139–49.
17 de Onis M, Villar J, Gulmezoglu M. Nutritional interventions to prevent intrauterine growth retardation: evidence from randomized controlled trials. *Eur J Clin Nutr* 1998;52(Suppl 1):S83–93.
18 Iams JD, Romero R, Culhane JF, Goldenberg RL. Primary, secondary, and tertiary interventions to reduce the morbidity and mortality of preterm birth. *Lancet* 2008;371:164–75.
19 Zhang J, Villar J, Sun W, Merialdi M, Abdel-Aleem H, Mathai M, et al. Blood pressure dynamics during pregnancy and spontaneous preterm birth. *Am J Obstet Gynecol* 2007;197:162–e1–e6.
20 Merchant KM, Villar J. Maternal height and newborn size relative to risk of intrapartum caesarean delivery and perinatal distress. *BJOG* 2001;108:689–96.
21 Lynch CD, Zhang J. The research implications of the selection of a gestational age estimation method. *Paediatr Perinat Epidemiol* 2007;21(Suppl. 2):86–96.
22 Robinson HP, Fleming JE. A critical evaluation of sonar ''crown–rump length'' measurements. *Br J Obstet Gynaecol* 1975;82:702–10.
23 Verburg BO, Steegers EAP, De Ridder M, Snijders RJM, Smith E, Hofman A, et al. New charts for ultrasound dating of pregnancy and assessment of fetal growth: longitudinal data from a population-based cohort study. *Ultrasound Obstet Gynecol* 2008;31:388–96.
24 Thorsell M, Kaijser M, Almstrom H, Andolf E. Expected day of delivery from ultrasound dating versus last menstrual period—obstetric outcome when dates mismatch. *BJOG* 2008;115:585–9.
25 Group MGRS, de Onis M. Complementary feeding in the WHO Multicentre Growth Reference Study. *Acta Paediatr Suppl* 2006;450:27–37.
26 Papageorghiou AT, Sarris I, Ioannou C, Todros T, Carvalho M, Pilu G, et al. for the International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st). Ultrasound methodology used to construct the fetal growth standards in the INTERGROWTH-21st Project. *BJOG* 2013; DOI: 10.1111/1471-0528.12313.
27 Sarris I, Ioannou C, Ohuma EO, Altman DG, Hoch L, Cosgrove C, et al. for the International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st). Standardisation and quality control of ultrasound measurements taken in the INTERGROWTH-21st Project. *BJOG* 2013; DOI: 10.1111/1471-0528.12315.
28 Belizan JM, Villar J, Nardin JC, Malamud J, De Vicurna LS. Diagnosis of intrauterine growth retardation by a simple clinical method: measurement of uterine height. *Am J Obstet Gynecol* 1978;131:643–6.
29 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.
30 Villar J, Knight HE, de Onis M, Bertino E, Gilli G, Papageorghiou AT, et al. Conceptual issues related to the construction of prescriptive standards for the evaluation of postnatal growth of preterm infants. *Arch Dis Child* 2010;95:1034–8.
31 Borghi E, de Onis M, Garza C, Van den Broeck J, Frongillo EA, Grummer-Strawn L, et al. Construction of the World Health Organization child growth standards: selection of methods for attained growth curves. *Stat Med* 2006;25:247–65.
32 Bhutta Z, Haroon A, Knight H, Albernaz A, Batra M, Bhat B, et al. The International Fetal and Newborn Growth Consortium (INTERGROWTH-21st): The Neonatal Manual [www.intergrowth21.org.uk]. Accessed 15 November 2010.
33 Villar J, Carroli G, Zavaleta N, Donner A, Wojdyla D, Faundes A, et al. Maternal and neonatal individual risks and benefits associated

with caesarean delivery: multicentre prospective study. *BMJ* 2007;335:1025.

34 Kramer MS, Ananth CV, Platt RW, Joseph KS. US Black vs White disparities in foetal growth: physiological or pathological? *Int J Epidemiol* 2006;35:1187–95.

35 Joseph KS, Fahey J, Platt RW, Liston RM, Lee SK, Sauve R, et al. An outcome-based approach for the creation of fetal growth standards: do singletons and twins need separate standards? *Am J Epidemiol* 2009;169:616–24.

36 Kramer MS, Papageorghiou AT, Culhane JF, Bhutta ZA, Goldenberg RL, Gravett M, et al. Challenges in defining and classifying the preterm birth syndrome. *Am J Obstet Gynecol* 2012;206:108–12.

37 Goldenberg RL, Gravett MG, Iams J, Papageorghiou AT, Waller SA, Kramer M, et al. The preterm birth syndrome: issues to consider in creating a classification system. *Am J Obstet Gynecol* 2012;206:113–8.

38 Villar J, Papageorghiou AT, Knight HE, Gravett MG, Iams J, Waller SA, et al. The preterm birth syndrome: a prototype phenotypic classification. *Am J Obstet Gynecol* 2012;206:119–23.

39 Goedhart G, van Eijsden M, van der Wal M, Bonsel G. Ethnic differences in term birthweight: the role of constitutional and environmental factors. *Paediatr Perinat Epidemiol* 2008;22:360–368.

40 Mortensen LH, Diderichsen F, Arntzen A, Gissler M, Cnattingius S, Schnor O, et al. Social inequality in fetal growth: a comparative study of Denmark, Finland, Norway and Sweden in the period 1981–2000. *J Epidemiol Community Health* 2008;62:325–31.

41 Ohuma EO, Hoch L, Cosgrove C, Knight HE, Cheikh Ismail L, Juodvirsiene L, et al. for the International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st). Managing data for the international, multicentre INTERGROWTH-21st Project. *BJOG* 2013; DOI 10.1111/1471-0528.12080.

42 de Onis M, Habicht JP. Anthropometric reference data for international use: recommendations from a World Health Organization Expert Committee. *Am J Clin Nutr* 1996;64:650–8.

43 Altman DG, Ohuma EO, for the International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st). Statistical considerations for the construction of prescriptive growth standards. *BJOG* 2013; DOI 10.1111/1471-0528.12031.

44 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 World Health Organ* 2002;80:189–95.

45 Owusu WB, Lartey A, de Onis M, Onyango AW, Frongillo EA. Factors associated with unconstrained growth among affluent Ghanaian children. *Acta Paediatr* 2004;93:1115–9.

46 WHO Multicentre Growth Reference Study Group. *WHO Child Growth Standards: Growth Velocity Based on Weight, Length and Head Circumference: Methods and Development*. Geneva: World Health Organization, 2009.

47 Kulier R, de Onis M, Gulmezoglu AM, Villar J. Nutritional interventions for the prevention of maternal morbidity. *Int J Gynaecol Obstet* 1998;63:231–46.

48 Bhutta Z. Balanced protein energy supplementation in pregnancy. 2nd Meeting of the Advisory Board. Improving maternal, neonatal and child health outcomes through better designed nutrition policies and programs. Atlanta, GA.

