Fetal growth and risk assessment: is there an impasse?

THE ISSUE: Fetal growth restriction is an indicator of placental insufficiency and is strongly associated with adverse perinatal outcome. There is a point that the recent dominance in the medical literature about which reference charts to use and dichotomization of fetal size at the 10th percentile overlooks the fact there is not a single cut-off in any growth chart that acts as an absolute divider between high and low risk for adverse outcome. Thus, the collective goal of all researchers to identify, monitor and effectively manage growth-restricted fetuses is better served by replacing dichotomisation of normal versus abnormal fetal growth at the 10th percentile by interpretation of fetal size in context with other known parameters of fetal risk - all as continuous parameters. The use of prospective comprehensive datasets should facilitate better risk assessment for the individual fetus, to help direct effective and appropriate interventions. The counterpoint is that the debate about which growth standard to use was necessary and has been settled through evidence that size, and therefore growth, need customized limits to allow adjustment for constitutional variation, and to help distinguish between normal and abnormal growth. Implementation of a more precise standard has led to better detection of fetuses that are at risk due to growth restriction, improved application of additional investigations, enhanced clinical confidence in management including timely delivery, and ultimately increased prevention of adverse outcomes.

Point

Wessel Ganzevoort, MD, PhD
Department of Obstetrics & Gynaecology, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands

Baskaran Thilaganathan, MD, PhD
Fetal Medicine Unit, St. George’s University Hospitals NHS Foundation Trust, London and Molecular and Clinical Sciences Research Institute, St George’s University of London, London, UK

Ahmet Baschat, MD
Center for Fetal Therapy, Department of Gynecology & Obstetrics, Johns Hopkins University, Baltimore, MD

Sanne J. Gordijn, MD, PhD
Department of Obstetrics & Gynaecology, University Medical Centre Groningen, University of Groningen, Groningen, The Netherlands

Counterpoint

Jason Gardosi, MD, FRCOG
Director, Perinatal Institute, Birmingham, UK

The small for gestational age (SGA; <10th percentile) fetus is at a 7-fold increased risk of dying before delivery.1 Size is an important indicator of fetal growth. Yet, physiologic variation needs to be adjusted for or customized to improve the association with pathologic smallness. The debate about which charts to use needed to be had, and ample evidence has emerged that proves that 1 size does not fit all. An improved standard helps to reassure when smallness is only constitutional and helps to better identify when it is potentially pathologic. It can prompt additional investigations to assess the presence and severity of placental insufficiency and increases clinical confidence when to deliver the at-risk fetus from a hostile intrauterine environment. There is therefore no impasse but, in fact, good progress to translate evidence into preventive strategies, which has already shown significant improvements in stillbirth rates in England.2

Size: an important indicator of growth

Size matters. Ultrasound biometry is required at the beginning of pregnancy to establish the gestation dates and, subsequently, to determine the growth that has occurred up to respective points in pregnancy, hence expressed as size (continued)
Dr Ganzevoor (continued)

predominant. We argue that this impasse distracts from the actual issue and wastes both clinical and academic resources. In this article, we highlight how the focus on fetal size as a proxy for fetal growth and adequate placental nutrition has been oversimplified as “above” or “below” the 10th percentile to distinguish between apparently normal and abnormal fetal growth. This dichotomization results in erroneous underdiagnosis of growth-restricted fetuses among those that apparently are grown normally with the risk of adverse outcomes because of lack of medical attention. Conversely, among constitutionally small fetuses, it leads to overdiagnosis of FGR and risk from unnecessary obstetric intervention. FGR can result from any disease that affects placental function and is associated with significant adverse short- and long-term outcomes. Accordingly, adequate detection and risk stratification is of paramount importance to guide perinatal care. With >10,000 citations on “prenatal diagnosis” or “definition” of FGR in current medical literature, we have achieved little progress beyond the initial landmark observation that fetal size apparently is expressed optimally by ascribing a percentile to its estimate. This article argues that we should progress beyond fetal size assessment alone and undertake a more comprehensive risk assessment with the use of contemporary techniques.

Which fetal growth chart?

Population-based reference charts

Population-based fetal size charts are created from retrospective datasets and by nature are descriptive and show how the fetuses in the observed population have grown. These references are skewed at the extremes of gestation where pathologic conditions, such as preeclampsia and preterm birth, are concentrated, because these abnormal pregnancies usually are not excluded in their development. This still holds true for more recent descriptive reference charts.

Customized fetal growth assessment

To overcome some of the methodologic drawbacks of population-based growth charts, Gardosi et al. constructed growth charts that attempt mathematically to predict normal variation in growth at term. This group introduced the idea of customization by correcting for maternal characteristics that individualize the expected growth potential of the fetus. Many variables have been used for correction such as fetal gender, maternal height, weight, ethnicity, and parity. At first glance, ethnicity seems to be an intuitive variable for customization, but ethnicity is often associated

Dr Gardosi (continued)

(or weight)-for-gestation, as standard deviations or, more commonly, as percentiles. In the absence of an inbuilt sensor to assess growth rate longitudinally, such measurements provide snapshots from which growth can be assessed. Birthweight-for-gestation is another snapshot that imputes growth during pregnancy and is a predictor of risk in the neonatal period and subsequent development.

Weight-for-gestation becomes most important at the extremes. There is little clinical relevance whether a single measurement of a fetus or neonate is at the 40th or 60th percentile; however, cut-offs of 3rd, 5th, and 10th percentile to define SGA have important associations with adverse outcomes. Size is a proxy for growth restriction; in many instances, there can be significant placental failure without the fetal weight dropping below any of these cut-off limits. For example, in a term multiethnic population, SGA at <3rd percentile fetuses have a relative risk of 4.2 for term stillbirth; however a similar number of deaths occurred in fetuses with less severe SGA (>3 and <10th percentile), with a still significant relative risk of 1.8. Although approximately 40% of stillbirths are below the 10th percentile (after adjustment for the average delay between fetal death and delivery), there is also a significantly larger proportion of cases of 10–50th percentile than 50–90th percentile in the distribution of stillbirths at term.

Birthweight and SGA are also important measures to audit performance in antenatal growth surveillance. Apart from research tools such as air displacement plethysmography or catch-up growth during infancy at population level, we have no ready indicators from which to determine whether a neonate had restricted growth in utero; in most cases, the condition has not been recognized antenatally. Thus, SGA birthweight has become the incomplete, yet easily available, method to assess the growth status at the end of pregnancy.

A customized standard improves precision

In contrast to biophysical assessment of placental function by Doppler flow, normal ranges of fetal biometry are affected by constitutional variation, which must be adjusted for to improve the association between physiologic and pathologic variation. Fetal sex and maternal height, weight in early pregnancy, parity, and ethnic origin are variables that have a significant influence on the growth of the fetus and the weight of the neonate.

Figure 1 demonstrates the stepwise improved prediction of birthweight within the mid-tertile weight distribution of a pregnancy cohort. The adjustment factors (coefficients) that are used to predict the weight at the end of a normal pregnancy are derived from a normal population of term deliveries and exclude pathologic factors that affect
with poorer socioeconomic status that may be the determinant for adverse perinatal outcomes.\textsuperscript{10} Definition and self-reported categorization of race is difficult, particularly in a multiethnic and continuously mingling society.\textsuperscript{11} Additionally, how much can change in transition from a generation of severe deprivation to a well-nourished and healthy generation is not known. We suggest that individual variables should be assessed comprehensively for their individual putative merits and their well-known existing relationship with adverse outcomes. In the available datasets of second and third generation migrants, we can test how many generations are required for adverse effects to subside. Similar interactions may be observed for maternal weight and parity, but fetal gender and maternal height may have a stronger argument to be used in customization.\textsuperscript{7} Even though there may be an academic discussion about the concept of customization, the initiative to use backcalculated growth charts from a healthy term cohort and serially to plot growth assessments in a systematic manner have significantly improved awareness about growth and identification of fetuses at risk.\textsuperscript{9,12,13}

Adjusting for ethnic origin

It is sometimes argued that ethnicity should not be adjusted for, because some minority groups have an increased risk of adverse pregnancy outcome. However, this is a superficial argument, because the reason for higher rates of adverse outcome in these groups are not likely to be due to growth restriction, which is often suspected wrongly if a population-based standard is used. A misdirected, exclusive focus on growth restriction may avoid attention and the increased awareness about the adverse effect of social deprivation.
management of other risk factors, such as social deprivation or diabetes mellitus, as potential causes of excess morbidity and death.

Ethnic differences in fetal growth and birthweight have been confirmed in a number of studies of low-risk populations. In an analysis of a cohort of South Asian pregnancies in England with 10,405 third-trimester ultrasound estimated fetal weights, we found that 56% of cases that were considered SGA by the commonly used Hadlock estimated fetal weight standard were not SGA if the standard was customized. Significantly, this subgroup of reclassified cases had the same mortality risk as those not designated SGA by either Hadlock or the customized standard. Figure 2 illustrates the effect of plotting the same EFW on 2 growth charts that are customized for different maternal characteristics.

The end-effect of a more precise standard is not only improved recognition of pregnancies that are at risk but also reduced false alarms and, hence, fewer unnecessary investigations and interventions. Use of the customized standard for birthweight and antenatal assessment is recommended by the Royal College of Obstetricians and Gynaecologists’ guidelines on the investigation and management of the SGA fetus.

The still promoted multinational Intergrowth 21st standard sought to prescribe a one-size-fits-all approach, with the implication that any population that did not fit that standard is stunted because of chronic malnutrition. However, this claim has been contradicted by a subsequent, conceptually similar multinational study by the World Health Organization and by country-specific studies in New Zealand, Hong Kong, and England as well as a 10-country comparison with 1.2 million births that found that customized definition of SGA identified significantly more babies who were at risk of stillbirth. In light of such evidence, it is difficult to see a reason why any clinician would still want to use a single standard for all mothers in their care.

Detection of slow growth
Recognition and management of SGA and/or slow growth should start with pre-pregnancy advice and awareness of risk factors for early pregnancy assessment. Lists of factors have been published by the RCOG and by NHS England, and include those relating to medical and obstetric history as well as maternal age, obesity and smoking status. The guidelines recommend that pregnancies without risk factors are to be monitored by serial fundal height measurement, while those with increased risk are monitored by serial ultrasound. One-off ‘routine’ third trimester scans are discouraged as there is no evidence for their effectiveness.
Prescriptive growth standards

To a certain extent, healthy populations across the world are expected to have similar fetal growth, because only 1 species of humans exists without large phylogenetic differences. Recent initiatives prospectively have followed healthy uncomplicated pregnancies with sequential ultrasound scans to develop prescriptive growth standards and to define how a healthy fetus grows. In the Intergrowth-21 study that was performed in 8 different countries, measurement variation between countries was significantly smaller than within-population variation. This uniformity suggests that prescriptive growth charts are the gold standard. However, there is persistent significant variation in fetal growth within each population when environmental constraints are not controlled for adequately. Apparently, these factors have a significant adverse influence on fetal growth and not all fetuses grow the same.

The way forward

Obviously, these approaches are conceptually different, and protagonists of either approach are in fundamental disagreement. Much effort is put into comparing how either approach performs retrospectively in large datasets. These analyses have inherent methodologic flaws. We postulate that, if the strength of each approach is appreciated openly and academically, we may come to a combined approach using prescriptive charts that use clinically validated with the use of effective customization. We propose a combined approach by merging datasets on an individual level to test whether variables that are included in customization or the concept of conditional percentiles may be used to determine and they may in fact be harmful by providing false reassurance.

There is increasing evidence of the usefulness of biomarkers and uterine artery Doppler in the detection of early onset FGR and pre-eclampsia, but not late onset FGR,
optimal growth for the individual.22 Next, the relationship between these variables and adverse outcomes can and must be explored, although this will remain problematic in retrospective datasets because of treatment paradox. Currently, many retrospective studies compare different strategies in their diagnostic capacity to identify small infants at birth or those infants with adverse outcomes. These comparisons are methodologically flawed because they analyze retrospectively whether these strategies accurately predict birthweight category or adverse outcomes, while ignoring the effects of treatment paradox. Moreover, these comparisons also overlook the issue of the balance of detection rates and false-positive rates that prospective and randomized trials are lacking.23 We propose that the combination of the 2 approaches will help achieve the penultimate goal to define normal individual fetal growth patterns. But we should not stop there.

What we are missing?
How relevant is fetal size? Size is a consequence of preceding fetal growth that reflects fetal nutrition. Current practice dichotomizes normal and abnormal fetal growth at the 10th percentile. But we need to move beyond “good” or “bad” fetal size. For most continuous measures in human physiologic condition, this is an oversimplification with significant sources of error. First, most fetuses who are identified as small for gestational age are constitutionally small and healthy. Second, many fetuses who have impaired growth and placental dysfunction are of apparent normal weight (Figure 1).24 Third, fetal size only reflects nutrient transfer function of the placenta and points to the magnitude of placental dysfunction only by association.25 Stillbirth risk decreases with higher birthweight percentiles, and the majority of fetuses who are born <3rd percentile are known to have been exposed to significant intrauterine hypoxemia. However, there is no percentile above which this risk is excluded (Figure 2).24 This is further complicated by the poor performance of ultrasound-based fetal growth assessment that detects only up to 50% of babies who are born weighing <10th percentile.26 For the aforementioned reasons, fetal size would be better utilized as a continuous variable in risk calculation.

Clinical outcomes of relevance
The objective of obstetric care is not only to diagnose fetal malnutrition but also to prevent the negative effects of placental dysfunction that include fetal hypoxemia, brain damage, and stillbirth. In postnatal life, protracted nutritional deprivation results in kwashiorkor or marasmus many weeks or months before infant death. In contrast, an infant survives only a very short period with respiratory failure. The placenta is uniquely responsible for many

(continued)

Dr Gardosi (continued)
which constitutes the majority of cases.29 Biophysical assessment of placental function including umbilical artery, venous and middle cerebral Doppler or cerebro-placental ratio (CPR) can be useful adjuncts in the assessment of third trimester growth,29 and help to identify the optimal time of delivery. However the diagnostic tree starts with SGA30 and the outcome is most strongly linked with growth, or the change in size between consecutive measurements.31,32

Cerebro-placental ratio (CPR) is being promoted in some quarters as an independent indicator of significant growth restriction. There is evidence that abnormal CPR is associated with perinatal morbidity and mortality.33,34 However, the clinical application of these findings remains to be determined. The largest prospective series on unselected pregnancies have failed to find predictive value of CPR for adverse outcome when performed at 30-3435 or at 35-37 weeks.36 Although a low CPR is associated with adverse outcome, as a screening test its predictive performance is relatively poor. Recent systematic reviews37,38 have shown that performance varies with population characteristics and specific outcome (being better in SGA suspected babies); hence size matters here too. Even if CPR can be demonstrated to be a marker of placental insufficiency independent of size, the effectiveness of a strategy based on CPR assessment in the overall population is still to be proven.29

Strategies for prevention
So although screening tools are still incomplete, there is enough evidence of the avoidability of many adverse outcomes, especially stillbirths, and hence the need to recognise the urgency to progress with preventive measures based on what we already know.

The Growth Assessment Protocol (GAP) synthesised such evidence and has recently been described in detail.8 In brief it includes 1. Early pregnancy risk assessment; 2. Hands-on training in fundal height measurement and referral pathways; 3. Standardised protocols for surveillance in low and high risk pregnancies; 4. Implementation and training in the use of customised growth charts; and 5. Audit tools for monitoring detection rates and analysing the reasons for missed cases.

The programme was implemented in the majority of UK hospitals over the last 8 years, and has resulted in a two- to three-fold increase in antenatal detection of babies that are SGA at birth (with the presumed simultaneous increase in detection of babies that are growth restricted but not SGA).8,27 Over the same period, there has been a year on year drop in stillbirth rates in England to their lowest ever levels,2 resulting in a 23% decrease which represents 860 fewer deaths per year. Significantly, evidence is emerging that it is specifically SGA stillbirths that are being prevented

(continued)
critical body functions, namely nutrition and respiration. To date, clinicians have used fetal size/nutrition as a proxy for placental respiratory failure and risk of stillbirth to avert this risk by timely delivery.27 Many studies evaluate diagnostic tools or interventions by their ability to identify or prevent small babies. The outcomes of relevance, however, should be the variables that indicate fetal hypoxemia (such as stillbirth), the inability to withstand uterine contractions, and long-term neurodevelopmental outcomes.

A comprehensive approach
Focusing on fetal size parameters in isolation to detect compromise is a grossly oversimplified diagnostic approach and, as such, is flawed.28 Placental function is reflected across a number of variables that are associated with adverse outcomes that can be examined prenatally.25 These include Doppler ultrasound scanning of the fetal umbilical, middle cerebral, and maternal uterine arteries, serum biomarkers, and growth trajectory.26,30 A recent international expert consensus recognized that parameters that indicate placental respiratory function should be included in the assessment.31,32 Currently, these risk factors are used in a categoric fashion, where they have a dose-dependent relationship with poor fetal growth and stillbirth. Risk factors and assessment variables are interdependent, which is a fact that often is disregarded in risk assessment tools that are recommended by national institutions. Finally, stillbirth is a time-dependent outcome rather than an overt disease; as such, it is extremely susceptible to the competing risk and intervention bias of elective delivery.

It is now possible to use contemporary software tools to generate competing algorithms to undertake comprehensive risk assessment. We propose the development of a predictive approach that takes into account relevant variables in a continuous, nondichotomous manner. Large datasets with longitudinal prospective data currently are generated from on-going clinical intervention trials. We need aggregate individual data from these datasets to allow prediction modelling followed by internal and external validation. Within these algorithms, institutions will be able to choose either type of growth chart (prescriptive or customized) depending on local, regional, or national interpretation. Once individual risk assessment is available, the rational next step is to feed intervention trials.

Conclusion
Currently, the 1 issue that dominates much of the debate among valued individual researchers is which growth chart to use to assess normal or abnormal fetal size. We argue that we need to abandon that gridlock because it obscures the bigger and more clinically relevant picture. The current assessment tools that are recommended by national institutions. Finally, stillbirth is a time-dependent outcome rather than an overt disease; as such, it is extremely susceptible to the competing risk and intervention bias of elective delivery.

It is now possible to use contemporary software tools to generate competing algorithms to undertake comprehensive risk assessment. We propose the development of a predictive approach that takes into account relevant variables in a continuous, nondichotomous manner. Large datasets with longitudinal prospective data currently are generated from on-going clinical intervention trials. We need aggregate individual data from these datasets to allow prediction modelling followed by internal and external validation. Within these algorithms, institutions will be able to choose either type of growth chart (prescriptive or customized) depending on local, regional, or national interpretation. Once individual risk assessment is available, the rational next step is to feed intervention trials.

Conclusion
Currently, the 1 issue that dominates much of the debate among valued individual researchers is which growth chart to use to assess normal or abnormal fetal size. We argue that we need to abandon that gridlock because it obscures the bigger and more clinically relevant picture. The current

REFERENCES
5. Gardosi J. Classification of stillbirth by relevant condition at death (ReCoDe); population based cohort study. BMJ 2009;335:1113–7.
standstill effectively prohibits progress towards the sketched horizon with a comprehensive risk assessment that will benefit the prenatal care for patients.

REFERENCES

Dr Gardosi (continued)