Effects of applying universal fetal growth standards in a Scandinavian multi-ethnic population

Running heading: Universal fetal growth standards

Line SLETNER[1], Torvid KISERUD[2, 3], Siri VANGEN[4, 5], Britt NAKSTAD[1, 6]& Anne Karen JENUM[7]

[1]Department of Pediatric and Adolescents Medicine, Akershus University Hospital, Lørenskog, [2]Department of Obstetrics and Gynecology, Haukeland University Hospital, Bergen, [3]Department of Clinical Science, University of Bergen, Bergen, [4]Norwegian National Advisory Unit on Women`s Health, Women's Clinic, Oslo University Hospital, Oslo, [5]Institute of Clinical Medicine, University of Oslo, Oslo, [6]Institute of Clinical Medicine, Campus Ahus, University of Oslo, Lørenskog, [7]Department of General Practice, Institute of Health and Society, University of Oslo, Oslo, Norway

Corresponding author:

Line Sletner

Department of Pediatric and Adolescents Medicine, Akershus University Hospital, 1478 Lørenskog, Norway.

E-mail: line.sletner@medisin.uio.no

Conflicts of interest

All the authors report no conflict of interest

Abstract

Introduction: The question whether universal growth charts can be used in multi-ethnic settings is of general interest. The Intergrowth-21st (IG-21) fetal growth and newborn size standards are suggested to represent optimal fetal growth regardless of country origin. Our aim was to examine whether women fulfilling the strict IG-21 inclusion criteria were healthier, showed less ethnic differences in fetal growth and newborn size, and less adverse perinatal outcomes. *Material and methods:* Data were drawn from a population-based multi-ethnic cohort of 823 presumably healthy pregnant women in Oslo, Norway. We assessed differences in fetal and neonatal gestational age specific z-scores and compared maternal health parameters, pregnancy- and birth complications between pregnancies fulfilling and not fulfilling the IG-21 criteria. *Results:* Only 21% of pregnancies enrolled in our cohort fulfilled the IG-21 criteria. Fetal growth deviated substantially from the new standards, in particular for ethnic Europeans. Ethnic differences persisted in pregnancies fulfilling the criteria.In South Asian fetuses estimated fetal weight was -0.60 SD (95% CI: -1.00, -0.20) lower at 24 gestational weeks, and birthweight was -0.62 SD (-0.95, -0.29) lower, compared with ethnic Europeans. Corresponding numbers for Middle-East/N-Africans were -0.13 (-0.62, 0.36) and -0.60 (-1.00, -0.20). Maternal health indicators and birth complications were similar in women fulfilling and not fulfilling the criteria, or the relation depended on ethnic origin. *Conclusions:* In an urban multi-ethnic Norwegian population, applying an extensive list of criteria to define "healthy" pregnancies excludes the majority of women but does not cancel ethnic differences in fetal growth.

Keywords

ultrasound biometry, fetal growth, newborn size, ethnic differences, maternal health, birth complications, growth standards

Abbreviations

AC abdominal circumference

CH-length crown-heel length

EFW estimated fetal weight

FL femur length

HC head circumference

IG-21 Intergrowth-21st fetal growth and newborn size standards

Key message

Is one growth standard universally applicable, as concluded by the Intergrowth-21st consortium, or are modifications needed to fit various populations or ethnicities? In our multi-ethnic population, strict inclusion criteria, as in Intergrowth-21st, did not uniformly identify women with better health or less adverse perinatal outcomes, and ethnic differences in fetal growth and newborn size persisted.

Introduction

Growth charts are important clinical tools, and can be constructed as "growth references" or "growth standards". A growth reference simply describes the growth in a given population. A growth standard, on the other hand, is based on the assumption that fetuses regardless of ethnic background grow similarly under optimal health, environmental and socioeconomic conditions. In 2014, new international fetal growth and newborn size standards suggested for universal use were introduced by the Intergrowth-21st (IG-21) consortium, based on samples of pregnant women from all world regions fulfilling strict inclusion criteria (1, 2). The IG-21 consortium concluded that the sample represented healthy pregnant women where nutritional and health needs were met, and hence that the derived growth charts would represent "optimal growth", regardless of ethnic origin (3, 4). However, two recent studies also aiming at producing fetal growth charts for general use, dispute the IG-21 conclusions (5, 6).

In multi-ethnic populations, it is essential to know whether one single growth chart can be used for all, as concluded from the IG-21 study, or if population-specific reference ranges are more appropriate. To address this question, we used the multi-ethnic population-based STORK-Groruddalen cohort in Oslo, Norway, which have detailed information on a wide range of maternal health parameters, almost all factors included in the IG-21 criteria, and fetal growth and newborn size. The aim of the present study was to test the validity of the IG-21 conclusions. If valid we could expect that: 1) applying such criteria would cancel ethnic differences in fetal growth and newborn size, and growth measures would align with the IG-21 sample mean at each time point. 2) women, across ethnic groups, fulfilling the strict IG-21-criteria would be generally healthier and have less adverse perinatal outcomes in comparison to those who do not.

Material and methods

Study sample

In Norway, women with low-risk pregnancies are provided antenatal care in primary health care, free of charge. Most women attend and rates for adverse perinatal outcomes are generally low, although some ethnic differences exist (7, 8). Groruddalen, Oslo, has a multi-ethnic population with a diverse socio-economic status, and the majority of pregnant women (75-85%) attend local Maternal and Child Health Clinics for antenatal care.

Stork Groruddalen is a population-based cohort study of pregnant women recruited during early pregnancy at three local Maternal and Child Health Clinics from May 2008 to May 2010 (9). Questionnaire data was collected by specially trained midwifes through interviews, supported by a professional interpreter using a translated questionnaires when needed. The women were eligible if they 1) were living in one of the three city districts, 2) would give birth at one of the two study hospitals, 3) were at less than 20 gestational weeks, 4) were not suffering from diseases necessitating intensive hospital follow-up during pregnancy (including pre-gestational diabetes), 5) could speak Norwegian or one of eight other languages, and 6) were able to provide informed written consent. Overall participation rate was 74% (ranging between 63 and 82% between ethnic groups) and the study cohort was found representative for women attending the Maternal and Child Health Clinics with respect to ethnicity and age (9).

Maternal variables

Factors included in the Intergrowth-21st study inclusion criteria

The two IG-21 study samples; Newborn Cross-Sectional Study Prescriptive Subpopulation and the Fetal Growth Longitudinal Study, used the same inclusion criteria, hereafter referred to as the "IG-21 criteria" (Table 1). Data was collected at eight study sites, representing all world regions (3). Large study samples were first selected at a population level, by recruiting from hospitals serving populations with middle- and high socio-economic status. Secondly, women were selected at an individual level, based on a number of clinical and nutritional indicators associated with intra-uterine growth restriction (Table 1)(3, 4). In the Stork-Groruddalen cohort we had information on all except four IG-21-criteria ("previous pregnancy affected by pre-eclampsia/eclampsia, HELLP syndrome or a related pregnancy-associated condition", "positive urine-analysis", "clinically significant atypical red cell alloantibodies" and "in an occupation with risk of exposure to chemicals or toxic substances, or very physically demanding activity") (9). The numbers for these four outcomes were presumably low in our population. Furthermore, occupational exposures/activity in relation to pregnancy is strictly regulated by law in Norway. Hence, the last of these criteria will not apply in this setting. As some IG-21-criteria cut-off values were site-specific (e.g. socioeconomic status), we used cut-offs from European sites (3, 10).

Other maternal variables

Maternal and offspring ethnic origin was defined by the pregnant participant's country of birth. If the participant's mother was born outside Europe or North-America, the country of origin was defined by the participant's mother's country of birth. Except for Pakistanis (30% second generation immigrants), very few women of non-European origin were born in Norway. Women were merged into five groups (see footnote Table 1); Europe (n=379), South Asia (n=200), East Asia (n=43), Middle-East/North-Africa (n=172) and Others (n=29).

Pre-pregnant body mass index (kg/m^2) was calculated using self-reported pre-pregnant weight and measured height (9). Maternal skin folds were measured twice (triceps, subscapular and suprailiac, with Holtain T/W Caliper 0-48mm (Holtain Ltd., Crymych; UK), and the mean of the sum was used (11). Weight gain from pre-pregnancy to 28 gestational weeks was calculated from measured weight at 28 gestational weeks and the self-reported pre-pregnant weight. Weight gain from gestational 28 gestational weeks to birth was calculated from total weight gain, self-reported at the post-natal visit, and weight at 28 gestational weeks. Parity was categorized as para 0, para 1 or multipara, provided pregnancies lasted at least 22 gestational weeks. Self-reported regular physical activity at inclusion was classified as "regular physical activity" or "no regular physical activity" according to specific criteria (see footnote Table 2) (9).

Fasting venous blood samples were drawn at inclusion and analysed at certified laboratories (9). Severe vitamin D deficiency was defined as 25(OH) vitamin D below 25 nmol/L. A standard 75 g oral glucose tolerance test was performed at 28±2 gestational weeks. Gestational diabetes mellitus was diagnosed according to the1999 World Health Organization criteria (fasting plasma glucose ≥7 mmol/l or 2-hour plasma glucose ≥7.8 mmol/l) (9, 12).

Pregnancy and birth complications

Data on birth complications and hypertensive disorders in pregnancy (systolic blood pressure >140 mmHg and/or diastolic blood pressure >90 mmHg in early pregnancy, or after 20 gestational weeks, accompanied by new-onset proteinuria) were retrieved from hospital records. A composite birth complication variable was constructed based on the presence of at

least one of three events (acute cesarean section, grade 3-4 perineal rupture and Apgar score < 7 at 5 min).

Offspring variables

Gestational age was derived from the first day of the woman's last menstrual period, unless this differed more than 14 days from that derived from the routine ultrasound measurements at 17-20 gestational weeks (<7% of women); then the latter was preferred. This differs slightly from the Intergrowth-21st study, where gestational age based on an early ultrasound scan (at 9-13 gestational weeks) was used if the last menstrual period dating differed more than seven days (2). Preterm birth was defined as birth before 37+0 gestational weeks.

Fetal ultrasound measurements

We measured head circumference (HC), abdominal circumference (AC), and femur length (FL) (13) and the mean of three measurements was entered in the statistics. Estimated fetal weight (EFW) was calculated using Combs formula: $EFW=(0.23718 \times AC^2 \times FL) + (0.03312 \times HC^3)$ (14). Four experienced clinicians (two obstetricians and two ultrasound midwives, trained at the same hospital) carried out the ultrasound examinations at 24, 32 and 37 gestational weeks according to a study-specific standardized protocol using Voluson Pro (GE-Healthcare) with a AB2-7 scan head. At each session, participants were randomly allocated to the available ultrasound examiner.
Interrater sessions were performed on 24 pregnant women. Two examiners participated in each session, and they performed in total 8-15 examinations each. The intra-class correlation coefficient ranged from 0.95 (FL) to 0.99 (HC, AC and EFW.

Neonatal measurements

Birthweight was routinely measured in grams on calibrated electronic scales immediately after birth. Within 72 hours after birth, detailed anthropometric measurements were taken, according to protocol, by specially trained study midwives on the majority (72%) of term neonates (9, 11). Bi-annual interrater sessions were performed. Crown-heel length (CH-length) and HC were measured to the nearest 0.1 cm.(11) .

Statistical analyses

Differences between pregnancies fulfilling and not fulfilling all but four IG-21-criteria (hereafter referred to as "pregnancies fulfilling the IG-21 criteria") in total and stratified by ethnic origin (three largest groups only), were tested by t-tests for continuous and chi-square tests for categorical variables.

We derived z-scores for each offspring in the Stork-G cohort, using the z-score calculators available at the Intergrowth-21st homepage, based on the IG-21 Fetal Growth Longitudinal Study sample (HC, AC and FL from 14-42 gestational weeks) and from the IG-21 Newborn Cross-Sectional Study sample (weight, HC and CH-length at birth). As standards for EFW during pregnancy and AC at birth are not available from the IG-21 samples, we derived z-scores using a Norwegian reference population and our own study sample to investigate ethnic differences (11, 13). We then plotted mean z-scores against the IG-21 standard mean at each time point, and assessed ethnic differences in mean fetal and neonatal size (mean z-scores) within our cohort, and, restricting the sample to pregnancies lasting more than 37 gestational weeks. SPSS version 22.0 for Windows (SPSS Inc., Chicago, IL, USA) was used. The significance level set to ≤0.050.

Ethical approval

The study protocol was approved by The Regional Committee for Medical and Health Research Ethics for South Eastern Norway (REK 18.10.07, 2007/894), and The Norwegian Data Inspectorate.

Results

Of 823 women included in the Stork-Groruddalen study only 173 (21%) fulfilled the IG-21 criteria, ranging from 16 to 25% for different ethnic groups (Table 1). Approximately 20% in all ethnic groups were excluded for reasons related to their last menstrual period date (Table 1). Other prevalent causes for exclusion were pre-pregnancy smoking, low education, anemia (Hemoglobin < 11 g/dl), age >35 years and obesity (body mass index >30 kg/m^2), but varied by ethnicity. Applying the IG-21-criteria excluded the majority of multiparous Middle East/N-African women, while this was not seen in the two other ethnic groups (Table 2).

Does applying the IG-21 criteria cancel ethnic differences in fetal growth and newborn size?

Compared with the IG-21 fetal growth standard, fetuses in the Stork-G cohort had smaller HC, longer FL, and larger AC during pregnancy (Figure 1a and Supporting Information Table S1). At birth, all three measurements were larger than the IG-21 newborn size standard. Restricting the sample to women fulfilling the IG-21-criteria had little impact on the observed deviations. Compared with the IG-21 newborn size standard, European newborns in particular had larger HC, as 36% of ethnic European neonates had HC above the 90th percentile at birth (Table S1 and Supporting Information Figure S1). This increased to 47% in women fulfilling the IG-21-criteria (Table S1 and Table S1).

Compared with ethnic Europeans, South Asian fetuses were slightly smaller, while Middle East/N-African fetuses had similar EFW in mid pregnancy (Figure 2). However, fetuses in both ethnic minority groups had higher mean FL at this stage of pregnancy (Figure S1 (a)) and grew more slowly until birth (Figure 1) for all three body measurements (Figure S1 (a)). When restricting the sample to women fulfilling the IG-21 criteria, ethnic differences persisted, and for birthweight even increased (Figure 1). The ethnic differences in FL at mid-pregnancy, however, decreased (Figure S1 (d)).

Do the IG-21-criteria identify pregnant women with better health and fewer adverse perinatal outcomes?

Overall, fewer women fulfilling the IG-21 criteria developed gestational diabetes, compared with those not fulfilling the criteria (Table 2). However, this difference was significant only in South Asians, while no such trend was observed in Europeans (Table 3 and 4). Women fulfilling the criteria were slightly less adipose, and had slightly lower serum triglyceride levels in early pregnancy (Table 2). However, this difference was not observed in South Asians (Table 3). Women fulfilling the criteria gained more weight during pregnancy (Table 2 and 3).

More ethnic minority women had severe vitamin D deficiency, depressive symptoms and were physically inactive than ethnic European women (Table 3), with no significant difference between women fulfilling or not fulfilling the IG-21 criteria. Thus, these ethnic differences persisted.

The caesarean section rate was less than half of that in the IG-21 study (Table 2)(4). Overall, we did not observe significant differences between women fulfilling or not fulfilling the

criteria for adverse birth outcomes, such as acute caesarean section, grade 3-4 perineal tear or low Apgar-score (Table 2 and 4).

Discussion

In this multi-ethnic cohort of presumably healthy pregnant women with expected low-risk pregnancies in Norway, only 21% fulfilled the inclusion criteria used for the IG-21 standards. Fetal growth and newborn size deviated notably from the IG-21 standards, in particular for ethnic Europeans. The fact that those selected according to such strict criteria still maintained ethnic differences in fetal growth and newborn size, challenges the concept that one growth standard fits all. Further, the criteria excluded women differently in various ethnic groups, and those fulfilling the IG-21 criteria were not consistently "healthier" or had fewer adverse perinatal outcomes, and importantly, the relations varied by ethnicity.

Several studies have demonstrated that some maternal factors, such as anthropometric parameters and current socioeconomic level, affect fetal growth and perinatal outcomes differently according to ethnic groups and settings, possibly due to different stages of sociodemographic transition (15-18). Additionally, a range of other maternal factors also influence fetal growth (19-22) interacting with each other, and with genetic factors, as part of the adaptive responses made by the fetus to tune its development appropriately to the current conditions (23). Hence, the nature of these relations, and whether they are similar across different populations, ethnicities and cultures, are scantily understood and therefore possibly not well represented in current inclusion and exclusion criteria.

This is in line with our findings that women fulfilling the IG-21 criteria may not necessarily be optimally nourished. I.e. more than half of South Asian women fulfilling the criteria had severe vitamin D deficiency, while this was rare in Europeans. Although the independent effect of vitamin D in the regulation of fetal growth has not been established, there are data supporting that maternal vitamin D status may influence obstetric outcomes and fetal tissue development (24). In contrast, South Asian women fulfilling the IG-21 criteria had less while European women had similar prevalence of gestational diabetes as women not fulfilling the criteria. This could indicate that the IG-21 criteria do not identify risk of gestational diabetes similarly in all ethnic groups and hence that the IG-21 growth standards may not reflects "optimal conditions for fetal growth" in all settings.

In our multi-ethnic cohort, we found ethnic differences in fetal and neonatal size irrespective of whether IG-21 criteria were fulfilled. This is in accordance with findings from the two recent studies aiming to develop new fetal growth charts; one from USA including women from four ethnic groups (White, Hispanic, Black and Asian) (5), and another from the World Health Organization including 10 countries worldwide (6). In both studies, differences based on ethnic origin or country were observed from early pregnancy, and were only partially explained by known maternal factors.

Customizing growth charts according to maternal factors including ethnicity has long been a promising method of refining the diagnostic application (25). The recent insights imply that such customization may need a more differential approach than a crude shift in mean and standard deviation to fit individual diagnostic needs (6). E.g. intrauterine growth may follow different trajectories according to population background and size of fetus (5).

In addition to not being an appropriate standard for expected healthy fetal growth for all, it is a concern that the use of a universal standard in clinical practice could lead to over-diagnosing both large and small for gestational age fetuses. Ethnic Norwegian newborns are among the largest in the world, with a mean birthweight of almost 3700 g at full term (26). E.g. almost half of ethnic European neonates in our cohort had HC >90 percentile. In other populations, specifically in Asia, where mean birth weight is lower, the opposite would be the case, as demonstrated in a Chinese population (27). This could have far-reaching consequences for clinical care. Furthermore, if the criteria aiming to identify women that offer "optimal conditions for fetal growth" do not identify women with better perinatal outcomes, one might question if adopting such standards could do more harm than good, at least for some ethnic groups.

Our study has several strengths, including the population-based design and a multi-ethnic setting. We had information on almost all variables used in the IG-21-study, including maternal data from early pregnancy and offspring growth data from mid gestation until birth, and a wide range of other variables on maternal clinical parameters and obstetric outcomes. However, our study has some limitations. Applying the criteria on our cohort reduced the number of participants substantially and thereby the statistical power. This limited our possibility to address differences in rare, but perhaps more specific perinatal outcomes, such as perinatal mortality.

Fetal FL was substantially longer in our cohort, compared with the mean from the IG-21-standard. This is in accordance with recent studies from France and Greece (28, 29). Furthermore, in our cohort HC was smaller during pregnancy, but larger at birth, compared the IG-21. Whether this discrepancy reflects a systematic difference in measurements performed before or after birth between the studies, is unknown. Anthropometric measurements are known to be sensitive to inter-observer variation. Although all our biometric data were measured according to a strict protocol, comparable to the one used in the Intergrowth-21st studies, such bias cannot be ruled out. However, this would not explain the observed ethnic differences within our cohort.

We are just starting to understand the complexity of fetal growth regulation governed by genetic as well as numerous environmental factors. Our study shows that when using very strict maternal factors to define "healthy" pregnancies, we end up with very few pregnancies. Even under such "optimal" conditions we do not observe less deviation from the growth standards, reduced ethnic differences in fetal growth or newborn size, consistently better maternal health or fewer adverse perinatal outcomes. Hence, we question the validity of using a universal standard.

Acknowledgements

The authors thank the ultrasonographers and study- and general staff at the Child Health Clinics in Stovner, Grorud and Bjerke districts in Oslo and at the delivery- and post-natal wards at Akershus University Hospital and Oslo University Hospital, for help with collecting the data.

Funding

The data collection was funded by The Research Council of Norway, The Norwegian Directorate of Health and collaborative partners in the city of Oslo, Stovner, Grorud and Bjerke administrative districts. The South-Eastern Norway Regional Health Authority has funded the post-doc fellowship for L Sletner. The study sponsors have not been involved in the collection, analysis and interpretation of the data; in the writing of the report; or in the decision to submit the article for publication.

References

1. Papageorghiou AT, Ohuma EO, Altman DG, Todros T, Cheikh Ismail L, Lambert A, et al. International standards for fetal growth based on serial ultrasound measurements: the Fetal Growth Longitudinal Study of the INTERGROWTH-21st Project. Lancet. 2014;384:869-79.

2. Villar J, Cheikh Ismail L, Victora CG, Ohuma EO, Bertino E, Altman DG, et al. International standards for newborn weight, length, and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21st Project. Lancet. 2014;384:857-68.

3. Villar J, Altman DG, Purwar M, Noble JA, Knight HE, Ruyan P, et al. The objectives, design and implementation of the INTERGROWTH-21st Project. BJOG. 2013;120 Suppl 2:9-26, v.

4. Villar J, Papageorghiou AT, Pang R, Ohuma EO, Cheikh Ismail L, Barros FC, et al. The likeness of fetal growth and newborn size across non-isolated populations in the INTERGROWTH-21st Project: the Fetal Growth Longitudinal Study and Newborn Cross-Sectional Study. Lancet Diabetes Endocrinol. 2014;2:781-92.

5. Buck Louis GM, Grewal J, Albert PS, Sciscione A, Wing DA, Grobman WA, et al. Racial/ethnic standards for fetal growth: the NICHD Fetal Growth Studies. Am J Obstet Gynecol. 2015;213:449 e1-e41.

6. Kiserud T, Piaggio G, Carroli G, Widmer M, Carvalho J, Neerup Jensen L, et al. The World Health Organization Fetal Growth Charts: A Multinational Longitudinal Study of Ultrasound Biometric Measurements and Estimated Fetal Weight. PLoS Med. 2017;14:e1002220.

7. Sorbye IK, Daltveit AK, Sundby J, Stoltenberg C, Vangen S. Caesarean section by immigrants' length of residence in Norway: a population-based study. Eur J Public Health. 2015;25:78-84.

8. Sorbye IK, Stoltenberg C, Sundby J, Daltveit AK, Vangen S. Stillbirth and infant death among generations of Pakistani immigrant descent: a population-based study. Acta Obstet Gynecol Scand. 2014;93:168-74.

9. Jenum AK, Sletner L, Voldner N, Vangen S, Morkrid K, Andersen LF, et al. The STORK Groruddalen research programme: A population-based cohort study of gestational diabetes, physical activity, and obesity in pregnancy in a multiethnic population. Rationale, methods, study population, and participation rates. Scand J Public Health. 2010;38:60-70.

10. Roseman F, Knight HE, Giuliani F, Lloyd S, Di Nicola P, Laister A, et al. Implementation of the INTERGROWTH-21st Project in the UK. BJOG. 2013;120 Suppl 2:117-22, v.

11. Sletner L, Nakstad B, Yajnik CS, Morkrid K, Vangen S, Vardal MH, et al. Ethnic differences in neonatal body composition in a multi-ethnic population and the impact of parental factors: a population-based cohort study. PLoS One. 2013;8:e73058.
12. Jenum AK, Morkrid K, Sletner L, Vangen S, Torper JL, Nakstad B, et al. Impact of ethnicity on gestational diabetes identified with the WHO and the modified International Association of Diabetes and Pregnancy Study Groups criteria: a population-based cohort study. Eur J Endocrinol. 2012;166:317-24.
13. Sletner L, Rasmussen S, Jenum AK, Nakstad B, Jensen OH, Vangen S. Ethnic differences in fetal size and growth in a multi-ethnic population. Early Hum Dev. 2015;91:547-54.
14. Combs CA, Jaekle RK, Rosenn B, Pope M, Miodovnik M, Siddiqi TA. Sonographic estimation of fetal weight based on a model of fetal volume. Obstet Gynecol. 1993;82:365-70.
15. Blumenshine P, Egerter S, Barclay CJ, Cubbin C, Braveman PA. Socioeconomic disparities in adverse birth outcomes: a systematic review. Am J Prev Med. 2010;39:263-72.
16. Makgoba M, Savvidou MD, Steer PJ. The effect of maternal characteristics and gestational diabetes on birthweight. BJOG. 2012;119:1091-7.
17. Lampl M, Thompson AL, Frongillo EA. Sex differences in the relationships among weight gain, subcutaneous skinfold tissue and saltatory length growth spurts in infancy. Pediatr Res. 2005;58:1238-42.
18. Leary S, Fall C, Osmond C, Lovel H, Campbell D, Eriksson J, et al. Geographical variation in relationships between parental body size and offspring phenotype at birth. Acta Obstet Gynecol Scand. 2006;85:1066-79.
19. Bjornerem A, Johnsen SL, Nguyen TV, Kiserud T, Seeman E. The shifting trajectory of growth in femur length during gestation. J Bone Miner Res. 2010;25:1029-33.
20. Bottomley C, Daemen A, Mukri F, Papageorghiou AT, Kirk E, Pexsters A, et al. Assessing first trimester growth: the influence of ethnic background and maternal age. Hum Reprod. 2009;24:284-90.
21. Kinare AS, Chinchwadkar MC, Natekar AS, Coyaji KJ, Wills AK, Joglekar CV, et al. Patterns of fetal growth in a rural Indian cohort and comparison with a Western European population: data from the Pune maternal nutrition study. J Ultrasound Med. 2010;29:215-23.
22. Lampl M, Gotsch F, Kusanovic JP, Gomez R, Nien JK, Frongillo EA, et al. Sex differences in fetal growth responses to maternal height and weight. Am J Hum Biol. 2010;22:431-43.
23. Hanson MA, Low FM, Gluckman PD. Epigenetic epidemiology: the rebirth of soft inheritance. Ann Nutr Metab. 2011;58 Suppl 2:8-15.
24. Moon RJ, Harvey NC, Cooper C. ENDOCRINOLOGY IN PREGNANCY: Influence of maternal vitamin D status on obstetric outcomes and the fetal skeleton. Eur J Endocrinol. 2015;173:R69-83.

25. Gardosi J, Mongelli M, Wilcox M, Chang A. An adjustable fetal weight standard. Ultrasound Obstet Gynecol. 1995;6:168-74.

26. Skjaerven R, Gjessing HK, Bakketeig LS. Birthweight by gestational age in Norway. Acta Obstet Gynecol Scand. 2000;79:440-9.

27. Cheng Y, Leung TY, Lao T, Chan YM, Sahota DS. Impact of replacing Chinese ethnicity-specific fetal biometry charts with the INTERGROWTH-21(st) standard. BJOG. 2016;123 Suppl 3:48-55.

28. Sotiriadis A, Eleftheriades M, Chatzinikolaou F, Hassiakos D, Chrousos GP, Pervanidou P. National curves of foetal growth in singleton foetuses of Greek origin. Eur J Clin Invest. 2016;46:425-33.

29. Stirnemann JJ, Fries N, Bessis R, Fontanges M, Mangione R, Salomon LJ. Implementing the INTERGROWTH-21st fetal growth standards in France: a 'flash study' of the College Francais d'Echographie Foetale (CFEF). Ultrasound Obstet Gynecol. 2017;49:487-92.

Supporting Information legends

Figure S1. Ethnic differences in fetal head circumference (HC), length, (measured as femur length (FL) during pregnancy and crown-heel length at birth) and abdominal circumference (AC) in ethnic South Asians (left panel) and Middle East/N-Africans (right panel), compared with ethnic Europeans (zero-line in both figures). Figure a and c present differences in total Stork-Groruddalen cohort and b and d in those fulfilling the Intergrowth-21st inclusion criteria. Results are presented as differences in mean z-score (95% CI). Numbers includes only live born, singleton neonates, with valid birth data and at least one ultrasound measurement, born at more than 37 gestational weeks (179 South Asian, 158 Middle East/N-African and 346 ethnic European in the total cohort and 44, 25 and 78, respectively, in pregnancies not fulfilling the IG-21 criteria).

Table S1. Mean fetal and neonatal z-scores, and the proportion classified large or small for gestational age, according to the standards developed by the Intergrowth-21st consortium, stratified by ethnic origin.

Figure legends

Figure 1. Mean z-scores of fetal/neonatal measures in the Stork Groruddalen cohort at 24, 32 and 37 weeks' gestation and at birth, according to the new Intergrowth-21st fetal growth and newborn size standards); (a) in the total cohort and (b) in pregnancies fulfilling the Intergrowth-21st (IG-21) inclusion criteria. Measures are mean z-scores (95% CI) of head circumference (HC, circles), length (measured as femur length (FL) in pregnancy and crown-heel length at birth, squares), abdominal circumference (AC, white triangles) and birth weight (black triangles). . The zero-line represents the mean in the Intergrowth-cohorts. Numbers include only live born, singleton neonates with valid birth data and at least one ultrasound measurement (164 pregnancies fulfilling, and 608 pregnancies not fulfilling the IG-21 criteria).

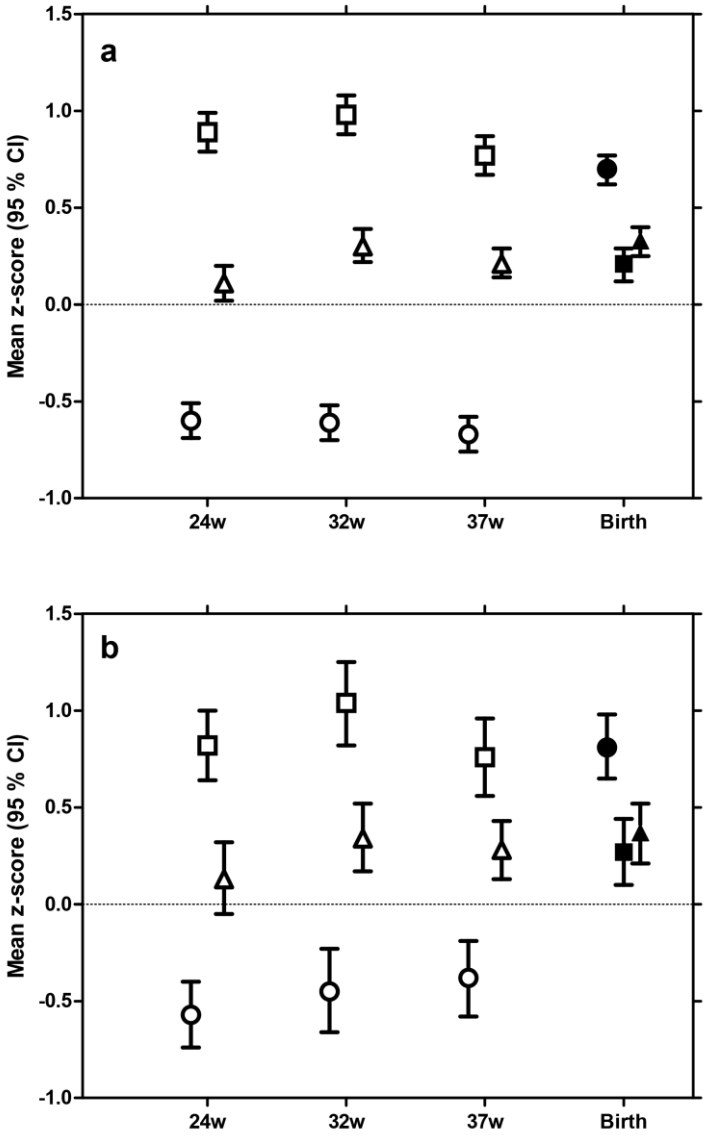

Figure 2. Ethnic differences in estimated fetal weight (EFW) and birth weight, in ethnic South Asians (white circles) and Middle East/N-Africans (black squares), compared with ethnic Europeans (zero-line) in the Stork-Groruddalen cohort; (a) in the total cohort and (b) in pregnancies fulfilling the Intergrowth-21st (IG-21) inclusion criteria. Results are presented as differences in mean z-score (95% CI). Numbers includes only live born, singleton neonates, with valid birth data and at least one ultrasound measurement, born at more than 37 gestational weeks (179 South Asian, 158 Middle East/N-African and 346 ethnic European in the total cohort and 44, 25 and 78 respectively in pregnancies not fulfilling the IG-21 criteria).

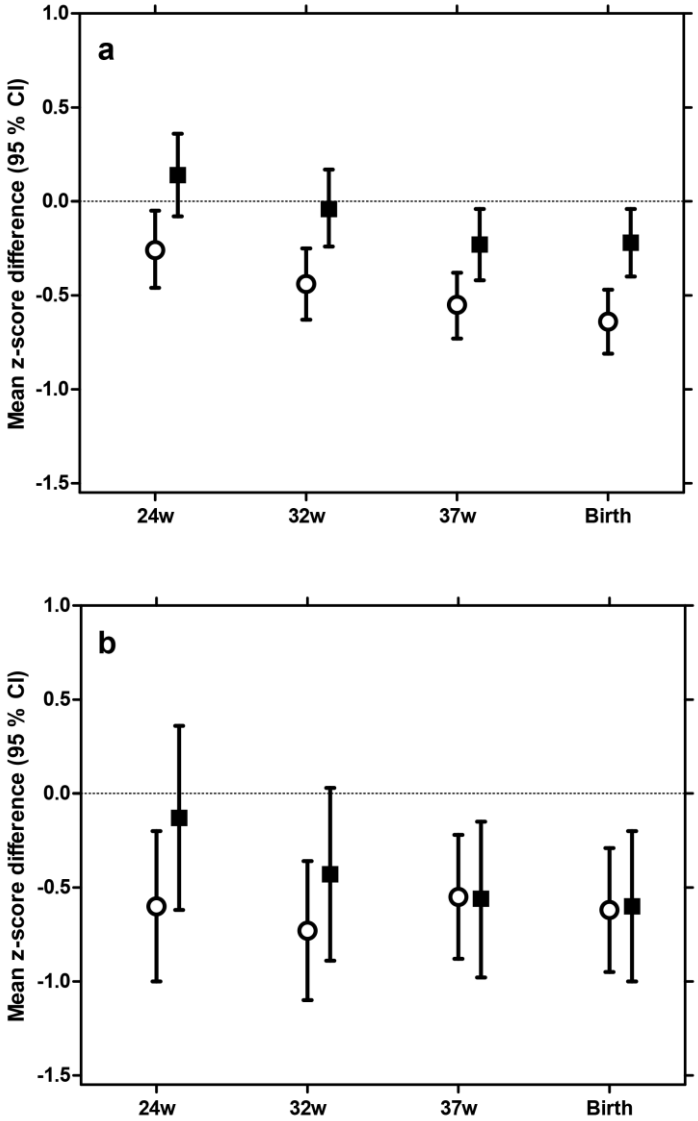

Table 1. Proportion of women in the Stork-Groruddalen cohort (n (%)) excluded if applying the Intergrowth-21st inclusion criteria.

	Overall	Europe[a]	South Asia[b]	East Asia[c]	Middle East/N-Afr[d]	Others[e]
	n=823	n=379	n=200	n=43	n=172	n=29
Demographics/anthropometrics						
Age <18 years[f]	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Age ≥35 years	127 (15)	66 (17)	18 (9)	11 (26)	28 (16)	4 (14)
Height<153 cm	41 (5)	3 (0.8)	16 (8.0)	10 (23)	8 (5)	4 (14)
Pre-pregnant BMI <18.5 kg/m^2	54 (7)	20 (5)	20 (10)	4 (9)	9 (5)	1 (3)
Pre-pregnant BMI ≥30 kg/m^2	113 (14)	47 (12)	19 (10)	2 (5)	37 (22)	8 (28)
Low education (primary or less)	133 (16)	15 (4)	35 (18)	8 (19)	72 (42)	3 (10)
Last menstrual period dating issues						
Any uncertainty related to LMP dating[g]	175 (21)	78 (21)	44 (22)	13 (30)	26 (21)	4 (14)
No recall of LMP date	27 (3)	18 (5)	3 (2)	0 (0)	6 (4)	0 (0)
Hormonal contraceptive use last 2 cycles	74 (9)	47 (12)	12 (6)	2 (5)	12 (7)	1 (3)
IVF/hormonal induction	13 (2)	10 (3)	1 (0.5)	0 (0)	2 (1)	0 (0)
Obstetric issues/history						
Twin pregnancy	11 (1)	4 (1)	2 (1)	1 (2)	3 (2)	1 (3)
Two or more previous miscarriages	52 (6)	24 (6)	15 (8)	3 (7)	7 (4)	3 (10)
Previous preterm delivery	22 (3)	5 (1)	13 (7)	0 (0)	4 (2)	0 (0)
Previous birthweight <2500 g	31 (4)	4 (1)	13 (7)	1 (2)	12 (7)	1 (3)
Previous birthweight >4500 g	11 (1)	3 (0.8)	0 (0)	0 (0)	8 (5)	0 (0)
Previous stillbirth/perinatal death	8 (1)	2 (0.5)	3 (2)	0 (0)	2 (1)	1 (3)
Previous preeclampsia	-	-	-	-	-	-
Medical issues at inclusion						
Medical history/medication[h]	59 (7)	26 (7)	18 (9)	1 (2)	14 (8)	0 (0)
Any smoking last 3 mo prior to pregnancy	143 (18)	117 (31)	4 (2)	4 (9)	16 (9)	2 (7)
Heavy alcohol use in pregnancy	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Positive urine-analysis	No info	-	-	-	-	-
Systolic blood pressure>140 mmHg	2 (0.2)	2 (0.5)	0 (0)	0 (0)	0 (0)	0 (0)
Diastolic blood pressure>90 mmHg	4 (0.5)	4 (1)	0 (0)	0 (0)	0 (0)	0 (0)
Anemia (Hb<11 g/dl)	112 (14)	30 (8)	37 (19)	7 (16)	33 (19)	5 (17)
Treatment for anemia	46 (6)	13 (3)	18 (9)	4 (9)	11 (6)	0 (0)
Sexually transmitted disease (Lues, HIV)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Atypical red cell alloantbodies	-	-	-	-	-	-
High risk occupation[i]	-	-	-	-	-	-
Total remaining after excl. on any criterion	173 (21)	81 (22)	49 (25)	8 (19)	28 (16)	5 (17)

[a] Europe (mainly Scandinavia).

[b] South Asia (Pakistan, Sri Lanka, India).

[c] East Asia (mainly Vietnam, the Philippines and Thailand).

[d] Middle-East/North-Africa (mainly Iraq, Turkey, Morocco, Somalia) and

[e] Others (other Sub-Sahara African and S/C-American countries).

[f] Not specified exclusion criteria in Stork-Groruddalen study, but one woman was excluded partly for this reason, social considerations.

[g] May be related to somewhat irregular cycles, contraceptive use last cycles, IVF-pregnancy, bleedings in early pregnancy, uncertain recall of exact date etc.

[h] Conditions during the last three months before pregnancy/during pregnancy with need of regular medication or treatment. Isolated pollen-allergy, eczema and hyperemesis were not excluded. The most common conditions were asthma, hypothyroidism, psychiatric disorders (mainly depression), urinary tract infections, hypertension and hormonal medication related to fertilization treatment.

[i] In an occupation with risk of exposure to chemicals or toxic substances, or very physically demanding activity. Such occupational activity during pregnancy is regulated by law in Norway.

LMP, Last menstrual period.

Table 2. Maternal characteristics and pregnancy and birth complications in women participating in the Stork-Groruddalen cohort (left panel), either fulfilling or not fulfilling all Intergrowth-21st (IG-21) inclusion criteria. The right panel represents comparable numbers from the two Intergrowth-21st study samples; the Fetal Growth Longitudinal Study (FGLS) and the Newborn Cross-Sectional Study Prescriptive Subpopulation (NCSS-PS).

	Stork-G			Intergrowth-21st	
	Fulfilling IG-21 criteria n=173	Not fulfilling IG-21 criteria n=650	p	FGLS-cohort n=4321	NCSS-PS-cohort n=20486
Maternal characteristics					
Incl. in Intergrowth-21st criteria					
Maternal age (years)	29.4 (3.6)	30.0 (5.2)	**0.02**	28.4 (3.9)	28.0 (4.0)
Years of formal education (years)	14.5 (2.4)	12.7 (3.7)	**<0.001**	15.0 (2.8)	14.2 (3.0)
Height (cm)	164.5 (6.1)	163.4 (6.8)	0.06	162.2 (5.8)	161.8 (5.6)
Pre-pregnant BMI (kg/m^2)	23.3 (2.6)	25.0 (5.2)	**<0.001**	23.3 (3.0)	23.4 (2.9)
Hemoglobin level (g/dl)	12.3 (0.7)	12.0 (1.0)	**<0.001**	12.5 (1.1)	12.3 (1.2)
Not incl. in Intergrowth-21st criteria					
Parity			**0.03**		
Para 0	92 (53)	289 (45)		2955 (68%)	12996 (63%)
Para 1	59 (34)	221 (34)		÷	÷
Multipara	22 (13)	140 (22)		÷	÷
Married or cohabiting	169 (98%)	607 (93%)	**0.03**	4204 (97%)	19877 (97%)
Weight (kg)	64.6 (9.2)	68.8 (15.3)	**<0.001**	61.3 (9.1)	61.3 (8.6)
Sum of skinfolds (mm)	68.4 (16.7)	73.3 (20.6)	**0.005**	÷	÷
Triglycerides (mmol/l)	1.20 (0.41)	1.35 (0.58)	**0.002**	÷	÷
Fasting glucose (mmol/l)	4.41 (0.43)	4.43 (0.42)	0.6	÷	÷
Severe vitamin D-deficiency[a]	36 (21)	142 (22)	0.8	÷	÷
Depressive sympt. (EPDS-score>10)[b]	18 (12)	78 (13)	0.6	÷	÷
No regular physical activity[c]	125 (72%)	490 (75%)	0.5	÷	÷
1. -2. trimester weight gain (kg)	9.0 (4.3)	8.5 (4.8)	0.3	÷	÷
3.trimester weight gain (kg)	6.0 (3.8)	5.1 (4.1)	**0.02**	÷	÷
Pregnancy complications					
Gestational diabetes mellitus[d]	13 (8%)	86 (14%)	**0.03**	÷	÷
Hypertensive disorder in pregnancy[e]	9 (5%)	37 (6%)	0.8	÷	÷
Preeclampsia[e]	3 (2%)	8 (1%)	—	31 (<1%)	256 (1%)
Birth outcomes					
Spontaneous initiation of labour[f]	136 (81%)	468 (74%)	0.1	2868 (66%)	13470 (66%)
Cesarean section[f]	26 (15%)	112 (17%)	0.5	1541 (36%)	7452 (36%)
Preterm birth (<37 gestational weeks)[f]	9 (6%)	44 (7%)	0.5	195 (5%)	1136 (6%)
Acute cesarean section[g]	19 (12%)	64 (11%)	0.7	÷	÷
Grade 3-4 perineal tear[g]	4 (3%)	19 (3%)	—	÷	÷
Apgar score < 7 at 5 min[g]	0 (0%)	8 (1%)	—	÷	÷
Composite three complications[g, h]	23 (16%)	78 (15%)	0.7	÷	÷
Gender boy[g]	73 (46%)	300 (52%)	0.2	2149 (50%)	10482 (51%)

Gestational age at birth (days)[g]	282 (9)	281 (9)	0.2	÷	÷
Birthweight (g)[g]	3511 (472)	3470 (505)	0.4	3.3 kg (0.4)	3.3 kg (0.5)
Head circumference[g]	35.1 (1.4)	34.8 (1.4)	**0.03**	33.9 (1.3)	33.9 (1.3)
Crown-heel-length[g]	50.1 (2.0)	50.0 (2.1)	0.2	49.4 (1.9)	49.3 (1.8)
Low birth weight (<2500 g)[g]	2 (1%)	13 (2%)	—	128 (3)	651 (3%)

Numbers are n (%) or mean (SD). Differences between women in the Stork-Groruddalen (Stork-G) fulfilling, compared women not fulfilling the IG-21 criteria, were assessed using chi-square tests for categorical and t-tests for continuous variables. Significant differences (p-values <0.05) are written in bold. All variables, except depressive symptoms (28 gestational weeks) and weight gain (28 gestational weeks (missing in 7% of women) and post-partum (missing in 24% of women) were assessed at inclusion (mean: 15 gestational weeks). There are few (<1-2%) missing values for variables reported at inclusion, most participants have no missing variables.

[a] Maternal 25 OH vit D < 25 nmol/L.

[b]Measured at 28 gestational weeks, by the Edinburgh Postnatal Depression Scale.

[c] Regular physical activity defined as either 1) moderately intensive activity for 30 min for ≥5 days/week, 2) moderately intensive activity for 2.5 h/week over ≥3 days, 3) vigorous-intensity activity for ≥20 min three times per week, or 4) activity of both moderate and vigorous intensity in 1 or 2.

[d] Gestational diabetes defined by WHO 1999 criteria; Fasting plasma glucose ≥7 or 2-hour plasma glucose ≥7.8 mmol/l by oral glucose tolerance screening test at 28 gestational weeks or diagnosed at any time during pregnancy.

[e] Hypertensive disorder in pregnancy defined as systolic blood pressure > 140 mmHg and/or diastolic blood pressure > 90 mmHg occurring after 20 gestational weeks , whether or not accompanied by new-onset proteinuria. Numbers presented are from hospital records. Preeclampsia represents women diagnosed with hypertensive disorder from before 34 gestational weeks and/or with important clinical significance, as defined by hospital records. Numbers are likely not fully comparable with the numbers from Intergrowth-21st.

[f] As in the Intergrowth-21st study, numbers are calculated from live born, singleton neonates with birth data. In total there were 13 abortions, 8 stillbirths/perinatal deaths, 11 twin-pairs and 8 were lost to follow-up with no birth data.

[g]As in the Intergrowth-21st study numbers are calculated from live born, singleton neonates, born at more than 37 gestational weeks, with birth data (84% and 82% of the total number of births in the respective group).

[h]Composite score of three complication; acute cesarean section, grade 3-4 perineal rupture and Apgar score < 7 at 5 min.

BMI, body mass index; EPDS, Edinburgh Postnatal Depression Scale.

Table 3. Maternal characteristics in women fulfilling, compared with women not fulfilling the Intergrowth-21st criteria (IG-21 criteria), stratified by ethnicity.

	Europe			South Asia			Middle East/N-Africa		
	Fulfilling IG-21 criteria	Not fulfilling IG-21 criteria	p	Fulfilling IG-21 criteria	Not fulfilling IG-21 criteria	p	Fulfilling IG-21 criteria	Not fulfilling IG-21 criteria	p
	n=83	n=296		n=49	n=151		n=28	n=144	
Characteristics included in IG-21 criteria									
Maternal age (years)	30.2 (2.9)	30.9 (4.8)	0.2	29.1 (3.9)	28.5 (4.7)	0.5	27.2 (4.1)	29.6 (5.7)	**0.03**
Years of formal education (years)	15.4 (2.3)	14.4 (2.7)	**0.004**	13.6 (2.3)	12.4 (2.8)	**0.008**	9.3 (2.0)	13.6 (2.3)	**<0.001**
Height (cm)	167.3 (5.8)	167.1 (5.7)	0.9	161.6 (5.1)	159.5 (5.8)	**0.03**	163.5 (5.3)	161.6 (5.5)	0.1
Pre-pregnant BMI (kg/m^2)	23.2 (2.6)	24.9 (5.1)	**0.004**	23.3 (2.7)	23.9 (4.6)	0.4	23.4 (2.7)	26.5 (5.7)	**0.005**
Hemoglobin level (g/dl)	12.4 (0.7)	12.3 (0.9)	0.3	12.2 (0.7)	11.6 (1.0)	**<0.001**	12.1 (0.7)	11.8 (1.0)	0.1
Characteristics not included in IG-21 criteria									
Parity			0.3			0.7			**0.004**
Para 0	49 (59%)	155 (52%)		23 (47%)	61 (40%)		15 (54%)	47 (33%)	
Para 1	29 (35%)	108 (37%)		14 (29%)	51 (34%)		11 (39%)	40 (28%)	
Multipara	5 (6%)	33 (11%)		12 (25%)	39 (26%)		2 (7%)	57 (40%)	
Married or cohabiting	83 (100%)	278 (94%)	**0.02**	48 (98%)	149 (99%)	0.7	26 (93%)	133 (92%)	0.9
Weight (kg)	66.4 (8.6)	71.5 (14.5)	**0.002**	62.3 (8.8)	62.6 (12.5)	0.9	64.3 (9.8)	71.9 (16.7)	**0.02**
Sum of skinfolds (mm)	65.9 (16.0)	70.7 (20.5)	0.06	73.0 (17.9)	74.3 (19.5)	0.7	67.4 (14.5)	77.7 (21.5)	**0.02**
Triglycerides, median (IQR) (mmol/l)	1.06 (0.44)	1.13 (0.58)	**0.02**	1.26 (0.76)	1.32 (0.65)	0.6	1.04 (0.50)	1.27 (0.66)	**0.03**
Fasting glucose (mmol/l)	4.38 (0.42)	4.41 (0.38)	0.7	4.51 (0.35)	4.46 (0.42)	0.5	4.31 (0.32)	4.48 (0.49)	0.07
Severe vitamin D-deficiency[a]	0 (0)	6 (2)	0.2	27 (55)	63 (42)	0.1	8 (29)	55 (39)	0.3
Depressive sympt. (EPDS-score>10)[b]	10 (13)	20 (8)	0.1	5 (11)	28 (20)	0.1	2 (10)	20 (16)	0.5
No regular physical activity[c]	46 (55%)	185 (62%)	0.4	45 (92%)	132 (87%)	0.4	23 (82%)	127 (88%)	0.4

1. -2. trimester weight gain (kg)	8.8 (4.0)	8.9 (4.7)	0.7	9.0 (4.5)	8.0 (4.6)	0.2	10.4 (4.4)	8.5 (5.2)	0.1
3.trimester weight gain (kg)	6.2 (3.6)	5.5 (3.8)	0.2	5.9 (4.0)	4.9 (3.9)	0.2	5.0 (4.3)	4.2 (5.2)	0.2

Numbers are n (%) or mean (SD). Differences between women fulfilling, compared women not fulfilling the IG-21 criteria, were assessed using chi-square tests for categorical and t-tests for continuous variables. Significant differences (p-values <0.05) are written in bold. All variables, except depressive symptoms (28 gestational weeks and weight gain (28 gestational weeks (missing in 7% of women) and post-partum (missing in 24% of women)), were assessed at inclusion (at mean 15 gestational weeks). There are few (<1-2%) missing values for variables reported at inclusion, most participants have no missing variables.

[a]25 OH vitamin D < 25 nmol/L.

[b]Measured at 28 gestational weeks, by the Edinburgh Postnatal Depression Scale.

[c]Moderately intensive activity for 30 min for ≥5 days/week, moderately intensive activity for 2.5 h/week over ≥3 days, vigorous-intensity activity for ≥20 min three times per week, or activity of both moderate and vigorous intensity.

EPDS, Edinburgh Postnatal Depression Scale.

Table 4. Pregnancy complications and perinatal outcomes in women fulfilling, compared with not fulfilling the Intergrowth-21st criteria (IG-21 criteria), stratified by ethnicity.

	Europe			South Asia			Middle East/N-Africa		
	Fulfilling IG-21 criteria n=83	Not fulfilling IG-21 criteria n=296		Fulfilling IG-21 criteria n=49	Not fulfilling IG-21 criteria n=151		Fulfilling IG-21 criteria n=28	Not fulfilling IG-21 criterai n=144	
Pregnancy complications									
Gestational diabetes mellitus[a]	8 (10%)	33 (12%)	0.6	3 (6%)	25 (18%)	**0.05**	1 (4%)	23 (18%)	0.09
Hypertensive disorder in pregnancy[b]	19 (6%)	7 (8%)	0.5	0 (0%)	11 (7%)	0.06	6 (4%)	2 (7%)	0.5
Birth outcomes									
Spontaneous initiation of labour	65 (80%)	216 (75%)	0.5	44 (90%)	109 (75%)	0.08	19 (76%)	105 (76%)	0.4
Cesarean section	15 (18%)	55 (19%)	0.9	5 (10%)	27 (17%)	0.2	3 (11%)	21 (15%)	0.6
Preterm birth (< 37gestational weeks)	3 (4%)	21 (7%)	0.3	4 (8%)	10 (7%)	0.7	2 (8%)	4 (3%)	—
Birth complications[c]									
Acute cesarean section[c]	12 (15%)	26 (10%)	0.2	4 (9%)	15 (11%)	0.7	3 (12%)	17 (13%)	0.9
Grade 3-4 perineal tear[c]	0 (0%)	10 (4%)	—	4 (9%)	3 (2%)	—	0 (0%)	4 (3%)	—
Apgar score < 7 at 5 min[c]	0 (0%)	4 (2%)	—	0 (0%)	1 (1%)	—	0 (0%)	2 (2%)	—
Composite three complications[c, d]	12 (15%)	38 (14%)	0.8	8 (17%)	22 (13%)	0.4	3 (12%)	22 (17%)	0.6
Gender boy[c]	43 (55%)	137 (51%)	0.5	24 (55%)	68 (50%)	0.6	13 (52%)	52 (39%)	0.2
Gestational age at birth (days)[c]	283 (10)	282 (9)	0.8	282 (8)	279 (10)	**0.02**	280 (9)	280 (10)	0.9
Birth weight (g)[c]	3669 (478)	3610 (471)	0.3	3398 (399)	3247 (467)	0.06	3350 (435)	3480 (525)	0.2
Head circumference[c]	35.5 (1.4)	35.1 (1.3)	**0.03**	34.8 (1.2)	34.3 (1.3)	**0.03**	34.5 (1.1)	34.9 (1.5)	0.3
Crown-heel-length[c]	50.6 (2.0)	50.3 (2.1)	0.3	49.8 (1.8)	49.5 (2.1)	0.5	49.6 (2.4)	49.7 (2.2)	0.8
Low birth weight (<2500 g)[c]	1 (1%)	2 (1%)	—	0 (0%)	6 (4%)	—	0 (0%)	2 (2%)	—

Numbers are n (%) or mean (SD). Differences between women fulfilling, compared women not fulfilling the IG-21 criteria, were assessed using chi-square tests for categorical and t-tests for continuous variables. Significant differences (p-values <0.05) are written in bold. Missing p-values is due to insufficient numbers of observations to perform chi-square test.

[a]By WHO 1999 criteria; Fasting plasma glucose ≥ 7 mmol/l or 2-hour plasma glucose ≥ 7.8 mmol/l).

[b]Systolic blood pressure > 140 mmHg and/or diastolic blood pressure > 90 mmHg occurring from early pregnancy, or occurring after 20 weeks gestation, accompanied by new-onset proteinuria. Numbers are from hospital records.

[c]Numbers are calculated from live born, singleton neonates, born at more than 37 gestational weeks, with birth data (as in the Intergrowth-21st study). In total there were 13 abortions, 8 stillbirths/perinatal deaths, 11 twin-pairs and 8 were lost to follow-up with no birth data.

[d]Composite score of three complication; acute cesarean section, grade 3-4 perineal rupture and Apgar score < 7 at 5 min.