Most of the Advantage of Customised Birthweight Standards Is a Threshold Effect: A Simulation of Growth-Restriction Misclassification

Dr. PriyadharshiniP.

Assistant Professor, Department of Community Medicine, KMCH Institute of Health Sciences and Research, Tamil Nadu, India.

Corresponding author: drpriya.cm@gmail.com

Received: 01 January 2026; Revised: 11 March 2026; Accepted: 17 March 2026; Published: 03 May 2026

Abstract

Background: Small for gestational age is defined by a centile threshold, and which babies fall below it depends on the standard used. Customised standards adjust for maternal height, weight, parity and fetal sex on the reasoning that a baby small for its own growth potential is the one at risk, and are reported to identify growth restriction better than population standards. Whether that advantage reflects better discrimination or simply flagging more babies has not been separated.

Objective: To quantify how population and customised birthweight standards classify the same births, to compare their detection of true pathological growth restriction at equal and at nominal thresholds, and to characterise the babies on whom they disagree.

Methods: Five hundred thousand births were simulated. Each fetus was assigned a growth potential determined by maternal height, booking weight, parity and fetal sex; 6 % were assigned pathological restriction reducing birthweight by 12 to 35 %; and physiological variation was added. Population centiles were computed from fetal sex alone and customised centiles from the individual growth potential. Detection of true restriction was compared at nominal centile thresholds of 3, 5 and 10, and again at thresholds matched so that both standards flagged the same proportion of births.

Results: At the nominal 10th centile the customised standard appeared substantially better, detecting 81.8 % of truly restricted babies against 67.6 % — but it flagged 12.8 % of all births against 7.5 %. When both standards were set to flag the same 7.53 % of births, detection was 71.5 % against 67.6 % and positive predictive value 56.7 % against 53.5 %. Most of the apparent advantage was therefore a threshold effect, with a genuine but modest discrimination gain of about four percentage points remaining. The standards disagreed about 6.0 % of all births: 5.64 % were flagged only by the customised standard, of whom 15.4 % were truly restricted, and 0.36 % only by the population standard, of whom 5.1 % were. The discordant groups differed systematically — population-only babies were born to mothers of mean height 152.0 cm and weight 53.1 kg, 17.8 % parous, against 165.4 cm, 69.3 kg and 64.4 % parous for customised-only babies.

Conclusions: Comparisons of birthweight standards at the same nominal centile are not like-for-like and overstate the customised advantage. The substantive difference is in which babies are selected rather than how many restricted babies are found: population standards preferentially flag the constitutionally small babies of small mothers, and miss restricted babies of large mothers.

Keywords: Small for Gestational Age, Fetal Growth Restriction, Customised Centiles, Birthweight Standards, Misclassification, Simulation.

1. INTRODUCTION

Small for gestational age is a definition, not a diagnosis. A baby is small for gestational age if its birthweight falls below a centile threshold on a chosen standard, and the label is used as a proxy for the condition that actually matters — fetal growth restriction, in which a fetus has failed to achieve its own growth potential and is at increased risk of stillbirth, hypoxic injury and neonatal morbidity.

The proxy is imperfect in two directions that are well recognised and rarely quantified together. A baby may be small because its parents are small, having grown exactly as it should; labelling it growth restricted

generates surveillance, anxiety and sometimes intervention for no benefit. And a baby with a large growth potential may be restricted while remaining above the threshold, because it started from a higher expected weight; such a baby is missed entirely.

Customised standards address this by computing an expected weight for each pregnancy from maternal height, booking weight, parity and fetal sex, and expressing the actual weight relative to it. The approach is intuitive, has been associated with better identification of babies at risk of adverse outcome, and is used in several national programmes. It is also contested, with critics arguing that the apparent advantage arises from the way comparisons are constructed rather than from better discrimination.

That dispute is difficult to settle with observational data because the true condition — whether a given fetus failed to achieve its potential — is unobservable. In simulation it is known by construction, which makes the comparison tractable. This study asks three questions: how much of the reported advantage of customised standards survives when both standards flag the same proportion of births; how much of the apparent difference is a threshold effect; and which babies the two standards disagree about.

2. METHODS

2.1 Generating the population

Five hundred thousand term births were simulated. Each pregnancy was assigned maternal height, booking weight and parity, and each fetus a sex. An optimal birthweight — the weight the fetus would achieve in the absence of pathology — was computed as a linear function of these variables.

Optimal weight (g) = 3400 + 18(height − 163) + 7(weight − 66) + 110 × parous + 120 × male (1)

A randomly selected 6 % of fetuses were assigned pathological growth restriction, reducing birthweight by a uniformly distributed 12 to 35 %. Physiological variation was then applied multiplicatively with a log-normal term of coefficient of variation 10.5 %. The resulting distribution of birthweight therefore contains constitutionally small babies, constitutionally large babies, restricted babies of every size, and the overlaps between them, with the truth known for each.

Table 1. Model parameters

Parameter	Distribution or value	Role
Maternal height	Normal, mean 163 cm, SD 6.5, truncated 145–185	Determinant of growth potential
Maternal booking weight	Normal, mean 66 kg, SD 12, truncated 42–120	Determinant of growth potential
Parity	55 % parous	Determinant of growth potential
Fetal sex	51.2 % male	Determinant of growth potential; used by both standards
Pathological growth restriction	6 % of fetuses	The condition to be detected
Weight deficit if restricted	Uniform, 12–35 %	Severity of restriction
Physiological variation	Log-normal, coefficient of variation 10.5 %	Irreducible scatter
Births simulated	500,000	Monte Carlo error negligible

Note. Parameter values are representative assumptions chosen to produce a realistic birthweight distribution, not estimates from any cohort. The coefficients in equation (1) are illustrative of the direction and rough magnitude of the maternal determinants of birthweight and are not taken from a published customisation algorithm.

2.2 The two standards

Population standard. The birthweight centile is computed from a distribution defined by fetal sex alone, with gestation held constant at term. This represents a conventional population reference, which adjusts for sex and gestation and for nothing about the mother.

Customised standard. The centile is computed by referring the actual weight to the individual optimal weight from equation (1), with the same proportional scatter. This represents an idealised customisation: it knows the true growth potential exactly.

The second assumption is deliberately generous and should be stated plainly. A real customisation algorithm estimates growth potential from a regression with substantial residual error, so its performance is necessarily worse than modelled here. Any advantage the customised standard fails to show under this assumption it will not show in practice, which is what makes the comparison informative.

2.3 Analyses

Detection of true pathological restriction was computed at nominal centile thresholds of 3, 5 and 10 for both standards. Because the two standards flag different proportions of births at the same nominal centile, the comparison was repeated with the customised threshold set so that both flagged an identical proportion. Discordance was tabulated at the 10th centile, and the maternal and fetal characteristics of each discordant group were described. All computation was performed in Python with a fixed random seed and reproduces exactly.

3. RESULTS

3.1 The comparison as usually made

Table 2. Detection at nominal centile thresholds

Standard and threshold	Births flagged (%)	Restriction detected (%)	Positive predictive value (%)	Restricted babies missed (%)
Population, below 3rd	2.8	40.7	87.1	59.3
Customised, below 3rd	5.5	63.9	68.9	36.1
Population, below 5th	4.1	51.8	75.8	48.2
Customised, below 5th	7.6	71.8	56.1	28.2
Population, below 10th	7.5	67.6	53.5	32.4
Customised, below 10th	12.8	81.8	38.1	18.2

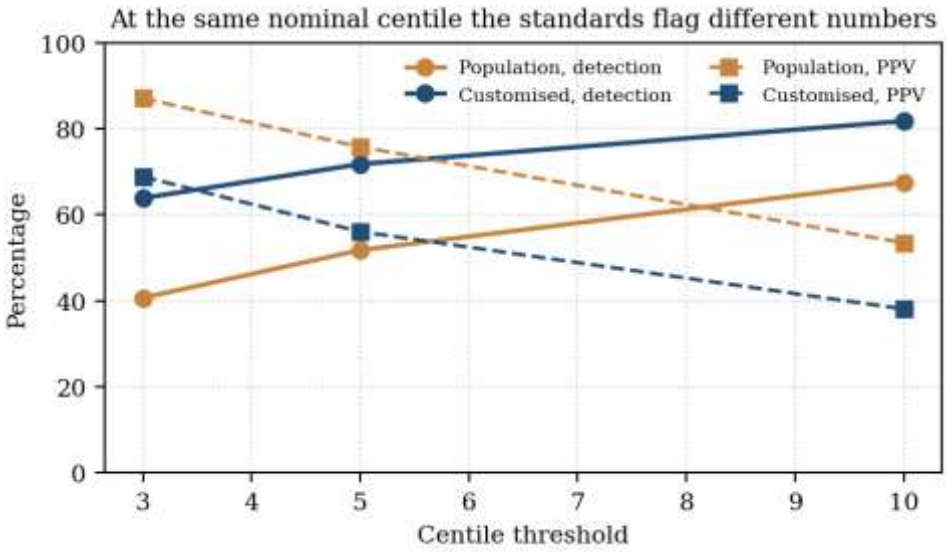

Figure 1. Detection and positive predictive value at nominal centile thresholds

Presented this way the customised standard looks decisively better: at the 10th centile it detects 81.8 % of truly restricted babies against 67.6 %, and misses 18.2 % against 32.4 %. This is the form in which the comparison is usually made and it is not a like-for-like comparison.

The second column explains why. At the same nominal centile the customised standard flags 12.8 % of all births and the population standard 7.5 %. A centile threshold does not correspond to a fixed proportion of the population once the reference distribution is individualised, because the customised centile removes maternal variation from the denominator and therefore redistributes babies across the scale. Comparing the two at "the 10th centile" compares a wider net with a narrower one.

3.2 The comparison at equal flagging

Table 3. Detection when both standards flag the same proportion of births

Births flagged	Standard	Threshold applied	Restriction detected (%)	Positive predictive value (%)	Restricted missed per 1000 births
2.79 %	Population	3rd centile	40.7	87.1	35.4
2.79 %	Customised	0.74th centile	42.7	91.4	34.2
4.08 %	Population	5th centile	51.8	75.8	28.8
4.08 %	Customised	1.72nd centile	55.1	80.6	26.8
7.53 %	Population	10th centile	67.6	53.5	19.4
7.53 %	Customised	4.90th centile	71.5	56.7	17.0

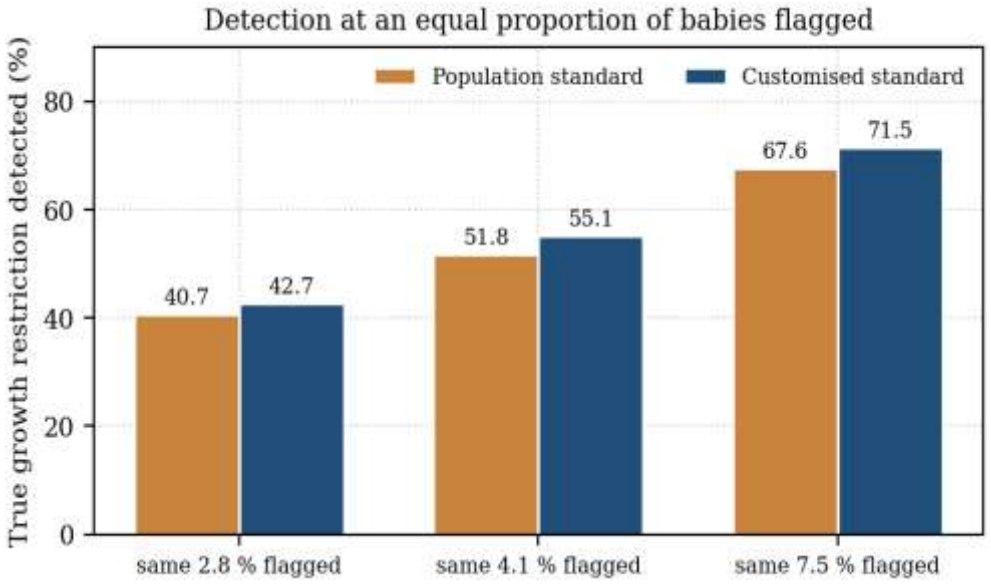

Figure 2. Detection of true growth restriction at an equal proportion of births flagged

This is the study's principal result. When both standards are set to flag 7.53 % of births, the customised standard detects 71.5 % of truly restricted babies against 67.6 % and achieves a positive predictive value of 56.7 % against 53.5 %. The advantage is real, consistent across all three flagging rates, and far smaller than the nominal-threshold comparison suggests: 3.9 percentage points of detection rather than 14.2.

Roughly three quarters of the apparent superiority of the customised standard at the 10th centile is therefore attributable to its flagging almost twice as many babies. A population standard applied at a lower threshold — or, equivalently, a service prepared to investigate 12.8 % of births rather than 7.5 % — would recover most of the difference without changing standard at all.

The residual advantage should not be dismissed. Reducing restricted babies missed from 19.4 to 17.0 per 1000 births is 2.4 fewer missed cases per 1000, which across a national birth cohort is a substantial

number, and it is obtained at no increase in the number investigated. It is simply a much more modest claim than the one usually made, and it is obtained here under an assumption — perfect knowledge of growth potential — that no real algorithm satisfies.

3.3 Which babies the standards disagree about

Table 4. Discordance at the 10th centile

Classification	Proportion of all births	Proportion truly restricted
Flagged by both standards	7.18 %	55.9 %
Flagged by the customised standard only	5.64 %	15.4 %
Flagged by the population standard only	0.36 %	5.1 %
Flagged by neither	86.83 %	—

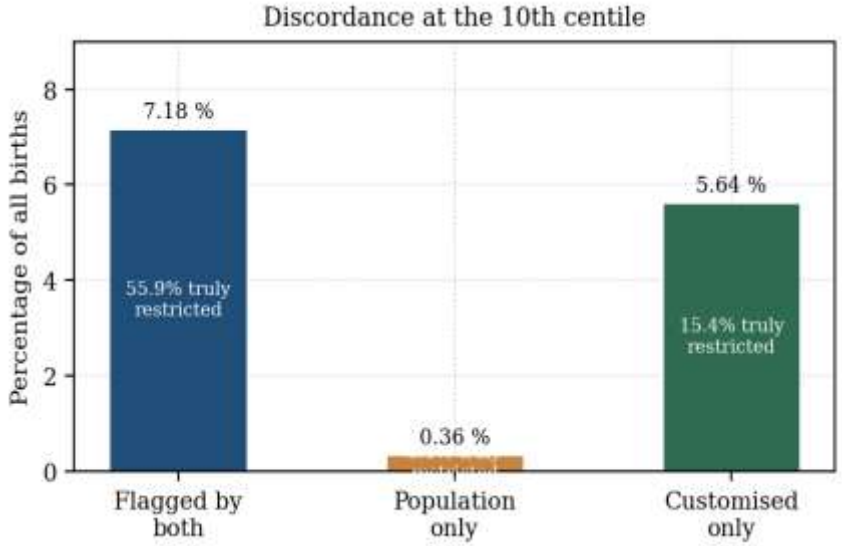

Figure 3. Discordance at the 10th centile, with the proportion truly restricted in each group

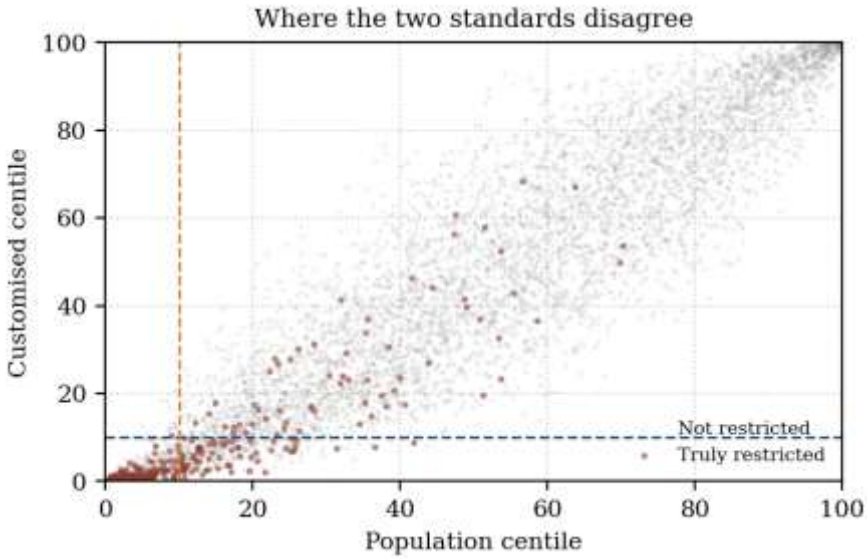

Figure 4. Population against customised centile for individual births, with truly restricted babies marked.

The two standards agree about 94 % of births and disagree about 6 %. The disagreement is markedly asymmetric: the customised standard adds 5.64 % of births that the population standard does not flag, while the population standard adds only 0.36 %.

The third column is what matters clinically. Among babies flagged only by the customised standard, 15.4 % are truly restricted — three times the 5.1 % among those flagged only by the population standard, and a yield that would justify investigation in most services. Among babies flagged by both, 55.9 % are restricted. The population-only group is therefore the least productive of the three, containing mostly babies who are small and well.

Table 5. Characteristics of the discordant groups

Group	n	Maternal height (cm)	Maternal weight (kg)	Parous (%)	Male (%)	Mean birthweight (g)
Flagged by population only	1,776	152.0	53.1	17.8	57.9	2,830
Flagged by customised only	28,218	165.4	69.3	64.4	49.4	3,014
All births	500,000	163.0	66.1	55.0	51.2	3,493

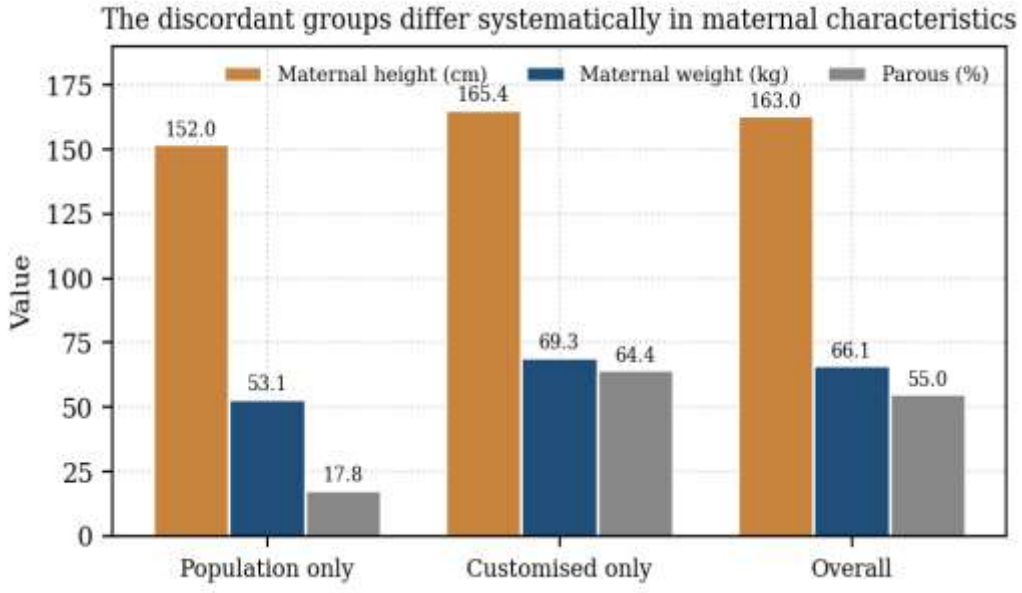

Figure 5. Maternal characteristics of the discordant groups against the whole cohort

Table 5 identifies who each standard is selecting. Babies flagged only by the population standard were born to mothers of mean height 152.0 cm and weight 53.1 kg, of whom only 17.8 % were parous — that is, to small, nulliparous mothers whose babies are expected to be small and, in 94.9 % of cases, are small for exactly that reason. Babies flagged only by the customised standard were born to mothers of mean height 165.4 cm and weight 69.3 kg, 64.4 % parous — mothers whose babies should have been larger, and in 15.4 % of cases were smaller than they should have been while remaining above the population 10th centile.

This is the substantive difference between the standards and it is a difference in selection rather than in aggregate performance. A population standard systematically directs surveillance toward the babies of small mothers and away from restricted babies of large mothers. Whether that pattern matters depends on the distribution of maternal size in the population served, which means the choice of standard has different consequences in different services.

4. DISCUSSION

4.1 Principal findings

Three findings follow. Comparisons of birthweight standards at the same nominal centile are not like-for-like, because the customised standard flagged 12.8 % of births at the 10th centile against the population standard's 7.5 %; roughly three quarters of its apparent advantage arises from that difference. A genuine discrimination advantage remains when flagging is equalised, but it is modest — 71.5 % against 67.6 % detection, and 2.4 fewer restricted babies missed per 1000 births. And the standards select systematically different babies, with the population standard preferentially flagging the constitutionally small offspring of small nulliparous mothers and missing restricted babies of larger, parous mothers.

The first finding qualifies a substantial literature. Studies reporting that customised centiles identify more babies at risk than population centiles, when both are applied at the tenth centile, are reporting a result that is partly definitional. The finding is not wrong and it is not a like-for-like comparison of discrimination, and the distinction has practical consequences because the two interpretations imply different actions: change the standard, or lower the threshold on the standard already in use.

4.2 Implications

Table 6. Implications

Finding	Implication	Recommendation
Standards flag different proportions at the same centile	Nominal-threshold comparisons are not like-for-like	Compare standards at equal flagging rates, and report the proportion flagged
Three quarters of the advantage is a threshold effect	Much of the benefit is available without changing standard	Consider whether the service can investigate a larger proportion
A modest discrimination advantage remains	Customisation adds something real	Do not dismiss customisation; quantify what it adds locally
Population-only flagged babies were 5.1 % restricted	That group is largely constitutionally small	Expect low yield from investigating small babies of small mothers
Customised-only flagged babies were 15.4 % restricted	That group is being missed by population standards	The gain from customisation is concentrated in babies of larger mothers
Selection differs by maternal size	Consequences depend on the population served	Evaluate the choice against local maternal anthropometry
Customisation was modelled with perfect knowledge	Real algorithms will do less well	Treat the reported advantage as an upper bound

The second row is the one most likely to change a service decision. If a unit's objective is to miss fewer growth-restricted babies, the analysis here suggests that most of the available gain comes from accepting a larger proportion of investigated pregnancies rather than from adopting a different standard — and that decision is under the unit's control, requires no change of software, and has a cost that can be calculated directly from its own birth numbers.

4.3 Limitations

The customised standard is modelled with perfect knowledge of each fetus's growth potential, which no real algorithm has. Published customisation models explain a limited share of the variance in birthweight, so their effective performance lies somewhere between the population standard and the idealised customisation modelled here. The 3.9-percentage-point advantage reported in Section 3.2 is therefore an upper bound on what customisation can deliver, and the real advantage is smaller by an amount this study cannot estimate.

Growth potential is modelled as a linear function of four maternal and fetal variables, and pathological restriction as an independent multiplicative deficit. Both are simplifications. In reality maternal size is itself partly a consequence of the maternal nutritional and social environment that also predisposes to placental insufficiency, so maternal smallness and fetal restriction are correlated rather than independent. That correlation, which this model omits, is the central objection to customisation raised by its critics: adjusting for maternal size may adjust away part of the pathology. Modelling it would reduce the customised advantage further, and possibly reverse it in the population-only group.

All parameters are representative assumptions rather than estimates. The prevalence of restriction, the size of the weight deficit, the physiological coefficient of variation and the coefficients in equation (1) all affect the absolute numbers, and a population with different maternal anthropometry would show different discordance. Gestational age is held constant at term, so preterm birth — where growth restriction is most consequential and where centile assignment is least reliable — is excluded entirely.

Finally, the outcome modelled is the presence of pathological restriction rather than any clinical event. Whether identifying a restricted baby leads to action that improves outcome is a separate question, and a standard that detects more restriction is only useful to the extent that detection changes management beneficially. Nothing here addresses that link.

5. CONCLUSION

At the 10th centile a customised birthweight standard detected 81.8 % of truly growth-restricted babies against a population standard's 67.6 %, while flagging 12.8 % of all births against 7.5 %. When both were set to flag the same proportion, the difference narrowed to 71.5 % against 67.6 %. Roughly three quarters of the customised standard's apparent advantage is therefore a consequence of flagging more babies rather than of discriminating between them better, and the residual advantage was obtained under the generous assumption that growth potential is known exactly.

What is not a threshold effect is the difference in which babies are selected. The two standards disagreed about 6 % of births, and the disagreement was systematic: babies flagged only by the population standard were born to small nulliparous mothers and were truly restricted in 5.1 % of cases, while those flagged only by the customised standard were born to larger parous mothers and were truly restricted in 15.4 %. The practical question for a service is therefore not which standard is better in the abstract but whether its population contains many mothers in the second group — and, before changing standard at all, whether the simpler option of investigating a larger proportion of pregnancies would deliver most of the same benefit.

Declarations

Ethical approval: Not applicable.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflicts of interest: The authors declare that there are no conflicts of interest regarding the publication of this paper.

Data availability: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Author contributions:

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Dr. PriyadharshiniP.	✓	✓	✓		✓	✓		✓	✓	✓	✓		✓	✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

6. REFERENCES

1. J. Gardosi, A. Chang, B. Kalyan, D. Sahota, and E. M. Symonds, "Customised antenatal growth charts," Lancet, vol. 339, no. 8788, pp. 283-287, 1992. doi.org/10.1016/0140-6736(92)91342-6
2. J. Gardosi, A. Francis, S. Turner, and M. Williams, "Customized growth charts: rationale, validation and clinical benefits," Am J Obstet Gynecol, vol. 218, no. 2S, pp. S609-S618, 2018. doi.org/10.1016/j.ajog.2017.12.011
3. J. Gardosi, V. Madurasinghe, M. Williams, A. Malik, and A. Francis, "Maternal and fetal risk factors for stillbirth: population based study," BMJ, vol. 346, 2013. doi.org/10.1136/bmj.f108
4. E. Carberry, A. Gordon, D. M. Bond, J. Hyett, C. H. Raynes-Greenow, and J. He, "Customised versus population-based growth charts as a screening tool for detecting small for gestational age infants in low-risk pregnant women," Cochrane Database Syst Rev, no. 5, 2014. doi.org/10.1002/14651858.CD008549.pub3
5. S. Iliodromiti, D. F. Mackay, and G. Smith, "Customised and noncustomised birth weight centiles and prediction of stillbirth and infant mortality and morbidity: a cohort study of 979,912 term singleton pregnancies in Scotland," PLoS Med, vol. 14, no. 1, 2017. doi.org/10.1371/journal.pmed.1002228
6. J. A. Hutcheon, X. Zhang, S. Cnattingius, M. S. Kramer, and R. W. Platt, "Customised birthweight percentiles: does adjusting for maternal characteristics matter?," BJOG, vol. 115, no. 11, pp. 1397-1404, 2008. doi.org/10.1111/j.1471-0528.2008.01870.x
7. J. Zhang, M. Merialdi, L. D. Platt, and M. S. Kramer, "Defining normal and abnormal fetal growth: promises and challenges," Am J Obstet Gynecol, vol. 202, no. 6, pp. 522-528, 2010. doi.org/10.1016/j.ajog.2009.10.889
8. J. Villar, C. Ismail, and L. Victora, "INTERGROWTH-21st. 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, vol. 384, no. 9946, pp. 857-868, 2014. doi.org/10.1016/S0140-6736(14)60932-6
9. N. H. Anderson, L. C. Sadler, C. Mckinlay, and L. Mccowan, "INTERGROWTH-21st vs customized birthweight standards for identification of perinatal mortality and morbidity," Am J Obstet Gynecol, vol. 214, no. 4, pp. 509-e510, 2016. doi.org/10.1016/j.ajog.2015.10.931
10. G. Chiossi, C. Pedroza, and M. M. Costantine, "Customized vs population-based growth charts to identify neonates at risk of adverse outcome: systematic review and Bayesian meta-analysis of observational studies," Ultrasound Obstet Gynecol, vol. 50, no. 2, pp. 156-166, 2017. doi.org/10.1002/uog.17381
11. S. J. Gordijn, I. M. Beune, and B. Thilaganathan, "Consensus definition of fetal growth restriction: a Delphi procedure," Ultrasound Obstet Gynecol, vol. 48, no. 3, pp. 333-339, 2016. doi.org/10.1002/uog.15884
12. L. M. Mccowan, F. Figueras, and N. H. Anderson, "Evidence-based national guidelines for the management of suspected fetal growth restriction: comparison, consensus, and controversy," Am J Obstet Gynecol, vol. 218, no. 2S, pp. S855-S868, 2018. doi.org/10.1016/j.ajog.2017.12.004
13. M. S. Kramer, R. W. Platt, and S. W. Wen, "A new and improved population-based Canadian reference for birth weight for gestational age," Pediatrics, vol. 108, no. 2, 2001. doi.org/10.1542/peds.108.2.e35
14. D. G. Altman and J. M. Bland, "Statistics notes: diagnostic tests 2 - predictive values," BMJ, vol. 309, no. 6947, 1994. doi.org/10.1136/bmj.309.6947.102
15. T. P. Morris, I. R. White, and M. J. Crowther, "Using simulation studies to evaluate statistical methods," Stat Med, vol. 38, no. 11, pp. 2074-2102, 2019. doi.org/10.1002/sim.8086