Eur J Pediatr (2016) 175:1047–1057
DOI 10.1007/s00431-016-2740-8

ORIGINAL ARTICLE

Defining small-for-gestational-age: prescriptive versus descriptive birthweight standards

Liset Hoftiezer[1] · Chantal W. P. M. Hukkelhoven[2] · Marije Hogeveen[3] · Huub M. P. M. Straatman[4] · Richard A. van Lingen[1]

Received: 15 December 2015 / Revised: 1 May 2016 / Accepted: 25 May 2016 / Published online: 2 June 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract Descriptive population-based birthweight standards possess low sensitivity in detecting infants with growth impairment. A prescriptive birthweight standard based on a 'healthy' subpopulation without risk factors for intrauterine growth restriction might be superior. We created two birthweight standards based on live born, singleton infants with gestational age 24–42 weeks and born in The Netherlands between 2000 and 2007. Inclusion criteria for the prescriptive birthweight standard were restricted to infants without congenital malformations, born to healthy mothers after uncomplicated pregnancies. We defined small-for-gestational-age (SGA) as birthweight <10th percentile and assessed the ability of both standards to predict adverse neonatal outcomes. The prescriptive birthweight standard identified significantly more infants as SGA, up to 38.0 % at 29 weeks gestation. SGA infants classified according to both standards as well as those classified according to the prescriptive birthweight standard only, were at increased risk of both major and minor adverse neonatal outcomes. The prescriptive birthweight standard was both more sensitive and less specific, with a maximum increase in sensitivity predicting bronchopulmonary dysplasia (+42.6 %) and a maximum decrease in specificity predicting intraventricular haemorrhage (−26.9 %) in infants aged 28–31 weeks.

Conclusion: Prescriptive birthweight standards could improve identification of infants born SGA and at risk of adverse neonatal outcomes.

What is Known:
- *Descriptive birthweight standards possess low sensitivity in detecting growth restricted infants at risk of adverse neonatal outcomes.*
- *Prescriptive standards could improve identification of very preterm small-for-gestational-age (SGA) infants at risk of intraventricular haemorrhage.*

What is New:
- *Prescriptive standards identify more preterm and term SGA infants at risk of major adverse neonatal outcomes.*
- *Late preterm and term SGA infants classified according to the prescriptive standard are at increased risk of minor adverse neonatal outcomes with potentially harmful implications.*

Keywords Small for gestational age · Intrauterine growth restriction · Reference standards · Sensitivity and specificity · Perinatal mortality and morbidity

Communicated by Patrick Van Reempts

Electronic supplementary material The online version of this article (doi:10.1007/s00431-016-2740-8) contains supplementary material, which is available to authorized users.

✉ Liset Hoftiezer
l.hoftiezer@isala.nl

[1] Princess Amalia Department of Paediatrics, Department of Neonatology, P.O. Box 10400, 8000 GK Isala, Zwolle, The Netherlands

[2] Perined, Utrecht, The Netherlands

[3] Amalia Children's Hospital, Department of Neonatology, Radboud University Medical Centre, Nijmegen, The Netherlands

[4] PHARMO Institute for Drug Outcomes Research, Utrecht, The Netherlands

Abbreviations

aOR	Adjusted odds ratio
BPD	Bronchopulmonary dysplasia
cPVL	Cystic periventricular leukomalacia
GA	Gestational age
HIE	Hypoxic ischemic encephalopathy
IRDS	Infant respiratory distress syndrome
IUGR	Intrauterine growth restriction

IVH	Intraventricular haemorrhage
NEC	Necrotising enterocolitis
PRN	The Netherlands Perinatal Registry
ROP	Retinopathy of prematurity
SGA	Small-for-gestational-age

Introduction

Intrauterine growth restriction (IUGR) is etiologically associated with various maternal, foetal and placental factors [6, 19, 30, 31]. Infants who have experienced IUGR are at higher risk of neonatal mortality and morbidity [10, 16, 32, 34, 41]. Moreover, they may be at increased risk of long-term adverse outcome, such as neurocognitive impairment, poor school performance and short stature [23, 27, 28]. Associations between IUGR and chronic diseases in (early) adulthood, such as coronary heart disease and type 2 diabetes mellitus, have also been described extensively [1]. There exists no universally agreed means of diagnosing IUGR. Small-for-gestational-age (SGA), often defined as birthweight below the gender-specific 10th percentile for gestational age (GA), is commonly used as a surrogate [2].

To distinguish between normal birthweight and birthweight deviation, reference values are required. In the absence of exclusion criteria regarding risk factors for IUGR, a reference is descriptive. A descriptive reference allows application to other neonates to establish whether or not their measurements are typical of the reference group. In the presence of exclusion criteria regarding risk factors for IUGR, a reference is prescriptive. A prescriptive reference constitutes a model to which a neonate should conform and indicates how growth should progress [4, 43].

It is well known that descriptive references possess low sensitivity in detecting IUGR [5]. It is less evident that a prescriptive standard is superior. The one study we identified had limited statistical power [13]. Studies that compared descriptive birthweight references to prescriptive foetal growth references consistently showed the benefits of the latter, i.e. superiority in detecting morbidity and mortality [5, 25, 29, 42]. However, foetal growth references differ from neonatal birthweight references in many methodological aspects and it is not evident that their prescriptive nature alone accounts for their superiority.

Considering the far-reaching consequences of IUGR, it is important to interpret birthweight for a given gestational age. The present study was initiated because of the observation of major discrepancies between the diagnosis of IUGR made in utero and the neonatal classification based upon descriptive birthweight standards. Inspired by previous studies, we hypothesised that prescriptive birthweight standards are superior to descriptive birthweight standards [5, 13, 25, 29, 42]. The aim of this study was to investigate to what extent a prescriptive birthweight standard based on a subpopulation free of risk factors for IUGR could improve the identification of infants born SGA at risk of adverse neonatal outcomes.

Materials and methods

Study design

We performed a two-phase retrospective cohort study. Firstly, we developed descriptive and prescriptive birthweight standards and classified neonates as either SGA or non-SGA accordingly. Secondly, we assessed the associations between the two SGA definitions and adverse neonatal outcomes and evaluated the birthweight standards in terms of diagnostic performance.

Study population

Data were extracted from the perinatal database of The Netherlands Perinatal Registry (PRN), which is a linked database of medical registries from four professional organisations that provide perinatal care. Among the items reported are maternal demographics and medical conditions, pregnancy complications and details concerning labour, birth and neonatal outcome, most of which are registered as diagnostic codes. Although participation is not 100 %, the database can be considered as an unbiased representation of the total population in The Netherlands [38, 40].

We included infants with GA between 24^{+0} and 42^{+0} weeks, who were alive at the onset of labour and were born in The Netherlands between January 1, 2000, and December 31, 2007. In over 95 % of the pregnancies, GA was certain, either confirmed by or based on early ultrasound. To construct descriptive birthweight standards, infants with missing GA, birthweight or gender data were excluded. In accordance with the current Dutch birthweight standards, we also excluded multiple pregnancies [40]. To construct prescriptive birthweight standards, we further restricted the inclusion criteria to infants without congenital malformations, born to healthy mothers after pregnancies without suggested risk factors for impaired foetal growth, and, in case of prematurity, after spontaneous onset of labour (Fig. 1).

For the second phase of our study, we selected a sample group, consisting of neonates who were born in hospitals that take part in both the obstetric and paediatric registration. Sampling was necessary to assure that infants without recorded adverse neonatal outcomes could indeed be considered healthy, instead of being falsely labelled as such due to incomplete registration. We extended the sample by adding (low risk) births attended by midwives if the mothers' postal codes were equal to those of mothers who gave birth in one of the selected hospitals. More details are provided in Fig. 1.

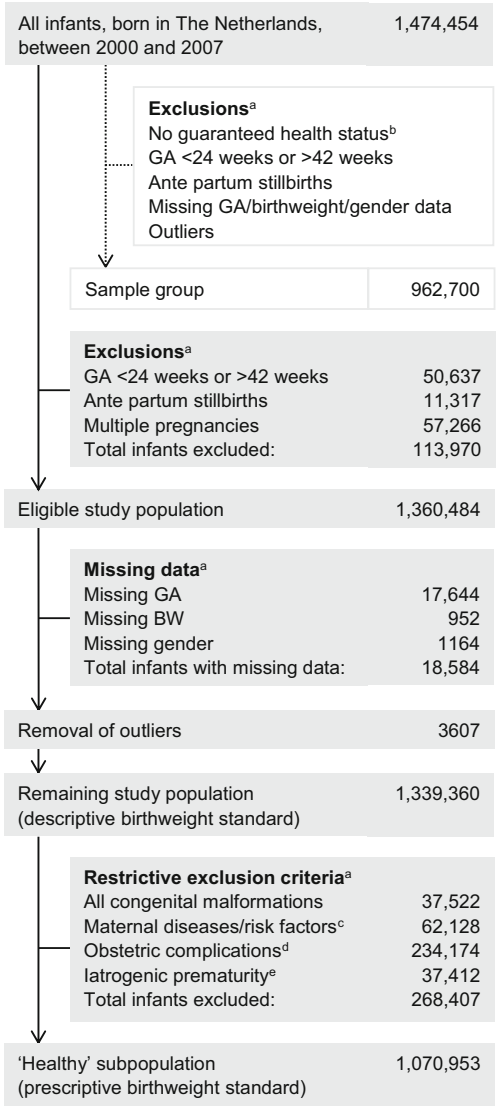

Fig. 1 Flow diagram of study participants. [a]More than one exclusion criterion or missing data variable may apply to the same pregnancy. [b]Health status was considered unknown when infants were born in hospitals that did not participate in the paediatric registration. [c]Pre-existent hypertension; pre-gestational diabetes mellitus; 'other' systemic diseases; cardiovascular disorders with haemodynamic consequences; thromboembolic disorders; respiratory disorders; epilepsy treated with anticonvulsants; malignant diseases; anaemia (haemoglobin level <9.6 g/dL (<6.0 mmol/L)); recurrent urinary tract infections; substance abuse; medication use. [d]Gestational hypertension; proteinuria; pre-eclampsia; eclampsia; placenta praevia; placental abruption; TORCHES; disorders of pregnancy 'not otherwise specified'. [e]Primary caesarean section or induction of labour <37 weeks GA

Selected neonates were classified as SGA or non-SGA (i.e. birthweight < p10 or birthweight > p10), separately according to both birthweight standards.

Data analysis

Baseline characteristics were analysed using descriptive statistics and presented in an appropriate manner.

We used Tukey's method to exclude outliers; observations were excluded if birthweight was more than twice the interquartile range below the first quartile or above the third quartile [22]. The LMS method was adopted to calculate birthweight percentiles [8]. The curves were smoothed using penalised B-splines. For more details, we refer to Appendix 1 (online supplementary material). Model fit was evaluated by visual assessment of the smoothed versus empirical percentiles (online supplementary Fig. S1).

In the absence of a gold standard, the standards were evaluated by their ability to 'predict' adverse neonatal outcomes associated with IUGR. Early perinatal mortality was defined as death within 7 days postpartum, including intrapartum death. Late perinatal mortality was defined as death between 8 and 28 days postpartum. Low 5-min Apgar score was defined as an Apgar score of ≤3 at 5 min postpartum [9]. The diagnosis of sepsis was established by a positive blood culture and/or based on clinical symptoms. Infant respiratory distress syndrome (IRDS) was diagnosed based on clinical and radiographic findings according to Giedion [15]. Bronchopulmonary dysplasia (BPD) was diagnosed in infants requiring oxygen supplementation beyond 36 weeks postconceptional age. Intraventricular haemorrhage (IVH) and cystic periventricular leukomalacia (cPVL) were diagnosed based on classic sonographic findings and classified according to Papile and De Vries, respectively [11, 35]. Necrotizing enterocolitis (NEC) was diagnosed based on a combination of clinical, laboratory and radiographic findings and staged according to Bell's classification [3]. Retinopathy of prematurity (ROP) was diagnosed based on fundoscopic examination and classified according to international classification [21]. Hypoxic ischemic encephalopathy (HIE) was diagnosed based on clinical presentation, presence of seizures and duration of symptoms and graded according to Sarnat's classification [37]. Hypoglycaemia was defined as a blood glucose level <2.5 mmol/L (45 mg/dL). Hyperbilirubinaemia was diagnosed when phototherapy was required. Hypothermia was defined as a body temperature <35.5 °C (95.9 °F). Polycythaemia was defined as a venous haematocrit >0.65.

Because the risk of adverse outcomes decreases substantially with increasing GA, we conducted separate analyses for different GA strata.

We performed multiple logistic regression analyses to relate each outcome variable to growth status (SGA versus reference category 'non-SGA'). Various potential confounders were identified and included in the model if statistically significant at the 0.05 level (two-tailed Wald test) [24]. We estimated adjusted odds ratios (aORs) and 95 % confidence intervals for each SGA definition. The diagnostic performance of both standards was assessed in terms of sensitivity and specificity and compared using McNemar's test for matched binary data [17, 18].

Statistical analyses were performed with SAS 9.2 software (SAS Institute, Inc., Cary, NC, USA).

Results

During the study period, 1,474,454 infants were born in The Netherlands. We excluded 113,970 (7.7 %) records because of inappropriate GA (i.e. <24^{+0} or >42^{+0} weeks), antepartum stillbirth, multiple pregnancy or missing data. Subsequently 3607 (0.3 %) outliers were excluded. The remaining study population comprised 1,339,360 live births (Fig. 1).

To obtain a 'healthy' subpopulation, another 268,407 (20.0 %) infants were excluded. In general, the exclusion rate decreased with increasing GA (online supplementary Tables S1 and S2). The principal cause of exclusion was maternal hypertension, defined as a diastolic blood pressure of >90 mmHg. The final healthy subpopulation comprised 1,070,953 records (Fig. 1).

Baseline characteristics are shown in Table 1. Compared to the healthy subpopulation, the excluded records comprised more nulliparous women (53.9 versus 43.8 %), preterm births (13.9 versus 3.9 %), low birthweights (<2000 g, 5.8 versus 0.6 %) and hospital deliveries (93.0 versus 55.3 %).

Figure 2a, b shows gender-specific 10th, 50th and 90th percentiles for both birthweight standards. The difference between the standards decreases both with increasing GA and with increasing birthweight. The maximum difference between the two 10th percentiles is 312 and 362 g for boys and girls, respectively (GA 31 weeks). Consequently, the heaviest SGA infants according to the prescriptive birthweight standard may be at most 362 g heavier than the heaviest SGA infants classified according to the descriptive birthweight standard. At 39 weeks GA, the two standards agree almost perfectly, as do the 97.7th percentiles (irrespective of GA; online supplementary Tables S1 and S2). The overall percentage of infants classified as SGA according to the descriptive standard was ≈10 %. The prescriptive birthweight standard classified significantly more (preterm) infants as SGA, up to 38.0 % at 29 weeks GA ($p < 0.0001$; Fig. 2c).

Both SGA definitions were significantly associated with multiple adverse neonatal outcomes (Table 2). All preterm infants including those aged 24^{+0}–25^{+6} weeks (online supplementary Table S3) were at significantly increased risk of low 5-min Apgar scores and early neonatal death, irrespective of which standard was used. Infants aged 26^{+0}–31^{+6} weeks were at increased risk of late neonatal death, BPD and ROP. The risk of developing NEC or sepsis was increased starting from GA 28^{+0} weeks. The risk of cPVL was decreased in infants aged 26^{+0}–27^{+6} weeks, but increased in infants aged 28^{+0} weeks and older. Infants aged 26^{+0}–31^{+6} weeks were at decreased risk of IVH, whereas infants aged 32^{+0}–36^{+6} weeks were at increased risk. The latter were also at increased risk of minor adverse outcomes hypoglycaemia, hyperbilirubinaemia, hypothermia and polycythaemia.

Term infants (GA 37^{+0}–39^{+0} weeks) were at increased risk of both early and late neonatal death, low 5-min Apgar score, sepsis, HIE and hypoglycaemia, hypothermia, hyperbilirubinaemia and polycythaemia (online supplementary Table S3). Infants aged >39^{+0} weeks were eliminated from further analyses because there was no substantial difference between the 10th percentiles.

To assess the additional benefit of using the prescriptive birthweight standard, separate analyses were conducted for those infants classified as 'extra' SGA (i.e. according to the prescriptive birthweight standard only). The results are also shown in Table 2. Overall, the additional benefit increased with increasing GA. Interestingly, while the risk of IRDS was significantly decreased in SGA infants aged 32^{+0}–36^{+6} weeks when classified according to the descriptive birthweight standard (aOR 0.57 [0.49–0.67]), the risk was significantly increased in extra SGA infants classified according to the prescriptive birthweight standard only (aOR 1.30 [1.16–1.46]).

The diagnostic performance of the two standards was assessed in terms of sensitivity and specificity and compared using McNemar's test (Fig. 3; online supplementary Fig. S2 and Tables S4 and S5). The prescriptive birthweight standard showed significantly improved sensitivity, when compared to the descriptive birthweight standard (generally $p < 0.0001$). The increased sensitivity was proportional to the relative increase in the percentage of SGA infants, with the highest sensitivity in the GA stratum with the highest SGA percentage according to the prescriptive birthweight standard (GA 28^{+0}–31^{+6} weeks). As a result of the inversely proportional relationship between sensitivity and specificity, the improved sensitivity was accompanied by a significant decrease in specificity ($p < 0.0001$) that was also most apparent in infants aged 28^{+0}–31^{+6} weeks (Fig. 3).

Discussion

Our study shows that prescriptive birthweight standards could improve the identification of SGA infants at risk of adverse neonatal outcomes, when compared to descriptive birthweight standards.

The two birthweight distributions we assessed were substantially different. Because risk factors for IUGR were more prevalent among preterm infants, comparatively many preterm infants did not fulfil the inclusion criteria of the prescriptive birthweight standard. The difference between the prescriptive and descriptive birthweight standard decreased with increasing GA, reflecting the higher incidence of IUGR and hence exclusion rate among preterm infants. Consequently, we found that SGA rates were much higher in preterm neonates when classified by the prescriptive birthweight standard. These results were consistent with a similar study performed

Table 1 Baseline characteristics of the two populations and the excluded records

	Study population *n* (%)	'Healthy' subpopulation *n* (%)	Excluded records[a] *n* (%)
Parity			
Nulliparous	613,306 (45.8)	468,600 (43.8)	144,706 (53.9)
Multiparous	725,940 (54.2)	602,255 (56.2)	123,685 (46.1)
Missing	114 (0.0)	98 (0.0)	16 (0.0)
Infant's gender			
Male	687,281 (51.3)	545,601 (50.9)	141,680 (52.8)
Female	652,079 (48.7)	525,352 (49.1)	126,727 (47.2)
Gestational age (weeks)			
>37	1,260,125 (94.1)	1,028,984 (96.1)	231,141 (86.1)
32–36	68,882 (5.1)	38,386 (3.6)	30,496 (11.4)
28–31	7667 (0.6)	2387 (0.2)	5280 (2.0)
24–27	2686 (0.2)	1196 (0.1)	1490 (0.6)
Mean (days) [SD]	277.3 [12.7]	278.4 [10.9]	272.5 [17.1]
Birthweight (g)			
0–999	3458 (0.3)	928 (0.1)	2530 (0.9)
1000–1999	18,891 (1.4)	5866 (0.6)	13,025 (4.9)
2000–2999	222,030 (16.6)	157,532 (14.7)	64,498 (24.0)
3000–3999	892,873 (66.7)	739,162 (69.0)	153,711 (57.3)
4000+	202,108 (15.1)	167,465 (15.6)	34,643 (12.9)
Mean (g) [SD]	3439.4 [571.3]	3485.0 [517.6]	3257.0 [719.8]
Place of delivery			
Home delivery	336,448 (25.1)	326,875 (30.5)	9573 (3.6)
Midwife-assisted in-hospital delivery	158,064 (11.8)	149,188 (13.9)	8876 (3.3)
Hospital delivery	841,764 (62.9)	592,088 (55.3)	249,676 (93.0)
Missing	3084 (0.2)	2802 (0.3)	282 (0.1)
Maternal age (years)			
<15	169 (0.0)	129 (0.0)	40 (0.0)
15–35	1,149,933 (85.9)	922,937 (86.2)	226,996 (84.6)
36+	188,986 (14.1)	147,663 (13.8)	41,323 (15.4)
Missing	272 (0.0)	224 (0.0)	48 (0.0)
Mean (years) [SD]	30.4 [4.8]	30.4 [4.8]	30.6 [4.9]
Ethnicity			
Dutch/other European	1,125,183 (84.0)	893,130 (83.4)	232,053 (86.5)
Other	205,953 (15.4)	170,273 (15.9)	35,680 (13.3)
Missing	8224 (0.6)	7550 (0.7)	674 (0.3)
Socioeconomic status			
High	313,646 (23.4)	252,815 (23.6)	60,831 (22.7)
Middle	606,673 (45.3)	483,957 (45.2)	122,716 (45.7)
Low	402,120 (30.0)	320,903 (30.0)	81,217 (30.3)
Missing	16,921 (1.3)	13,278 (1.2)	3643 (1.4)

[a] Excluded based on the presence of risk factors for IUGR

by Ferdynus et al. [13]. Previous studies that compared foetal growth standards to (descriptive) neonatal birthweight standards also reported similar results, supporting the hypothesis that prescriptive birthweight standards might approximate foetal growth standards [12, 29, 42]. So-called customised standards adjust birthweight further by considering the influence of supposedly physiological determinants of growth, mostly maternal characteristics [14]. Although customised standards have consistently shown improved prediction of neonatal mortality and morbidity, the actual contribution of maternal characteristics has been questioned [7, 20]. Moreover, several studies have demonstrated that maternal

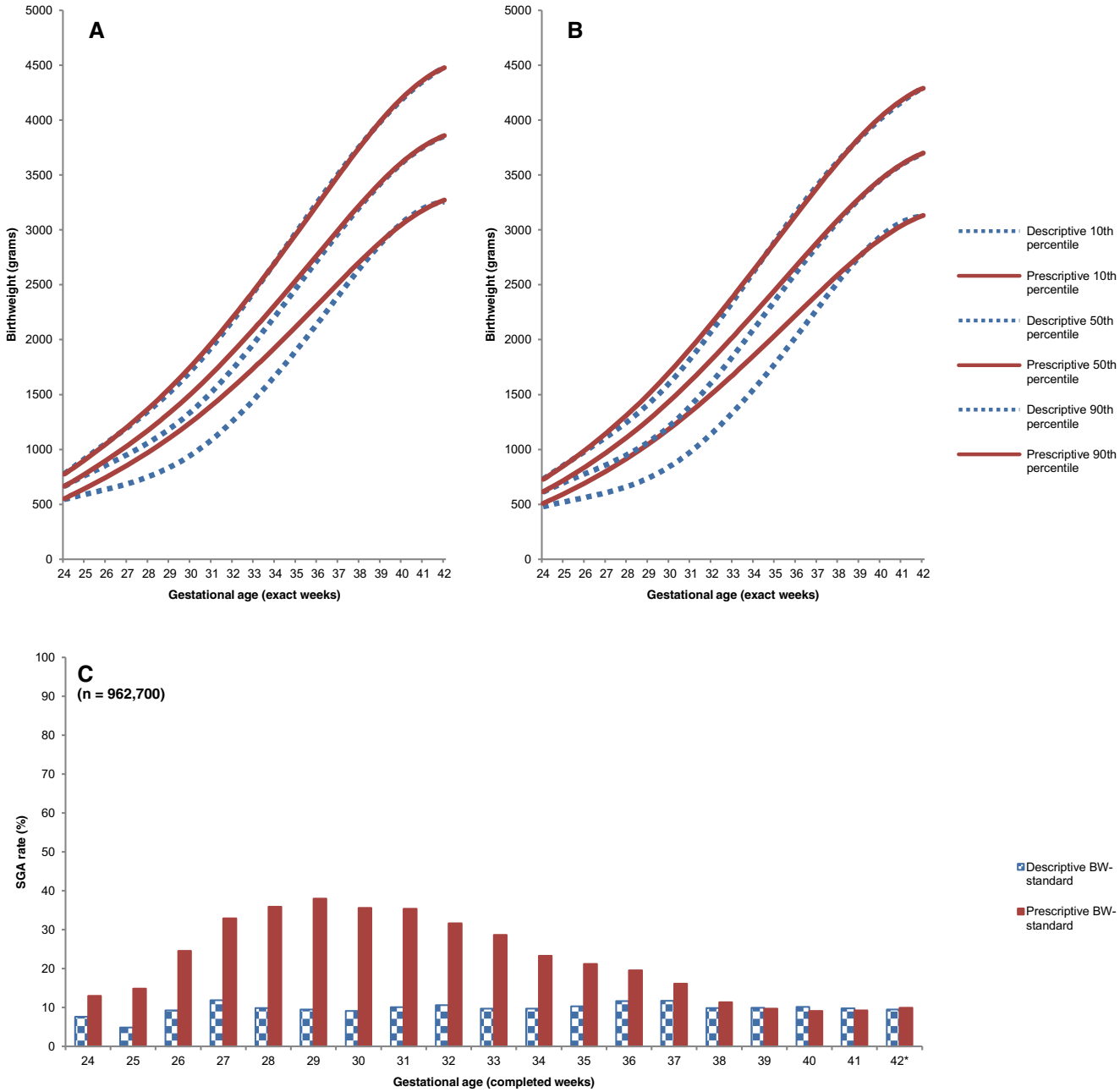

Fig. 2 Descriptive and prescriptive 10th, 50th and 90th percentiles for boys (**a**) and girls (**b**) and SGA rates in the sample group according to both standards (**c**). * Forty-two weeks includes only infants born at GA 42^{+0} weeks, whereas the other GAs include the entire week (e.g. 24^{+0} to 24^{+6} weeks). *BW* birthweight

influences on foetal growth may not be purely physiological [20, 44]. Because our database lacked information on most of the commonly used variables, we were not able to investigate whether customisation has an additional advantage over the application of restrictive inclusion criteria.

In our study, SGA was significantly associated with many adverse neonatal outcomes, including several minor adverse outcomes with potentially harmful implications [34, 41]. In general, associations were strongest when SGA was classified according to the descriptive standard. Since these infants were the smallest infants classified by both standards, they may be expected to be at greatest risk. However, extra SGA infants classified according to the prescriptive birthweight standard only were still at (significantly) increased risk of adverse neonatal outcomes. The additional benefit of the prescriptive birthweight standard increased with increasing GA, reflecting the shape of the 'area between the curves' and the (increasing) number of subjects within succeeding GA strata. Similar to Ferdynus et al., our results support the hypothesis that prescriptive birthweight standards can improve the identification

Table 2 Adjusted ORs and confidence intervals for separate GA strata

Adverse neonatal outcome	GA 26^{+0}–27^{+6} weeks						GA 28^{+0}–31^{+6} weeks		
	SGA descriptive		SGA prescriptive		'Extra' SGA[a]		SGA descriptive		SGA prescriptive
	aOR[b]	*p*	aOR[b]	*p*	aOR[b]	*p*	aOR[b]	*p*	aOR[b]
Low 5-min Apgar score	1.52 [1.03, 2.24]	0.0343	1.49 [1.11, 2.00]	0.0077	1.34 [0.95, 1.90]	0.0988	1.69 [1.27, 2.24]	0.0003	1.36 [1.10, 1.69]
Early neonatal mortality	2.55 [1.87, 3.48]	<0.0001	2.30 [1.81, 2.93]	<0.0001	1.94 [1.45, 2.59]	<0.0001	1.81 [1.40, 2.35]	<0.0001	1.36 [1.12, 1.66]
Late neonatal mortality	1.84 [1.19, 2.85]	0.0064	1.56 [1.11, 2.19]	0.0107	1.33 [0.87, 2.01]	0.1863	2.41 [1.61, 3.63]	<0.0001	1.41 [1.01, 1.97]
Sepsis[c]	0.92 [0.71, 1.21]	0.5572	0.95 [0.79, 1.14]	0.5454	0.96 [0.77, 1.19]	0.7183	1.55 [1.35, 1.77]	<0.0001	1.32 [1.21, 1.44]
IRDS	0.85 [0.64, 1.13]	0.2557	0.99 [0.81, 1.20]	0.8946	1.07 [0.85, 1.35]	0.5827	0.93 [0.81, 1.06]	0.3571	1.09 [1.00, 1.19]
BPD	2.21 [1.40, 3.49]	0.0007	1.67 [1.15, 2.42]	0.0066	1.29 [0.81, 2.06]	0.2795	3.19 [1.86, 4.12]	<0.0001	3.79 [2.65, 5.43]
IVH	0.57 [0.41, 0.80]	0.0010	0.53 [0.42, 0.66]	<0.0001	0.55 [0.42, 0.71]	<0.0001	0.85 [0.69, 1.04]	0.1108	0.82 [0.72, 0.93]
cPVL	0.24 [0.06, 0.99]	0.0485	0.63 [0.35, 1.14]	0.1286	0.85 [0.45, 1.61]	0.6191	1.60 [1.02, 2.50]	0.0404	1.19 [0.86, 1.63]
NEC	1.30 [0.86, 1.97]	0,2197	1.31 [0.97, 1.76]	0.0765	1.26 [0.88, 1.79]	0.2066	3.36 [2.66, 4.25]	<0.0001	2.36 [1.93, 2.87]
ROP	1.60 [1.22, 2.31]	0.0015	1.31 [1.03, 1.66]	0.0269	1.11 [0.83, 1.49]	0.4909	3.17 [2.40, 4.19]	<0.0001	2.38 [1.88, 3.02]
HIE	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Hypoglycaemia	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Hyperbilirubinaemia	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Hypothermia	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Polycythaemia	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Adverse neonatal outcome	GA 28^{+0}–31^{+6} weeks			GA 32^{+0}–36^{+6} weeks					
	SGA prescriptive	'Extra' SGA[a]		SGA descriptive		SGA prescriptive		'Extra' SGA[a]	
	p	aOR[b]	*p*	aOR[b]	*p*	aOR[b]	*p*	aOR[b]	*p*
Low 5-min Apgar score	0.0048	1.19 [0.93, 1.52]	0.1784	2.34 [1.90, 2.88]	<0.0001	2.09 [1.74, 2.51]	<0.0001	1.60 [1.24, 2.08]	0.0004
Early neonatal mortality	0.0020	1.14 [0.91, 1.44]	0.2539	2.15 [1.75, 2.63]	<0.0001	1.97 [1.64, 2.36]	<0.0001	1.55 [1.20, 2.01]	0.0008
Late neonatal mortality	0.0469	1.02 [0.68, 1.54]	0.9257	1.89 [1.11, 3.23]	0.0198	1.52 [0.93, 2.47]	0.0949	0.94 [0.44, 2.00]	0.8678
Sepsis[c]	<0.0001	1.23 [1.12, 1.36]	<0.0001	1.39 [1.26, 1.54]	<0.0001	1.47 [1.35, 1.59]	<0.0001	1.45 [1.31, 1.61]	<0.0001
IRDS	0.0534	1.13 [1.03, 1.24]	0.0121	0.57 [0.49, 0.67]	<0.0001	0.94 [0.85, 1.04]	0.2122	1.30 [1.16, 1.46]	<0.0001
BPD	<0.0001	3.22 [2.18, 4.76]	<0.0001	n/a	n/a	n/a	n/a	n/a	n/a
IVH	0.0018	0.83 [0.72, 0.95]	0.0072	1.48 [1.14, 1.93]	0.0030	1.63 [1.33, 2.01]	<0.0001	1.66 [1.27, 2.16]	0.0002
cPVL	0.2938	1.03 [0.72, 1.50]	0.8602	2.76 [1.55, 4.92]	0.0006	2.33 [1.39, 3.91]	0.0014	1.59 [0.76, 3.33]	0.2135
NEC	<0.0001	1.76 [1.40, 2.21]	<0.0001	3.08 [2.34, 4.05]	<0.0001	3.53 [2.74, 4.55]	<0.001	2.85 [2.07, 3.93]	<0.0001
ROP	<0.0001	1.90 [1.44, 2.50]	<0.0001	n/a	n/a	n/a	n/a	n/a	n/a
HIE	n/a	n/a	n/a	1.08 [0.71, 1.66]	0.7163	1.46 [1.08, 1.99]	0.0149	1.77 [1.22, 2.57]	0.0028
Hypoglycaemia	n/a	n/a	n/a	1.75 [1.62, 1.89]	<0.0001	1.54 [1.45, 1.65]	<0.0001	1.29 [1.19, 1.40]	<0.0001
Hyperbilirubinaemia	n/a	n/a	n/a	1.04 [0.97, 1.11]	0.3087	1.13 [1.08, 1.19]	<0.0001	1.20 [1.13, 1.28]	<0.0001
Hypothermia	n/a	n/a	n/a	2.38 [1.87, 3.03]	<0.0001	2.20 [1.79, 2.71]	<0.0001	1.77 [1.34, 2.34]	<0.0001
Polycythaemia	n/a	n/a	n/a	1.87 [1.45, 2.42]	<0.0001	1.93 [1.57, 2.37]	<0.0001	1.78 [1.36, 2.34]	<0.0001

aOR adjusted odds ratio, *IRDS* infant respiratory distress syndrome, *BPD* bronchopulmonary dysplasia, *IVH* intraventricular haemorrhage, *cPVL* cystic periventricular leukomalacia, *NEC* necrotizing enterocolitis, *ROP* retinopathy of prematurity, *HIE* hypoxic ischemic encephalopathy

[a] 'Extra' SGA refers to infants who were classified as SGA according to the prescriptive birthweight standard only

[b] Adjusted for infant's gender (reference category 'female'), ethnicity ('Dutch/Caucasian'), maternal age ('15–35 years'), socioeconomic status ('middle class'), parity ('multipara'), multiple pregnancy ('singleton') and congenital malformations ('absent') if significantly associated at the 0.05 level (two-tailed Wald test)

[c] Includes both perinatal infections and nosocomial infections. Overall, 7.04 % of the pregnancies was complicated by prolonged premature rupture of membranes (>24 h). In preterm infants, this percentage was higher (24^{+0}–25^{+6} weeks: 22.6 %; 26^{+0}–27^{+6} weeks: 17.1 %; 28^{+0}–31^{+6} weeks: 16.5 %; 32^{+0}–36^{+6} weeks: 15.3 %) and comparatively more cases occurred in the non-SGA groups

of SGA infants at risk. Interestingly, we found that the risk of having IVH was significantly decreased in infants aged 28^{+0}–31^{+6} weeks, whereas Ferdynus found an increased risk [13]. Results from other studies concerning the association between SGA and adverse outcomes are similarly inconsistent. It is likely that these inconsistencies arise at least partially from differences in study methodologies, for example differences in SGA definition, sample size and GA range [39]. Despite the large sample size ($n = 127{,}584$), Ferdynus et al. included relatively few premature infants. As a result, the statistical power of this study was limited and other results did not reach statistical significance [13].

Birthweight standards are inexpensive and convenient tools that are used in clinical practise to identify infants who exhibit signs of IUGR. As such, they can be viewed as screening tools, the performance of which is routinely expressed in

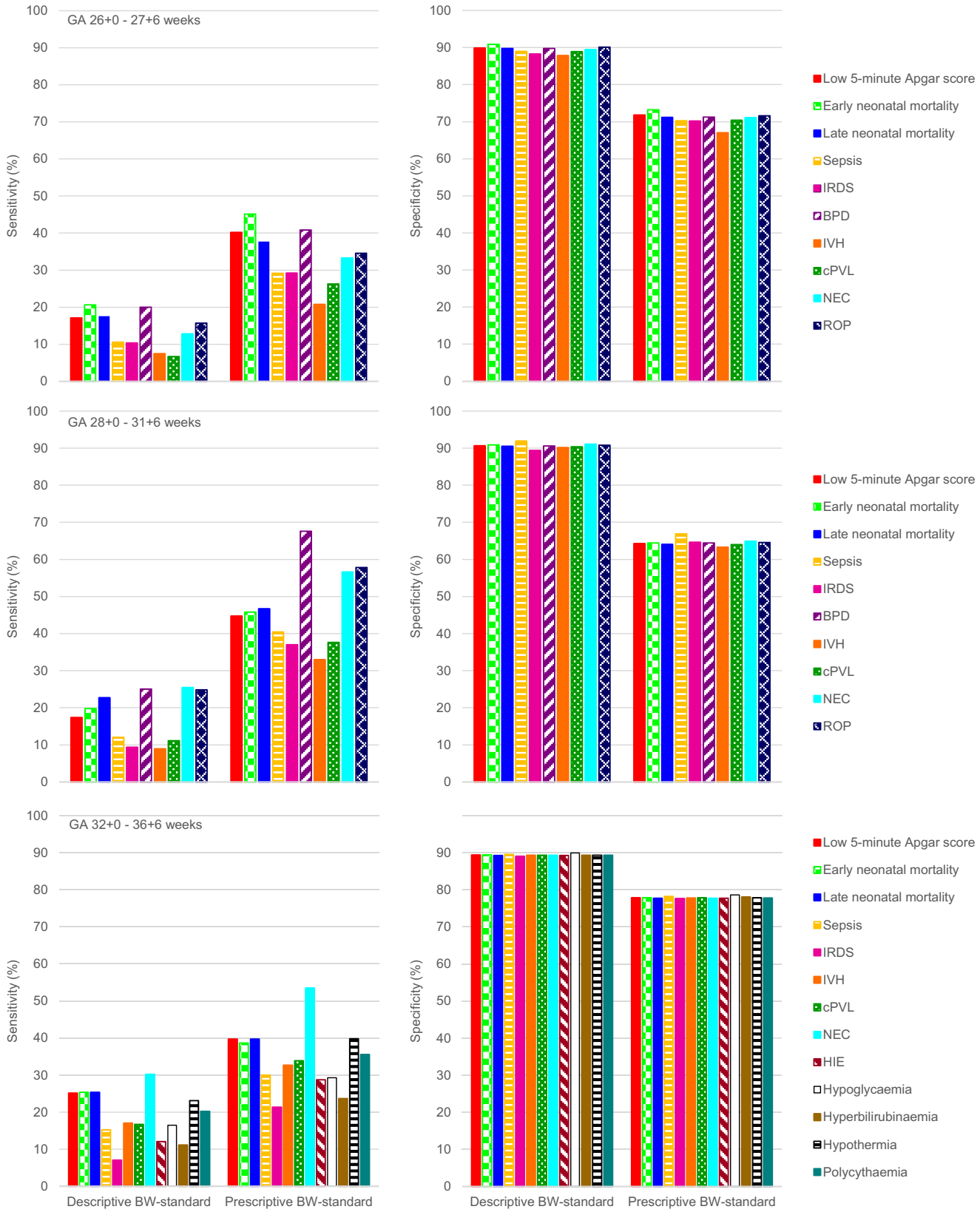

Fig. 3 Test characteristics for separate GA strata. *BW* birthweight

terms of sensitivity and specificity [18]. The results presented in this study do not indisputably indicate that either one of the two standards performs best. While the prescriptive birthweight standard was significantly more sensitive, the

descriptive birthweight standard was significantly more specific. Ideally, a screening instrument has both high sensitivity and specificity. In this particular setting, however, a false-positive result (i.e. low specificity) does not result in unnecessary or potentially harmful treatment, whereas a false-negative result (i.e. low sensitivity) means an increased risk of adverse neonatal outcomes is not acknowledged. All considered, the decrease in specificity may not be negligible but was deemed less clinically relevant.

As stated in the Introduction, the (negative) influence of IUGR is not only visible in early life but extends far into adulthood. The so-called developmental origins hypothesis postulates that undernutrition during foetal life and infancy permanently changes gene expression and thereby establishes body composition, functional capacity, setting of hormones and metabolism and responses to adverse environmental influences later in life. The result is an increased risk of cardiovascular disease, hypertension, stroke and type 2 diabetes mellitus [1]. Birthweight, as the ultimate outcome and quantifiable result of foetal growth, might be able to function as a starting point for risk assessment both early and later in life.

Strengths and limitations

One of the principal strengths of our study was the use of a large population. This allowed us to investigate the association between SGA and many adverse neonatal outcomes, and across separate GA strata. Because we used the same statistical method to calculate birthweight percentiles for both populations, we can attribute differences between the two standards to actual differences in birthweight distribution, thus demonstrating the effect of applying restrictive inclusion criteria.

Previous studies showed that despite the increased number of SGA infants, the association with adverse neonatal outcomes was still significant [13, 42]. By investigating extra SGA neonates as a separate group, we demonstrated the additional benefit of the prescriptive birthweight standard. Furthermore, we acknowledged the trade-off between identifying more SGA infants truly at risk, yet simultaneously increasing the number of false-positive results [18].

The biggest challenge was to define the characteristics of a healthy subpopulation. Our choice of exclusion criteria was based on an extensive literature search. Its application was to some (unknown) extent constrained by lacunas in the database; information on some of the risk factors we identified was either limited (e.g. smoking) or not available at all. If more stringent exclusion criteria could have been applied, this would have led to more removals, but not necessarily greater precision in the calculation of standards [26, 36]. Conditions that might have caused the opposite effect on growth (i.e. macrosomia) were not excluded. Although this may have caused some bias towards higher birthweights, since the upper percentiles were not influenced by the exclusion of IUGR infants, we assumed that excluding macrosomic infants would not effect the lower percentiles. Because the exclusion criteria concerned potential risk factors, it is possible that some healthy infants were excluded as well. However, birthweight distributions are more likely influenced by the inappropriate inclusion of an unhealthy infant than by the inadvertent exclusion of a healthy infant [33, 36].

Our study was limited by the need to select a sample group, to assure that infants without recorded adverse outcomes could indeed be considered healthy, instead of falsely being labelled as such due to incomplete (paediatric) registration. Aside from a distortion of the true ratios of hospital deliveries versus home deliveries and midwife-assisted in-hospital deliveries, baseline characteristics did not differ between women who were included in the sample and those who were not (results not shown). The use of retrospective data has the advantage of enabling the inclusion of a large number of infants. An obvious disadvantage is that the data were not collected with this particular study in mind and in the interpretation of our findings this has to be kept in mind. For example, the database did not contain data about antenatal corticosteroid treatment, which may be a confounder of the association between SGA and adverse neonatal outcomes. Fortunately, these limitations may not be of major consequence, since our analyses were executed on a matched sample and we were interested in the magnitude of the associations relative to each other, aiming to find the 'relatively' best standard. Because the current analyses were executed on a sample group of which the majority of infants also served as the reference population for the (two) birthweight standards, our findings should be confirmed in an independent population. It is our intention to repeat the analyses when we have a sufficiently large independent sample.

Conclusion

To our knowledge, this study is the largest study comparing descriptive versus prescriptive birthweight standards. We demonstrated that birthweight standards, defined by a subpopulation free of risk factors for IUGR, could improve identification of infants born SGA and at risk of several adverse neonatal outcomes. Clinical management should not be based on birthweight standards alone, and the results of this study endorse this statement by proving that neither one of the two standards performs great. However, for certain epidemiological purposes, it is also important that a standard correctly identifies infants born SGA, who may be at increased risk of adverse outcomes. Descriptive references mistakenly classify infants as non-SGA, which might, as a result, obscure or enhance the true association between potential risk factors and SGA, or between SGA and (long-term) adverse outcomes.

Improved understanding of these associations will aid in clinical decision-making and care and might ultimately improve both short-term and long-term outcome.

Our results provide the rationale for the development of prescriptive birthweight standards for the Dutch population. We believe the results also justify promotion of this concept in other countries. Above all, we feel that clinicians and researchers should be aware of the difference between prescriptive and descriptive birthweight standards, and the implications their use can have for both clinical practise and research outcomes.

Acknowledgments We would like to thank The Netherlands Perinatal Registry (PRN) for granting us unlimited access to the required data.

Authors' contributions RvL and MH conceived the idea of the manuscript. LH, CH, MH, and RvL were responsible for the study concept and design. LH, CH, and HS analysed the data. LH wrote the initial draft of the manuscript, which was revised by LH, CH, MH, HS, and RvL. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Compliance with ethical standards

Ethical standards The manuscript does not contain clinical studies or patient data.

Conflict of interest The authors declare that they have no conflicts of interest.

Funding No funding was secured for this study.

References

1. Barker DJ (2006) Adult consequences of fetal growth restriction. Clin Obstet Gynecol 49:270–283
2. Battaglia FC, Lubchenco LO (1967) A practical classification of newborn infants by weight and gestational age. J Pediatr 71:159–163
3. Bell MJ, Ternberg JL, Feigin RD, Keating JP, Marshall R, Barton L, Brotherton T (1978) Neonatal necrotizing enterocolitis. Therapeutic decisions based upon clinical staging. Ann Surg 187:1–7
4. Bertino E, Milani S, Fabris C, De Curtis M (2007) Neonatal anthropometric charts: what they are, what they are not. Arch Dis Child Fetal Neonatal Ed 92:F7–F10
5. Burkhardt T, Schäffer L, Zimmermann R, Kurmanavicius J (2008) Newborn weight charts underestimate the incidence of low birthweight in preterm infants. Am J Obstet Gynecol 199:139.e1–139.e6
6. Campbell MK, Cartier S, Xie B, Kouniakis G, Huang W, Han V (2012) Determinants of small for gestational age birth at term. Paediatr Perinat Epidemiol 26:525–533
7. Clausson B, Gardosi J, Francis A, Cnattingius S (2001) Perinatal outcome in SGA births defined by customised versus population-based birthweight standards. BJOG 108:830–834
8. Cole TJ (1990) The LMS method for constructing normalized growth standards. Eur J Clin Nutr 44:45–60
9. Committee on Obstetric Practice, ACOG; American Academy of Pediatrcis; Committee on Fetus and Newborn, ACOG (2006) ACOG Committee Opinion Number 333, May 2006 (replaces No. 174, July 1996): the Apgar score. Obstet Gynecol 107:1209–1212
10. Damodaram M, Story L, Kulinskaya E, Rutherford M, Kumar S (2011) Early adverse perinatal complications in preterm growth-restricted fetuses. Aust N Z J Obstet Gynaecol 51:204–209
11. De Vries LS, Eken P, Dubowitz LM (1992) The spectrum of leukomalacia using cranial ultrasound. Behav Brain Res 49:1–6
12. Ferdynus C, Quantin C, Abrahamowicz M, Burguet A, Sagot P, Gouyon JB (2013) Comparison of the ability of alternative birthweight and fetal weight standards to identify preterm newborns at increased risk of perinatal death. BJOG 120:1456–1464
13. Ferdynus C, Quantin C, Abrahamowicz M, Platt R, Burguet A, Sagot P, Binquet C, Gouyon JB (2009) Can birth weight standards based on healthy populations improve the identification of small-for-gestational-age newborns at risk of adverse neonatal outcomes? Pediatrics 123:723–730
14. Gardosi J, Chang A, Kalyan B, Sahota D, Symonds EM (1992) Customised antenatal growth charts. Lancet 339:283–287
15. Giedion A, Haefliger H, Dangel P (1973) Acute pulmonary X-ray changes in hyaline membrane disease treated with artificial ventilation and positive end-expiratory pressure (PEP). Pediatr Radiol 1:145–152
16. Gilbert WM, Danielsen B (2003) Pregnancy outcomes associated with intrauterine growth restriction. Am J Obstet Gynecol 188:1596–1599, discussion 1599–1601
17. Hawass NE (1997) Comparing the sensitivities and specificities of two diagnostic procedures performed on the same group of patients. Br J Radiol 70:360–366
18. Hemming K, Bonellie S, Hutton JL (2011) Fetal growth and birthweight standards as screening tools: methods for evaluating performance. BJOG 118:1477–1483
19. Hendrix N, Berghella V (2008) Non-placental causes of intrauterine growth restriction. Semin Perinatol 32:161–165
20. Hutcheon JA, Zhang X, Platt RW, Cnattingius S, Kramer MS (2011) The case against customised birthweight standards. Paediatr Perinat Epidemiol 25:11–16
21. International Committee for the Classification of Retinopathy of Prematurity (2005) The international classification of retinopathy of prematurity revisited. Arch Ophthalmol 123:991–999
22. Joseph KS, Kramer MS, Allen AC, Mery LS, Platt RW, Wen SW (2001) Implausible birth weight for gestational age. Am J Epidemiol 153:110–113
23. Karlberg J, Albertsson-Wikland K (1995) Growth in full-term small-for-gestational-age infants: from birth to final height. Pediatr Res 38:733–739
24. Kramer MS (1987) Determinants of low birth weight: methodological assessment and meta-analysis. Bull World Health Organ 65:663–737
25. Lackman F, Capewell V, Richardson B, daSilva O, Gagnon R (2001) The risks of spontaneous preterm delivery and perinatal mortality in relation to size at birth according to fetal versus neonatal growth standards. Am J Obstet Gynecol 184:946–953
26. Lawrence C, Fryer JG, Karlberg P, Niklasson A, Ericson A (1989) Modelling of reference values for size at birth. Acta Paediatr Scand Suppl 350:55–69
27. Leitner Y, Fattal-Valevski A, Geva R, Eshel R, Toledano-Alhadef H, Rotstein M, Bassan H, Radianu B, Bitchonsky O, Jaffa AJ, Harel S (2007) Neurodevelopmental outcome of children with intrauterine growth retardation: a longitudinal, 10-year prospective study. J Child Neurol 22:580–587
28. Løhaugen GC, Østgård HF, Andreassen S, Jacobsen GW, Vik T, Brubakk AM, Skranes J, Martinussen M (2013) Small for gestational age and intrauterine growth restriction decreases cognitive function in young adults. J Pediatr 163:447–453

29. Marconi AM, Ronzoni S, Bozzetti P, Vailati S, Morabito A, Battaglia FC (2008) Comparison of fetal and neonatal growth curves in detecting growth restriction. Obstet Gynecol 112:1227–1234
30. Maulik D (2006) Fetal growth restriction: the etiology. Clin Obstet Gynecol 49:228–235
31. McCowan L, Horgan RP (2009) Risk factors for small for gestational age infants. Best Pract Res Clin Obstet Gynaecol 23:779–793
32. McIntire DD, Bloom SL, Casey BM, Leveno KJ (1999) Birth weight in relation to morbidity and mortality among newborn infants. N Engl J Med 340:1234–1238
33. Niklasson A, Ericson A, Fryer JG, Karlberg J, Lawrence C, Karlberg P (1991) An update of the Swedish reference standards for weight, length and head circumference at birth for given gestational age (1977–1981). Acta Paediatr Scand 80:756–762
34. Pallotto EK, Kilbride HW (2006) Perinatal outcome and later implications of intrauterine growth restriction. Clin Obstet Gynecol 49:257–269
35. Papile LA, Burstein J, Burstein R, Koffler H (1978) Incidence and evolution of subependymal and intraventricular hemorrhage: a study of infants with birth weight less than 1,500 gm. J Pediatr 92:529–534
36. Rooth G, Meirik O, Karlberg P (1985) Estimation of the "normal" growth of Swedish infants at term. Preliminary report. Acta Paediatr Scand Suppl 319:76–79
37. Sarnat HB, Sarnat MS (1976) Neonatal encephalopathy following fetal distress. A clinical and electroencephalographic study. Arch Neurol 33:696–705
38. Brouwers HAA, Bruinse HW, Dijs-Elsinga J, van Huis AM, de Miranda E, Ravelli ACJ, Tamminga P, de Winter GA (2013) Stichting Perinatale Registratie Nederland Grote Lijnen 1999–2012 (The Netherlands Perinatal Registry Trends 1999–2012). The Netherlands Perinatal Registry, Utrecht
39. Tsai LY, Chen YL, Tsou KL, Mu SC, Taiwan Premature Infant Developmental Collaborative Study Group (2015) The impact of small-for-gestational-age on neonatal outcome among very-low-birth-weight infants. Pediatr Neonatol 56:101–107
40. Visser GH, Eilers PH, Elferink-Stinkens PM, Merkus HM, Wit JM (2009) New Dutch reference curves for birthweight by gestational age. Early Hum Dev 85:737–744
41. Yu VY, Upadhyay A (2004) Neonatal management of the growth-restricted infant. Semin Fetal Neonatal Med 9:403–409
42. Zaw W, Gagnon R, da Silva O (2003) The risks of adverse neonatal outcome among preterm small for gestational age infants according to neonatal versus fetal growth standards. Pediatrics 111:1273–1277
43. Zhang J, Merialdi M, Platt LD, Kramer MS (2010) Defining normal and abnormal fetal growth: promises and challenges. Am J Obstet Gynecol 202:522–528
44. Zhang X, Mumford SL, Cnattingius S, Schisterman EF, Kramer MS (2010) Reduced birthweight in short or primiparous mothers: physiological or pathological? BJOG 117:1248–1254