BMJ Open

Applying measures of discriminatory accuracy to revisit traditional risk factors for being small for gestational age in Sweden: a national cross-sectional study

Journal:	BMJ Open
Manuscript ID:	bmjopen-2014-005388
Article Type:	Research
Date Submitted by the Author:	02-Apr-2014
Complete List of Authors:	Juárez, Sol; Lund University, Centre for Economic Demography Wagner, Phillip; Lund University, Unit of Social Epidemiology, Faculty of Medicine Merlo, Juan; Lund University, Unit of Social Epidemiology, Faculty of Medicine
Primary Subject Heading :	Reproductive medicine
Secondary Subject Heading:	Reproductive medicine
Keywords:	Fetal medicine < OBSTETRICS, Prenatal diagnosis < OBSTETRICS, Maternal medicine < OBSTETRICS



Sol Pía Juárez

Centre for Economic Demography, Lund University Unit of Social Epidemiology, Department of Clinical Sciences, Lund University, Malmö P.O. Box 7083 SE-220 07 Lund, Sweden Tel. +46 46-222 00 00 Sol.juarez@ekh.lu.se

Phillip Wagner

Unit of Social Epidemiology, Department of Clinical Sciences, Lund University, Malmö Skåne University Hospital (SUS Malmö) Jan Waldenströms gata 35 SE- 214 21 Malmö, Sweden

Juan Merlo

Unit of Social Epidemiology, Department of Clinical Sciences, Lund University, Malmö Skåne University Hospital (SUS Malmö) Jan Waldenströms gata 35 SE- 214 21 Malmö, Sweden

Keywords: risk factors, SGA, discriminatory accuracy

Abstract

Objectives Small for gestational age (SGA) is considered an indicator of intrauterine growth restriction, and multiple maternal and newborn characteristics have been identified as risk factors for SGA. This knowledge is mainly based on measures of average association—like the odds ratio (OR)—that quantify differences in average risk between exposed and unexposed groups. Nevertheless, average associations do not assess the discriminatory accuracy of the risk factors (ie, its ability to discriminate the babies who will develop SGA from those that will not). Therefore, applying measures of discriminatory accuracy rather than measures of association only, our study revisits known risk factors of SGA and discusses their role from a public health perspective.

Design Cross-sectional study. We measured maternal (ie, smoking, hypertension, age, marital status, education) and delivery (ie, sex, gestational age, birth order) characteristics and performed logistic regression models to estimate both ORs and measures of discriminatory accuracy, like area under the receiver operating characteristic curve (AU-ROC) and the net reclassification improvement.

Setting Data were obtained from the Swedish Medical Birth Registry.

Participants Our sample included 731 989 babies born during 1987–1993.

Results We replicated the expected associations. For instance, smoking (OR=2.57) and hypertension (OR=4.02) were strongly associated with SGA. However, both variables show a very small discriminatory accuracy (AU-ROC \approx 0.5). The discriminatory accuracy increased, but remained unsatisfactorily low (AU-ROC=0.6), when including all variables studied in the same model.

Conclusions We found that traditional risk factors for SGA alone or in combination have a low accuracy for discriminating babies with SGA from those without SGA. A proper understanding of these findings is of fundamental relevance to address future research and to design policymaking recommendations in a more informed way.

- Our study emphasises the use and interpretation of measures of discriminatory accuracy (ie, capacity to distinguish between SGA and non-SGA babies) when evaluating risk factors.
- We confirm statistical associations between maternal and newborn characteristics and risk for SGA, but we underline that the discriminatory capacity of all the risk factors studied was very low.
- This low discriminatory capacity suggests that we know very little about the determinants of SGA in the population and that more effort should be paid to understanding individual heterogeneity of effects.
- Our finding is of fundamental relevance to address future research and to design • policymaking recommendations in a more informed way.

Small for gestational age (SGA) is commonly identified as a proxy of intrauterine growth restriction (IUGR).¹ This disorder has been associated with neonatal mortality and morbidities² as well as with major medical problems across the life course, such as a higher risk of neurodevelopmental impairments,^{3,4} autism,⁵ schizophrenia,⁶ impaired cognitive function,⁷ coeliac disease in boys,⁸ and reduced bone mass during early infancy,⁹ as well as Barrett's oesophagus and oesophagitis^{10,11} and others.^{12,13} Therefore, the identification of maternal and newborn characteristics (denominated as 'risk factors' in the rest of this work) associated with an increased risk for SGA is of obvious relevance in public health and preventive medicine.

Two reviews, one from 1987¹⁴ and other from 2009,¹⁵ pointed out that SGA is associated with a broad number of genetic, obstetric, demographic, and socioeconomic factors as well as maternal morbidities and toxic exposures before and during pregnancy. However, the identification of these risk factors has been exclusively based on measures of average association (eg, odds ratio) but without considering their accuracy for discriminating babies with, from those without, SGA. Indeed, it is a common practice to use measures of average association to gauge the ability of a factor to discriminate future cases of disease.¹⁶ For example, it is known that maternal hypertension during pregnancy gives a fourfold increased risk of delivering an SGA baby. Therefore, this variable is implicitly used as a predictive test to classify who will and who will not deliver an SGA baby. However, in spite of this popular belief, measures of association alone are inappropriate for this discriminatory purpose.^{16–23}

Although measures of discriminatory accuracy are extensively applied in other fields of epidemiology like, for instance, the identification of new biomarkers for cardiovascular diseases,¹⁷⁻²⁰ these measures are still unusual in public health and epidemiology.²¹ In fact, as far we know, they have never been explicitly used to formally revisit established maternal and newborn risk factors for SGA.

With this background our study aims to revisit the role of current risk factors for SGA in public health. We do it in two steps. Firstly, using measures of average association, we aim to replicate previous findings and identify maternal and newborn risk factors for SGA. Secondly, we apply measures of discriminatory accuracy to assess the ability of those risk factors (alone or in

combination) to discriminate babies with, from those without, SGA in the whole population and in different subgroups according to gestational age.

DATA AND METHODS

Study design, setting, and participants

This is a cross-sectional study based on a population-based register. We identified all the 811 599 babies born alive and recorded at the Swedish Medical Birth Registry (MBR) between 1 January 1987 and 31 December 1993. The MBR collects detailed and standardised information on nearly all pregnancies in Sweden culminating in delivery.^{22,23} Using a unique personal identification number, the Swedish authorities (National Board of Health and Welfare and Statistics Sweden) linked the MBR to the Register of the Total Population and the Swedish 1990 population census and created a research database. This database was delivered to us without the personal identification numbers to protect the anonymity of the subjects. The Regional Ethics Review Board in southern Sweden approved the construction of the database.

For the purposes of our study, we selected singletons, because it is known that multiple births (n=19 167) have a different intrauterine growth pattern from gestational weeks 28-30.²⁴ We excluded 13 539 babies born with significant congenital anomalies according to the MBR. Following previously established criteria,²⁵ we also excluded babies with inconsistent information on birth weight according to gestational age (n=9195) and babies weighing less than 500 g (n=51) as well as 15 observations with missing information on maternal age and birth order. The final sample contained 768 059 babies. Thereafter, we stratified the population by gestational age into pre-term (<37 gestational weeks) term (\geq 37 and <42 gestational weeks) and post-term babies (>42 gestational weeks) (Figure 1).

[Figure 1 about here]

Variables

The outcome variable combined birth weight and gestational age to dichotomise as being SGA or not, and using the last category as the reference. This variable was available at the MBR, where

it is routinely calculated following standard intrauterine growth curves.²⁶ Infants were defined as SGA if they weighed less than 2 standard deviations below the expected birth weight for gestational age and gender, according to a Swedish intrauterine growth curve.³¹

In our analyses we included child and maternal characteristics that are known to be associated with low birth weight and SGA.

As child characteristics we used $sex^{14,27}$ and *birth order*^{28,29} classified into three categories (ie, firstborn, second, and third or more). Among maternal characteristics we included *education*,^{30,31} categorised into low (primary education or less), middle (secondary school), and high education (graduate and PhD); *marital status*,^{32,33} categorised into single, widowed, or divorced, and married or cohabiting; and *maternal age at delivery*,³⁴⁻³⁶ categorised into four groups (ie, <20, 20–24, 25–34, and >35 years old), as well as information on *smoking habits*,³⁷⁻⁴⁰ categorised into non-smoking, light smoking (fewer than 9 cigarettes per day), heavy smoking (more than 9 cigarettes per day), and missing information. Finally, we included information about presence of *hypertension during pregnancy* (yes vs. no),^{15,41} and *maternal origin*, classified as being born in Sweden or not.⁴²

Statistical methods

To examine the average association between, on the one hand, the categorical variables mentioned above, and on the other, being SGA, we simply calculated odds ratios and 95% confidence intervals obtained from logistic regression analyses.

The discriminatory accuracy of a risk factor is better appraised by measuring *the true positive fraction* (TPF) and *the false positive fraction* (FPF). For a dichotomous risk factor, the TPF expresses the probability of being exposed to the risk factor when the SGA occurs (ie, cases that are exposed to the risk factor), and the FPF indicates the probability of being exposed to the risk factor. In the ideal scenario the TPF should be 1 and the FPF should be 0, even if a lower TPF or a higher FPF. For instance, if the identification of the risk factors conveys pharmacological treatment, we should try to keep the FPF as low as possible.

For the evaluation of the discriminatory accuracy of the combination of risk factors within a risk score (ie, predicted probability) we obtained *the receiver operating characteristic* (ROC) *curve*.

BMJ Open

The ROC curve is constructed by plotting the TPF against the FPF for different risk score thresholds.^{16,43,44} A traditional measure of discriminatory accuracy is the *area under the ROC curve* (AU-ROC) or C statistic.^{16,43,45-47} The AU-ROC extends from 0.5 to 1.0. An AU-ROC=0.5 means that the discriminatory accuracy of the candidate risk factor is similar to that obtained by flipping a coin. That is, a risk factor with an AU-ROC=0.5 is useless. An AU-ROC=1.0 means complete accuracy.

In series of simple logistic regression models we identified the single variables with the highest discriminatory accuracy. Thereafter, we performed models combining several variables at the same time and assessed the change of discriminatory accuracy when adding covariates to an initial, simpler model. We appraised the incremental value of a model by the difference between AU-ROCs.

We performed the analyses in the whole population, stratifying by gestational age (ie, preterm, term, and post-term). We performed the statistical analyses using STATA 12.0 (College Station, TX: StataCorp LP) and SPSS 21.0 (Armonk, NY: IBM Corp.).

RESULTS

Table 1 shows the maternal and individual characteristics of the population of newborns by SGA status. We see that SGA is much more prevalent among preterm babies (10.14%) than among term (1.87%) and post-term (3.03%) babies. Females show higher prevalence of SGA than males among preterm, and slightly lower prevalence among those born post-term. Regardless of gestational age, firstborns had a higher risk of SGA than their siblings. SGA was more frequent in mothers younger than 20 years of age; among divorced, widowed, and single women; and among those who were born outside Sweden and those with low educational achievement. In babies born SGA, hypertension was more frequent among preterm than among post-term.

BMJ Open: first published as 10.1136/bmjopen-2014-005388 on 30 July 2014. Downloaded from http://bmjopen.bmj.com/ on June 12, 2025 at Department GEZ-LTA Erasmushogeschool . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

				SGA	stratified b	y gestatio	nal age	
	SGA		Term		Preterm)	Post-ter	m
	768 059		676 961		36 080		55 018	
	(2.3%)		(1.9%)		(10.1%)		(3.0%)	
	Ν	%	Ν	%	Ν	%	Ν	%
Newborn sex								
Male	393 538	2.33	343 549	1.86	19 635	9.20	30 354	3.24
Female	374 521	2.34	333 412	1.87	16 445	11.26	24 664	2.78
Birth order								
First	325 326	3.30	279 426	2.66	18 796	11.67	27 104	4.10
Second	268 998	1.63	242 276	1.31	9 774	8.79	16 948	2.01
<u>></u> Third	173 735	1.63	155 259	1.30	7,510	8.07	10 866	1.97
Maternal age								
<20	40 735	3.06	35 367	2.49	2485	9.93	2883	4.30
20–24	156 736	2.39	138 232	1.91	7842	9.82	10 662	3.24
25–34	481 689	2.19	425 980	1.76	20 796	9.59	34 913	2.89
>35	88 899	2.69	77 382	2.08	4957	12.04	6 560	2.90
Hypertension	3166	8.69	2648	5.44	383	32.11	135	5.93
Smoking								
Non-smoker	541 962	1.77	479 851	1.37	22 603	9.14	39 508	2.42
Light smoker	109 799	3.62	76 327	3.05	5884	11.86	7588	4.4(
Heavy smoker	67 454	4.41	58 790	3.79	4234	11.76	4430	5.64
Missing	48 844	2.90	41 993	2.12	3359	11.79	3492	3.78
Marital status								
Married or								
cohabiting	385 173	2.02	342 330	1.61	16,50	9.36	26 343	2.69
Divorced or								
widowed	27 151	3.03	23 475	2.43	1706	10.67	1970	3.55
Single	355 735	2.63	311 156	2.10	17 874	10.81	26 705	3.33
Maternal origin								-
Born in Sweden	663 043	2.25	584 683	1.79	30 745	10.19	47 615	2.79
Not born in	05 202	2.00	04.000	2 2 2	4740	40.07	c c 2 2	4
Sweden	95 393	2.83	84 022	2.30	4/49	10.07	6 622	4.27
Missing	9623	2.67	8256	3.02	586	8.02	/81	7.30
Maternal educatio	on 101.000	4.67	400.400	4 = 2	7067	10.00	40.044	a
University	181 083	1.95	160,102	1.52	7367	10.26	13 614	2.57
Secondary	400 363	2.22	353 813	1.77	18 457	9.96	28 093	2.74
Primary	140 186	3.21	122 919	2.31	7631	10.78	9636	3.84
Missing	46 427	2.34	40 127	2.67	2625	9.22	3675	4.82

Table 1. Prevalence of SGA in the whole population of babies and in strata of gestational age in

Table 2 indicates that the risk for being SGA was similar in boys and girls. However, as expected, not being a firstborn reduced the risk of being SGA. With respect to maternal characteristics, mothers younger than 20 years and those 35 years and older had higher risk of delivering an SGA baby compared to 20- to 24-year-old mothers. Mothers who experienced hypertension during pregnancy had a higher risk of delivering SGA babies. Compared to non-smoker mothers, light and heavy smoker mothers had higher risk of delivering an SGA baby. Divorced and widowed mothers as well as single mothers were more likely to deliver an SGA baby than married and cohabiting mothers. Mothers with primary and secondary education had higher risk of delivering SGA babies than mothers with a university degree. Similarly, mothers who were not born in Sweden were at higher risk of delivering an SGA baby.

Table 2. Measures of association between offspring and maternal characteristics, and being small for gestational age (SGA), in the whole population of babies and in strata of gestational age in Sweden 1987–1991. Values are odds ratios (OR) and [95% confidence intervals (IC)]. Crude models

		SGA			TERM		SGA PRETERM			SGA POST-TERM			
	Unad	justed r	nodel	Unad	Unadjusted model			Unadjusted model			Unadjusted model		
	OR	IC-95%	6	OR	IC-95%	6	OR	IC-95%	6	OR	IC-95%	6	
Newborn sex													
Male (ref)													
Female	1.01	[0.98	1.04]	1.01	[0.97	1.04]	0.86	[0.84	0.88]	0.85	[0.77	0.94]	
Birth order													
First (ref)													
Second	0.48	[0.47	0.50]	0.49	[0.47	0.51]	0.60	[0.59	0.62]	0.48	[0.42	0.54]	
<u>></u> Third	0.49	[0.47	0.51]	0.48	[0.46	0.50]	0.72	[0.70	0.74]	0.47	[0.41	0.54]	
Gestational age	0.80	[0.79	0.81]	0.87	[0.85	0.88]	0.86	[0.85	0.87]	1.77	[1.61	1.95]	
Maternal age													
25–34 (ref)													
<20	1.41	[1.33	1.49]	1.42	[1.33	1.53]	1.44	[1.38	1.50]	1.51	[1.25	1.83]	
20–24	1.09	[1.05	1.13]	1.09	[1.04	1.14]	1.16	[1.13	1.19]	1.12	[0.99	1.27]	
>35	1.23	[1.18	1.29]	1.18	[1.12	1.25]	1.31	[1.27	1.36]	1.00	[0.86	1.17]	
Hypertension													
No (ref) vs. yes	4.02	[3.55	4.55]	3.05	[2.58	3.61]	2.73	[2.45	3.04]	2.02	[0.99	4.13]	
Smoking													
Non-smoker (ref)													
Light smoker	2.08	[2.01	2.16]	2.27	[2.17	2.37]	1.30	[1.26	1.33]	1.86	[1.64	2.12]	

Heavy smoker Missing	2.56 1.66	[2.46 [1.57	2.67] 1.76]	2.84 1.56	[2.70 [1.45	2.98] 1.67]	1.53 1.70	[1.48 [1.64	1.58] 1.76]	2.42 1 59	[2.10 [1 32	2.79] 1.92]
Marital status	1.00	[1.37	1.70]	1.50	[1.13	1.07]	1.70	[1.01	1.70]	1.55	[1.52	1.52]
Married or cohabit	inσ											
(ref)												
Divorced or widowed	1.52	[1.41	1.63]	1.52	[1.40	1.66]	1.51	[1.43	1.59]	1.33	[1.04	1.71]
Single	1.31	[1.27	1.35]	1.31	[1.27	1.36]	1.19	[1.17	1.22]	1.25	[1.13	1.38]
Maternal origin												
Born in Sweden												
(ref)												
Not born in	1 27	[1 21	1 3 2 1	1 30	[1 23	1 361	1 07	[1 0/	1 1 1 1	1 56	[1 37	1 77]
Sweden	1.27	[1.21	1.52]	1.50	[1.25	1.50]	1.07	[1.04	1.11]	1.50	[1.57	1.//]
Missing	1.66	[1.49	1.84]	1.71	[1.51	1.94]	1.35	[1.24	1.47]	2.74	[2.08	3.61]
Maternal												
education												
University (ref)												
Secondary	1.14	[1.10	1.19]	1.17	[1.12	1.23]	1.13	[1.10	1.17]	1.07	[0.94	1.22]
Primary	1.49	[1.42	1.56]	1.53	[1.45	1.62]	1.35	[1.31	1.39]	1.51	[1.30	1.76]
Missing	1.67	[1.57	1.77]	1.78	[1.66	1.92]	1.42	[1.36	1.49]	1.92	[1.59	2.31]

Figure 2 shows the values for the AU-ROC of the variables included in Table 2. Overall, their discriminatory accuracy was rather low. Newborn sex had the lower discriminatory accuracy. Hypertension, despite being the risk factor most strongly associated with SGA (OR 4.02), had a very low discriminatory accuracy (AU-ROC 0.51%). Birth order and smoking were the variables with highest accuracy (AU-ROC 0.59%).

[Figure 2 around here]

Figure 3 shows the AU-ROC for SGA of different risk factors after stratification by preterm, term, and post-term. As in the non-stratified analysis, the discriminatory accuracy of the variables was low. Smoking at term showed the highest discrimination (AU-ROC 0.60%).

[Figure 3 around here]

BMJ Open

Figure 4 shows that the discriminatory accuracy of the model A, including only birth order and smoking, was slightly improved (just 0.05 proportion units), when all variables were included in the full model B. Among preterm babies model B improved the discriminatory accuracy of the model by 0.1 proportion units, while this improvement was much lower among SGA term and SGA post-term babies.

[Figure 4 around here]

DISCUSSIONS

We were able to verify a number of recognised maternal and newborn risk factors for SGA. For instance, we found that smoking (OR=2.56), and especially maternal hypertension (OR=4.02), were 'strongly' associated with being SGA. However, even if the magnitude of the ORs was of a size normally considered as undoubtedly relevant in epidemiology, none of those traditional risk factors for SGA provided enough accuracy to discriminate babies with SGA from other babies. In fact, the AU-ROC for maternal hypertension was slightly higher than 0.5, which means that the accuracy of this variable for discriminating babies with SGA from those without SGA was rather similar to that obtained by flipping a coin. That is, we need to recognise that, although on average, mothers with hypertension were four times more likely to have an SGA baby, many mothers with hypertension. Our findings, therefore, seriously question the utility of maternal hypertension during pregnancy for planning strategies of prevention against SGA. This statement, however, does not mean that hypertension during pregnancy is irrelevant to understanding the origin of SGA, but rather that we need to determine who among hypertensive mothers is actually prone to deliver an SGA baby.

There is a tacit but fallacious belief that the discriminatory accuracy of a risk factor is high when it is supported by a 'strong' association (eg, an OR of 4, as in the case of maternal hypertension). However, for an association to be an accurate instrument for discrimination, it must be of a magnitude rarely identified in epidemiologic studies.^{16,48-50} Following our example, a low discriminatory accuracy only indicates that any attempt of intervention based on the existence of the risk factor will be inefficient and even inappropriate, because health professionals will unnecessarily treat many mothers. The decision to start an intervention should seriously take into

BMJ Open: first published as 10.1136/bmjopen-2014-005388 on 30 July 2014. Downloaded from http://bmjopen.bmj.com/ on June 12, 2025 at Department GEZ-LTA Erasmushogeschool . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

BMJ Open

 account the existence of important (physical or emotional) side effects in the false-positive women. That is, it is always important to consider the principle of *primum non nocere*.²¹

Compared to the other variables studied, birth order and smoking presented a higher discriminatory capacity. However, their discriminatory accuracy was still very low in absolute terms (AU-ROC \approx 0.59). Also, combining all the variables in the same model did not substantially increase the discriminatory accuracy (AU-ROC=0.69). In other words, our results indicate that we actually do not know so much about what determines being SGA.

The existence of a low discriminatory accuracy suggests that around the population average risk there is considerable individual heterogeneity. Therefore, a logical consequence should be to identify which women are most susceptible to the risk factors. Hence, we explored the discriminatory accuracy of the chosen risk factors in different strata of gestational age at birth. We found that the combination of all variables in the same model had a minor improvement for discriminatory accuracy among those born at term or preterm, but not among babies born postterm.

Our finding suggests the existence of individual heterogeneity of responses to some specific variables, so the discriminatory accuracy depended, for instance, on whether the baby was preterm, term, or post-term. In fact, smoking, birth order, maternal origin, and marital status had a lower discriminatory capacity among preterm than among babies at term. On the contrary, newborn sex and hypertension had a higher discriminatory accuracy among preterm than among term babies. In the same way, newborn sex and maternal origin had a lower discriminatory capacity among post-term babies, while for smoking we found the opposite relation. The variation of the magnitude of the discriminatory accuracy by gestational age at birth expresses the existence of individual heterogeneity.

In addition, the usual definition of SGA may be biased, and thereby contribute to reducing the discriminatory accuracy of the traditional risk factors. It is likely that all babies born at the 32^{nd} gestational week or earlier might to some extent be 'immature', and therefore a preterm baby with normal weight for gestational age has a different risk profile than a baby with normal weight for gestational age but born at term. Therefore, we adopted a relatively unusual approach, and stratified by gestational age.

Our findings have important research and policymaking implications. A possible reason for the

low discriminatory accuracy of many average associations is that average effects are a mixture of individual level effects and therefore mix interindividual heterogeneity (ie, some individuals respond intensively to the exposure, while others are resilient or might even respond in the opposite direction). The approach based on discriminatory accuracy understands average effects as an idealised mean value that does not necessarily represent the heterogeneity of individual effects.²¹ Some scholars prefer to conceive individual outcome as the expression of a stochastic phenomenon that is best estimated by the average risk using a probabilistic approach.⁵¹ Our understanding instead is that individual outcome reflects the interindividual heterogeneity of responses that can be potentially determined; lack of knowledge could be amended by a better understanding of individual responses.⁵² See elsewhere for a longer explanation of these ideas.^{21,53,54} From this perspective, reducing exposure to a risk factor would only be effective when acting on the susceptible, but not on the resilient, individuals. For instance, we need to better capture babies who suffer from IUGR, since, so far, we are incapable of distinguishing between babies who are constitutionally small from those who are pathologically growth restricted.¹ By stratifying between preterm, term, and post-term, we might be able to better approach the underlying heterogeneity.

From the policymaking perspective, our findings suggest that hitherto there has not been enough knowledge to identify any specific risk factor or combination of them that could discriminate with accuracy children with and without SGA status. Our findings support policymaking oriented to lifestyle modification, as according to the principle of *primum non nocere*,²¹ they have mostly positive consequences, even for 'false-positive' mothers. For instance, persuading women to quit smoking reduces the risk of SGA in some babies, but it improves general wellbeing in everyone. However, other risk factors with low discriminatory accuracy that lead to pharmacological treatment or screening might result in unnecessary side effects and cost. In the long run, an uncritical use of variables with low discriminatory accuracy may hinder the identification of pertinent risk factors and susceptible individuals and damage the scientific credibility of modern epidemiology.^{21,53,54}

Our conclusions are based on classical measures of discriminatory accuracy such as the AU-ROC curve. These measures have been criticised as insensitive to small changes in predicted individual risk.⁵⁵ Some authors propose more specific measures of reclassification, like the net reclassification improvement (NRI), and the integrated discrimination improvement (IDI).⁵⁶⁻⁵⁹

BMJ Open: first published as 10.1136/bmjopen-2014-005388 on 30 July 2014. Downloaded from http://bmjopen.bmj.com/ on June 12, 2025 at Department GEZ-LTA Erasmushogeschool . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

We applied NRI and IDI in a sensitivity analysis (results not shown in tables). For example, using NRI, we observed a reclassification of 30%. However, this figure does not add substantial information to our results, since NRI (as well as IDI) refers to the misclassification occurring all along the risk scale, instead of capturing the misclassification which takes place around the fixed threshold. Furthermore, the new NRI and IDI measures have also been criticised,⁶⁰ and some authors⁶¹ have explicitly advised against their use in common epidemiological practice because, unlike IDI and NRI, traditional measures of discrimination like the AU-ROC curve have the advantage that prognostic performance cannot be manipulated.⁶¹ Therefore, we preferred to quantify discriminatory accuracy by analysing ROC curves and AU-ROCs.

Our analyses are based on a national medical registry covering almost the entire population of residents in Sweden. Nearly all births are registered in the MBR, because giving birth at home is very unusual in Sweden. In addition, estimation of SGA is routinely calculated at the MBR following standard intrauterine growth curves.²⁶ However, our study also has a number of limitations. Because of lack of data, we could not assess many other variables identified in the literature as 'risk factors', such as genetic or nutritional factors.^{15,62} In spite of the quality of the MBR, the information regarding smoking is based on a self-reported questionnaire (anamnesis) administered by the midwife at the first antenatal visit (ie, between 10 and 12 gestational weeks), which to some extent might bias the result by including misclassification of exposure.⁶³ However, a study conducted in Sweden comparing self-reported nicotine exposure and plasma levels of cotinine in early and late pregnancy concluded that self-reported smoking information had acceptable validity.⁶⁴

Unfortunately, we could not identify those mothers who suffered from preeclampsia, for which the discriminatory accuracy concerning SGA may be higher than for hypertension in our model. Further analysis on this aspect is required. Another limitation of our study is that we calculated the discriminatory accuracy in the same sample used for constructing the predicted model. This procedure, however, might overestimate the discriminatory accuracy of the models, so the low discriminatory accuracy found may be an underestimation.

In conclusion, applying measures of discriminatory accuracy rather than measures of association only, our study revisits known risk factors of SGA and discusses their role from a public health perspective. We found that neither models including simple variables nor models including

several variables at the same time have a good discriminatory accuracy to discriminate babies with SGA from those without SGA. This finding is of fundamental relevance to address future research and to design policymaking recommendations in a more informed way.

As noted elsewhere,^{53,54} there is need of a new epidemiological approach that systematically provides information on the discriminatory accuracy and interindividual heterogeneity of effects and does not rely only on average measures of association.⁶⁵ In this line, new statistical methods like logic regression seem promising.^{66,67} A fundamental change is needed in the way traditional risk factors are currently interpreted in public health epidemiology. If the *discriminatory accuracy* of most classical risk factors is very low, what happens with the vast majority of recommendations given so far in epidemiology and public health? Are health professionals misleading the community by raising the alarm about risks that may be harmless for most individuals? What are the ethical repercussions of using risk factors with low discriminatory accuracy? Are there problems of inefficiency, medicalisation, and stigmatisation? We believe that these questions have a high significance for both the community and the future of public health research.

for .

BMJ Open: first published as 10.1136/bmjopen-2014-005388 on 30 July 2014. Downloaded from http://bmjopen.bmj.com/ on June 12, 2025 at Department GEZ-LTA Erasmushogeschool .

Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies

Acknowledgments

We would like to thank Dr. Karin Kallén for her very helping comments on a previous version of this paper. This work was supported by the Swedish Research Council (VR) (Dnr #2013-2484, PI Juan Merlo). The funders had no role in study design, data collection, and analysis; decision to publish; or preparation of the manuscript.

Authors' contributions

Merlo had the original idea of applying measures of discriminatory accuracy for the interpretation of risk factors and discussed it with Wagner and Juarez. Merlo and Juárez initiated the study. Merlo, Juárez, and Wagner contributed to the design of the study; Juárez performed the analyses under the supervision of Merlo and Wagner. Juárez wrote the first draft of the manuscript, and Merlo contributed to the writing of the final version. All authors made substantial contributions to the interpretation of the results and manuscript revision. All authors approved the final version of the manuscript.

Competing interest statement

The authors have declared that no competing interests exist.

REFERENCES

- **1.** Urquia ML, Ray JG. Seven caveats on the use of low birthweight and related indicators in health research. *J Epidemiol Community Health.* 2012;66(11):971-975.
- Grisaru-Granovsky S, Reichman B, Lerner-Geva L, et al. Mortality and morbidity in preterm small-for-gestational-age infants: a population-based study. *Am J Obstet Gynecol.* . 2012;206(2):150e151-e156.
- **3.** Streimish IG, Ehrenkranz RA, Allred EN, et al. Birth weight- and fetal weight-growth restriction: impact on neurodevelopment. *Early Hum Dev.* 2012;88(9):765-771.
- **4.** Eikenes L, Martinussen MP, Lund LK, et al. Being born small for gestational age reduces white matter integrity in adulthood: a prospective cohort study. *Pediatr Res.* . 2012;72(6):649-654.
- **5.** Moore GS, Kneitel AW, Walker CK, Gilbert WM, Xing G. Autism risk in small- and large-for-gestational-age infants. *Am J Obstet Gynecol.* 2012;206(4):314.
- **6.** Nielsen PR, Mortensen PB, Dalman C, et al. Fetal growth and schizophrenia: A nested case-control and case-sibling study. *Schizophr Bull.* 2012;12(): .
- 7. Løhaugen GC, Ostgård HF, Andreassen S, et al. Small for gestational age and intrauterine growth restriction decreases cognitive function in young adults. *J. Pediatr.* 2013; ():.
- **8.** Wingren CJ, Agardh D, Merlo J. Revisiting the risk of celiac disease in children born small for gestational age: a sibling design perspective. *Scand J Gastroenterol.* 2012;47(6):632-639.
- **9.** van de Lagemaat M, Rotteveel J, van Weissenbruch MM, Lafeber HN. Small-for-gestational-age preterm-born infants already have lower bone mass during early infancy. *Bone.* 2012;51(3):441-446.
- **10.** Forssell L, Cnattingius S, Bottai M, et al. Increased risk of Barrett's esophagus among Individuals born preterm or small for gestational age. *Clin Gastroenterol Hepatol.* . 2013; (): .
- **11.** Forssell L, Cnattingius S, Bottai M, Lagergren J, Ekbom A, Akre O. Risk of esophagitis among individuals born preterm or small for gestational age. *Clin Gastroenterol Hepatol.* . 2012;10(12):1369-1375.
- Dalla Pozza RD, Bechtold S, Putzker S, Bonfig W, Netz H, Schwarz HP. Young adults born small for gestational age: is reduced baroreceptor sensitivity a risk factor for hypertension? *Clin Cardiol.* . 2006;29(5):215-218.
- **13.** Crispi F, Figueras F, Cruz-Lemini M, Bartrons J, Bijnens B, Gratacos E. Cardiovascular programming in children born small for gestational age and relationship with prenatal signs of severity. *Am J Obstet Gynecol.* 2012;207(2):121.e121–121.e129.
- **14.** Kramer MS. Determinants of low birth weight: methodological assessment and meta-analysis. *Bull. World Health Organ.* 1987;65(5):663-737.
- **15.** McCowan L, Horgan RP. Risk factors for small for gestational age infants. *Best Pract Res Clin Obstet Gynaecol.* . 2009;23(6):779-793.
- **16.** Pepe MS, Janes H, Longton G, Leisenring W, Newcomb P. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *Am J Epidemiol.* May 1 2004;159(9):882-890.
- **17.** Zethelius B, Berglund L, Sundstrom J, et al. Use of multiple biomarkers to improve the prediction of death from cardiovascular causes. *New England Journal of Medicine*. May 15 2008;358(20):2107-2116.
- **18.** Cooney MT, Dudina AL, Graham IM. Value and limitations of existing scores for the assessment of cardiovascular risk: a review for clinicians. *J Am Coll Cardiol*. Sep 29 2009;54(14):1209-1227.
- **19.** Law MR, Wald NJ, Morris JK. The performance of blood pressure and other cardiovascular risk factors as screening tests for ischaemic heart disease and stroke. *J Med Screen.* 2004;11(1):3-7.

- 20. 21. 2014; In press. Stockholm: Centre of epidemiology;2002. J Soc Med. Jun 1990;18(2):143-148. 2008;8(244): . 2006;23(5):325-328. 2008;Part C(84):1-15. in Ireland. PLoS ONE. 2013;8(3):e57441. .
 - Melander O, Newton-Cheh C, Almgren P, et al. Novel and conventional biomarkers for prediction of incident cardiovascular events in the community. JAMA : the journal of the American Medical Association. Jul 1 2009;302(1):49-57. Merlo J. Multilevel analysis of individual heterogeneity: a fundamental critique of the current probabilistic risk factor epidemiology (invited commentary). American Journal of Epidemiology.
- 22. Facts about mothers, childbirth and newborn children. Medical Birth Registry, 1973 to 2000.
- 23. Cnattingius S, Ericson A, Gunnarskog J, Kallen B. A quality study of a medical birth registry. Scand
- 24. Cunningham FG. Williams Obstetrics. Third ed: McGraw-Hill; 2005.

1

2 3

4

5

6 7

8

9

10

11

12

13

14 15

16

17

18

19

20

21

22 23

24

25

26

27

28

29

30 31

32

33

34

35

36

37

38 39

40

41

42

43

44

45 46

47

48

49

50

51

52

53 54

55

56

- 25. Källén B. A birth weight for gestational age standard based on data in the Swedish Medical Birth Registry, 1985-1989. European Journal of Epidemiology. 1995;11(5):601-606.
- 26. Marsál K, Persson PH, Larsen T, Lilja H, Selbing A, Sultan B. Intrauterine growth curves based on ultrasonically estimated foetal weights. Acta Paediatr. 1996;85(7):843-848.
- 27. Alexander GR, Kogan MD, Himes JH. 1994-1996 U.S. singleton birth weight percentiles for gestational age by race, Hispanic origin, and gender. Matern. Child Health J. 1999;3(4):225-232.
- 28. Swamy GK, Edwards MJ, Gelfand A, James SA, Miranda ML. Maternal age, birth order, and race: differential effects on birthweight. J. Epidemiol. Community Health. 2012;66(2):136-142.
- 29. Elshibly EM, Schmalisch G. The effect of maternal anthropometric characteristics and social factors on gestational age and birth weight in Sudanese newborn infants. BMC Public Health.
- 30. Kramer MS, Séguin L, Lydos J, Goulet L. Socio-economic disparities in pregnancy outcome: why do the poor fare so poorly? *Paediatric and Perinatal Epidemiology*. 2000;14:194-210.
- 31. Auger N, Park AL, Harper S, Daniel M, Roncarolo F, Platt RW. Educational inequalities in preterm and term small-for-gestational-age birth over time. Ann Epidemiol. . 2012;22(3):160-167.
- 32. Shah PS, Zao J, Ali S, births. KSGoDopL. Maternal marital status and birth outcomes: a systematic review and meta-analyses. Matern. Child Health J. 2011;15(7):1097-1109.
- Castro T. Single motherhood and low birthweight in Spain: Narrowing social inequalitues in 33. health? Demographic Research. 2010;22(27):863-890.
- 34. Fraser AM, Brockert JE, Ward RH. Association of young maternal age with adverse repreductive outcomes. The New England Journal of Medicine. 1995;332(17):1113-1118.
- 35. Odibo A, Nelson D, Stamilo DM, Sehdev HM, Macones GA. Advanced maternal age is an independent risk factor for intrauterine growth restriction. American Journal of Perinatology.
- 36. Campbell MK, Cartier S, Xie B, Kouniakis G, Huang W, Han V. Determinants of small for gestational age birth at term. Paediatr Perinat Epidemiol. 2012;26(6):525-533.
- 37. Rogers JM. Tobacco and pregnancy: overview of exposures and effects. Birth defects research
- 38. Kabir Z, Daly S, Clarke V, Keogan S, Clancy L. Smoking Ban and Small-For-Gestational Age Births
- 39. Mitchell EA, Thompson JM, Robinson E, et al. Smoking, nicotine and tar and risk of small for gestational age babies. Acta Paediatr. 2002;91(3):323-328.
- 40. Juárez SP, Merlo J. Revisiting the Effect of Maternal smoking during pregnancy on offspring birthweight: a quasi-experimental sibling analysis in Sweden. PLoS ONE. 2013;8(4):e61734.
- 41. Catov JM, Nohr EA, Olsen J, Ness RB. Chronic hypertension related to risk for preterm and term small for gestational age births. *Obstet Gynecol.* 2008;112(2):290-296.

BMJ Open

2		
3	42.	Li X, Sundquist K, Sundquist J. Risks of small-for-gestational-age births in immigrants: A
4		nationwide epidemiological study in Sweden. Scand. J. Public Health. 2012;40:634-640.
5 6	43.	Pepe MS. The statistical evaluation of medical tests for classification and prediction. Oxford :
7		New York: Oxford University Press: 2003.
8	44	Zweig MH Broste SK Reinhart RA ROC curve analysis: an example showing the relationships
9		among serum linid and apolipoprotein concentrations in identifying natients with coronary
10		artery disease <i>Clin Chem</i> Aug 1002:28/8 Dt 1):1/25-1/28
11	45	Bene MS Japas II. Cu IM. Letter by Dana et al regarding article. "Use and misuse of the receiver
12	45.	Pepe IVIS, Janes H, Gu JW. Letter by Pepe et al regarding article, Ose and misuse of the receiver
13		operating characteristic curve in risk prediction". <i>Circulation</i> . Aug 7 2007;116(6):e132; author
14		reply e134.
15	46.	Royston P, Altman DG. Visualizing and assessing discrimination in the logistic regression model.
16		<i>Statistics in Medicine</i> . Oct 30 2010;29(24):2508-2520.
17	47.	Gerds TA, Cai T, Schumacher M. The performance of risk prediction models. <i>Biometrical journal</i> .
10		Biometrische Zeitschrift. Aug 2008;50(4):457-479.
20	48.	Boyko EJ, Alderman BW. The use of risk factors in medical diagnosis: opportunities and cautions.
20		Journal of clinical epidemiology. 1990;43(9):851-858.
22	49.	Wald NJ, Hackshaw AK, Frost CD. When can a risk factor be used as a worthwhile screening test?
23		<i>BMJ.</i> Dec 11 1999;319(7224):1562-1565.
24	50.	Khoury MJ, Newill CA, Chase GA. Epidemiologic evaluation of screening for risk factors:
25		application to genetic screening. American journal of public health. Oct 1985;75(10):1204-1208.
26	51.	Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction.
27		<i>Circulation</i> . Feb 20 2007:115(7):928-935.
28	52.	Zernicka-Goetz M. Huang S. Stochasticity versus determinism in development: a false
29	•	dichotomy? Nature reviews. Genetics. Nov 2010:11(11):743-744
31	53	Merlo I. Wagner P. The tyranny of the averages and the indiscriminate use of risk factors in
32	55.	nublic health: a call for revolution. European journal of enidemiology, 2013:28/1
33		Supplement):1/8
34	E /	Morle L Wagner D. Juarez S. Mulinari S. Hedblad P. The turanny of the averages and the
35	54.	indicariminate use of risk factors and nonulation attributable fractions in Dublis Usalth, the case
36		af agreenent beart diagon. Warking generation 2012 00 26. Unit for Cogicl Enidemic land
37		of coronary neart disease. Working paper version 2013-09-26. <i>Unit for Social Epidemiology,</i>
38		Department of Clinical Sciences, Faculty of Medicine, Lund University.
39		2013; <u>http://www.med.lu.se/english/klinvetmalmo/unit_for_social_epidemiology/abstracts_c_</u>
40		working papers c.
42	55.	Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction.
43		<i>Circulation</i> . Feb 20 2007;115(7):928-935.
44	56.	Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Evaluating the added predictive
45		ability of a new marker: From area under the ROC curve to reclassification and beyond. Statistics
46		<i>in Medicine</i> . Jan 30 2008;27(2):157-172.
47	57.	Pencina MJ, D'Agostino RB, Pencina KM, Janssens AC, Greenland P. Interpreting incremental
48		value of markers added to risk prediction models. Am J Epidemiol. Sep 15 2012;176(6):473-481.
49	58.	Pencina MJ, D'Agostino RB, Sr., Steyerberg EW. Extensions of net reclassification improvement
50		calculations to measure usefulness of new biomarkers. Stat Med. Jan 15 2011;30(1):11-21.
52	59.	Hilden J. On NRI. IDI. and "good-looking" statistics with nothing underneath. Epidemiology.
53		2014:25(2):265-267.
54	60.	Pene MS. Problems with risk reclassification methods for evaluating prediction models. Am I
55	501	Enidemiol Jun 1 2011:173(11):1327-1335
56	61	Hidden L Gerds T. Evaluating the impact of novel biomarkers: Do not rely on IDL and NPL
57	J 1.	Department of Riostatistics University of Congebragen Departs 2012 Personal Depart 12/09
58		Department of biostatistics, oniversity of copenhagen, Denniark. 2012, Research Report 12/08.
59		
60		

- **62.** Kramer MS. Determinants of low birth weight: methodological assessment and meta-analysis. *Bulletin of World Health Organization.* 1984;65(5):663-737.
- **63.** Frisell T, Öberg S, Kuja-Halkola R, Sjölander A. Sibling comparison designs: bias from non-shared confounders and measurement error. *Epidemiology*. 2012;23(5):713-720.
- **64.** Cnattingius S, Ericson A, Gunnarskog J, Källen K. A quality study of a medical birth registry. *Scand. J. Soc. Med.* 1990;18:143-148.
- **65.** Grove A. Rethinking clinical trials. *Science*. Sep 23 2011;333(6050):1679.

- **66.** Kooperberg C, Ruczinski I, LeBlanc ML, Hsu L. Sequence analysis using logic regression. *Genet Epidemiol.* 2001;21(1):s626-631.
- 67. Janes H, Pepe M, Kooperberg C, Newcomb P. Identifying target populations for screening or not Νν, . ng logic reg... screening using logic regression. Stat Med. . 2005;15(24):1321-1338.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml



Figure 1. Flow diagram showing the individuals excluded from the study population.

BMJ Open: first published as 10.1136/bmjopen-2014-005388 on 30 July 2014. Downloaded from http://bmjopen.bmj.com/ on June 12, 2025 at Department GEZ-LTA Erasmushogeschool .

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

Page 22 of 27

BMJ Open: first published as 10.1136/bmjopen-2014-005388 on 30 July 2014. Downloaded from http://bmjopen.bmj.com/ on June 12, 2025 at Department GEZ-LTA Erasmushogeschool .

rasmushogescl

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

BMJ Open



	AU-ROC	CI-95%			AU-ROC	CI-95%	
Newborn sex	0.501	[0.497	0.504]	Marital status	0.537	[0.533	0.540]
Hypertension	0.506	[0.505	0.507]	Maternal education	0.541	[0.537	0.545]
Maternal origin	0.517	[0.514	0.520]	Birth order	0.589	[0.586	0.593]
Maternal age	0.523	[0.519	0.527]	Smoking	0.593	[58.94	0.597]

Figure 2. Area under the ROC curve (AU-ROC) to compare the discriminatory accuracy of different models to distinguish between SGA and non-SGA babies.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml



Newbornsex	Hypertension	Mate	ernal ge	Maternal origin	Marital status	Maternal education	Birth o	rder	Smoking
		AU-ROC	CI-95%				AU-ROC	CI-95%	
Newborn sex	Pret	0.53	[0.519	0.537]	Marital status	Pret	0.52	[0.511	0.528]
	Term	0.50	[0.498	0.506]		Term	0.54	[0.532	0.540]
	Post	0.52	[0.507	0.531]		Post	0.53	[0.516	0.541]
Hypertension	Pret	0.51	[0.509	0.516]	Maternal education	Pret	0.51	[0.502	0.521]
	Term	0.50	[0.503	0.504]		Term	0.55	[0.541	0.549]
	Post	0.50	[0.499	0.503]		Post	0.55	[0.534	0.563]
Maternal age	Pret	0.52	[0.508	0.526]	Birth order	Pret	0.55	[0.537	0.555]
	Term	0.52	[0.517	0.525]		Term	0.59	[0.587	0.596]
	Post	0.52	[0.507	0.532]		Post	0.59	[0.578	0.602]
Maternal origin	Pret	0.50	[0.497	0.509]	Smoke	Pret	0.53	[0.526	0.543]
	Term	0.52	[0.518	0.524]		Term	0.60	[0.596	0.605]
	Post	0.54	[0.527	0.546]		Post	0.58	[0.568	0.593]

Pret= preterm; Post= Post-term

Figure 3. Area under the ROC curve (AU-ROC) for specific maternal and newborn characteristics.

BMJ Open: first published as 10.1136/bmjopen-2014-005388 on 30 July 2014. Downloaded from http://bmjopen.bmj.com/ on June 12, 2025 at Department GEZ-LTA Erasmushogeschool .

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies



Figure 4. Area under the ROC curve (AU-ROC) for specific maternal and newborn characteristics after stratifying by gestational age (preterm, term, and post-term).

$ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \\ 11 \\ 12 \\ 13 \\ 14 \\ 15 \\ 16 \\ 17 \\ 18 \\ 19 \\ 20 \\ 21 \\ 22 \\ 23 \\ 24 \\ 25 \\ 26 \\ 27 \\ 28 \\ 29 \\ 30 \\ 31 \\ 32 \\ 33 \\ 34 \\ 35 \\ 36 \\ 37 \\ 38 \\ 39 \\ 40 \\ 41 \\ 42 \\ 43 \\ 44 \\ 45 \\ 46 \\ 47 \\ 48 \\ 49 \\ 50 \\ 51 \\ 52 \\ 53 \\ 54 \\ 55 \\ 56 \\ 56 \\ 56 \\ 56 \\ 56 \\ 56 \\ 56 \\ $	
51 52 53 54 55 56 57 58 59 60	

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of cross-sectional studies

Section/Topic	ltem #	Recommendation	Reported on page #
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	1
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	4-5
Methods			
Study design	4	Present key elements of study design early in the paper	4
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(a) Give the eligibility criteria, and the sources and methods of selection of participants	5
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5-6
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	6
Bias	9	Describe any efforts to address potential sources of bias	12
Study size	10	Explain how the study size was arrived at	Figure 1 and page 5
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	6
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	6-7
		(b) Describe any methods used to examine subgroups and interactions	6-7
		(c) Explain how missing data were addressed	6
		(d) If applicable, describe analytical methods taking account of sampling strategy	No applicable
		(e) Describe any sensitivity analyses	11
Results			

BMJ Open: first published as 10.1136/bmjopen-2014-005388 on 30 July 2014. Downloaded from http://bmjopen.bmj.com/ on June 12, 2025 at Department GEZ-LTA Erasmushogeschool Protected by כקצענוקואסנוקאלוחק/לפאנפאנפאנפאנפאנפאנפאנפאנפאנפאנאן/אוואטט, אלעאנאנאנא אופאנאואס, ארסוסנופג.

BMJ Open

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	6
		(b) Give reasons for non-participation at each stage	No applicable
		(c) Consider use of a flow diagram	20
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	7
		(b) Indicate number of participants with missing data for each variable of interest	7
Outcome data	15*	Report numbers of outcome events or summary measures	Table 1
Main results	16	(<i>a</i>) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Table 2
		(b) Report category boundaries when continuous variables were categorized	Table 1
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	No relevant
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	12
Discussion			
Key results	18	Summarise key results with reference to study objectives	8-9
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	12
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	12
Generalisability	21	Discuss the generalisability (external validity) of the study results	12-13
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	1

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

Erssmushoper old by complete in the second state of the second state of the second state of the second s

AT-LAS instruction of the start of the start

BMJ Open

Applying measures of discriminatory accuracy to revisit traditional risk factors for being small for gestational age in Sweden: a national cross-sectional study

Journal:	BMJ Open
Manuscript ID:	bmjopen-2014-005388.R1
Article Type:	Research
Date Submitted by the Author:	13-Jun-2014
Complete List of Authors:	Juárez, Sol; Lund University, Centre for Economic Demography Wagner, Phillip; Lund University, Unit of Social Epidemiology, Faculty of Medicine Merlo, Juan; Lund University, Unit of Social Epidemiology, Faculty of Medicine
Primary Subject Heading :	Reproductive medicine
Secondary Subject Heading:	Reproductive medicine
Keywords:	Fetal medicine < OBSTETRICS, Prenatal diagnosis < OBSTETRICS, Maternal medicine < OBSTETRICS



Sol Pía Juárez

Centre for Economic Demography, Lund University Unit of Social Epidemiology, Department of Clinical Sciences, Lund University, Malmö P.O. Box 7083 SE-220 07 Lund, Sweden Tel. +46 46-222 00 00 Sol.juarez@ekh.lu.se

Phillip Wagner

Unit of Social Epidemiology, Department of Clinical Sciences, Lund University, Malmö Skåne University Hospital (SUS Malmö) Jan Waldenströms gata 35 SE- 214 21 Malmö, Sweden

Juan Merlo

Unit of Social Epidemiology, Department of Clinical Sciences, Lund University, Malmö Skåne University Hospital (SUS Malmö) Jan Waldenströms gata 35 SE- 214 21 Malmö, Sweden

Keywords: risk factors, SGA, discriminatory accuracy

Abstract

Objectives Small for gestational age (SGA) is considered an indicator of intrauterine growth restriction, and multiple maternal and newborn characteristics have been identified as risk factors for SGA. This knowledge is mainly based on measures of average association—(i.e. odds ratio (OR)—that quantify differences in average risk between exposed and unexposed groups. Nevertheless, average associations do not assess the discriminatory accuracy of the risk factors (i.e, its ability to discriminate the babies who will develop SGA from those that will not). Therefore, applying measures of discriminatory accuracy rather than measures of association only, our study revisits known risk factors of SGA and discusses their role from a public health perspective.

Design Cross-sectional study. We measured maternal (ie, smoking, hypertension, age, marital status, education) and delivery (ie, sex, gestational age, birth order) characteristics and performed logistic regression models to estimate both ORs and measures of discriminatory accuracy, like area under the receiver operating characteristic curve (AU-ROC) and the net reclassification improvement.

Setting Data were obtained from the Swedish Medical Birth Registry.

Participants Our sample included 731 989 babies born during 1987–1993.

Results We replicated the expected associations. For instance, smoking (OR=2.57), having had a previous SGA baby (OR=5.48) and hypertension (OR=4.02) were strongly associated with SGA. However, they show a very small discriminatory accuracy (AU-ROC \approx 0.5). The discriminatory accuracy increased, but remained unsatisfactorily low (AU-ROC=0.6), when including all variables studied in the same model.

Conclusions Traditional risk factors for SGA alone or in combination have a low accuracy for discriminating babies with SGA from those without SGA. A proper understanding of these findings is of fundamental relevance to address future research and to design policymaking recommendations in a more informed way.

Strengths and limitations of this study

- Our study emphasises the use and interpretation of measures of discriminatory accuracy (ie, capacity to distinguish between SGA and non-SGA babies) when evaluating risk factors.
- We confirm statistical associations between maternal and newborn characteristics and risk for SGA, but we underline that the discriminatory capacity of all the risk factors studied was very low.
- This low discriminatory capacity suggests that we know very little about the determinants of SGA in the population and that more effort should be paid to understanding individual heterogeneity of effects.
- Our finding is of fundamental relevance to address future research and to design policymaking recommendations in a more informed way.

BMJ Open: first published as 10.1136/bmjopen-2014-005388 on 30 July 2014. Downloaded from http://bmjopen.bmj.com/ on June 12, 2025 at Department GEZ-LTA Erasmushogeschool .

Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies

Small for gestational age (SGA) is commonly identified as a proxy of intrauterine growth restriction (IUGR).¹ This disorder has been associated with neonatal mortality and morbidities² as well as with major medical problems across the life course, such as a higher risk of neurodevelopmental impairments,^{3,4} autism,⁵ schizophrenia,⁶ impaired cognitive function,⁷ coeliac disease in boys,⁸ and reduced bone mass during early infancy,⁹ as well as Barrett's oesophagus and oesophagitis^{10,11} and others.^{12,13} Therefore, the identification of maternal and newborn characteristics (denominated as 'risk factors' in the rest of this work) associated with an increased risk for SGA is of obvious relevance in public health and preventive medicine.

Two reviews, one from 1987¹⁴ and other from 2009,¹⁵ pointed out that SGA is associated with a broad number of genetic, obstetric, demographic, and socioeconomic factors as well as maternal morbidities and toxic exposures before and during pregnancy. However, the identification of these risk factors has been exclusively based on measures of average association (eg, odds ratio) but without considering their accuracy for discriminating babies with, from those without, SGA. Indeed, it is a common practice to use measures of average association to gauge the ability of a factor to discriminate future cases of disease.¹⁶ For example, it is known that maternal hypertension during pregnancy gives a 5.5-fold increased risk of delivering an SGA baby¹⁷. Therefore, this variable is implicitly used as a predictive test to classify who will and who will not deliver an SGA baby. However, in spite of this popular belief, measures of association alone are inappropriate for this discriminatory purpose insofar as there are different scenarios of sensitivity/specificity for a given OR.^{16–23}

Although measures of discriminatory accuracy are extensively applied in other fields of epidemiology like, for instance, the identification of new biomarkers for cardiovascular diseases,¹⁸⁻²¹ these measures are still unusual in public health and epidemiology.²² In fact, as far we know, they have never been explicitly used to formally revisit established maternal and newborn risk factors for SGA.

With this background our study aims to revisit the role of current risk factors for SGA in public health. We do it in two steps. Firstly, using measures of average association, we aim to replicate previous findings and identify maternal and newborn risk factors for SGA. Secondly, we apply measures of discriminatory accuracy to assess the ability of those risk factors (alone or in

BMJ Open: first published as 10.1136/bmjopen-2014-005388 on 30 July 2014. Downloaded from http://bmjopen.bmj.com/ on June 12, 2025 at Department GEZ-LTA Erasmushogeschool . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

BMJ Open

combination) to discriminate babies with, from those without, SGA in the whole population and in different subgroups according to gestational age.

DATA AND METHODS

Study design, setting, and participants

This is a cross-sectional study based on a population-based register. We identified all the 811 599 babies born alive and recorded at the Swedish Medical Birth Registry (MBR) between 1 January 1987 and 31 December 1993. The MBR collects detailed and standardised information on nearly all pregnancies in Sweden culminating in delivery.^{23,24} Using a unique personal identification number, the Swedish authorities (National Board of Health and Welfare and Statistics Sweden) linked the MBR to the Register of the Total Population and the Swedish 1990 population census and created a research database. This database was delivered to us without the personal identification numbers to protect the anonymity of the subjects. The Regional Ethics Review Board in southern Sweden approved the construction of the database.

For the purposes of our study, we selected singletons, because it is known that multiple births (n=19 167) have a different intrauterine growth pattern from gestational weeks 28-30.²⁵ We excluded 13 539 babies born with significant congenital anomalies according to the MBR. Following previously established criteria,²⁶ we also excluded babies with inconsistent information on birthweight according to gestational age (n=9195) and babies weighing less than 500 g (n=51) as well as 15 observations with missing information on maternal age and birth order. The final sample contained 768 059 babies. Thereafter, we stratified the population by gestational age into pre-term (<37 gestational weeks) term (\geq 37 and <42 gestational weeks) and post-term babies (>42 gestational weeks) (Figure 1).

[Figure 1 about here]

Variables

The outcome variable combined birthweight and gestational age to dichotomise as being SGA or not, and using the last category as the reference. This variable was available at the MBR, where

it is routinely calculated following standard intrauterine growth curves.²⁷ Infants were defined as SGA if they weighed less than 2 standard deviations below the expected birthweight for gestational age and gender, according to a Swedish intrauterine growth curve.³¹

In our analyses we included child and maternal characteristics that are known to be associated with low birthweight and SGA.

As child characteristics we used $sex^{14,28}$ and *birth order*^{29,30} classified into three categories (ie, firstborn, second, and third or more). Among maternal characteristics we included *birth interval between newborns*^{14,31}, categorised into <1 year, 1-2 years, >2 years, 'only child' (i.e., when we know that the newborns have previous siblings but we do not have their information in our setting) and first child (i.e., we know the newborn has not a previous sibling); whether the mother has a *previous child with SGA*³² categorised into yes, no, 'only child' and first child; *education*,^{33,34} categorised into low (primary education or less), middle (secondary school), and high education (graduate and PhD); *marital status*,^{35,36} categorised into single, widowed, or divorced, and married or cohabiting; and *maternal age at delivery*,³⁷⁻³⁹ categorised into four groups (ie, <20, 20–24, 25–34, and >35 years old), as well as information on *smoking habits*,⁴⁰⁻⁴³ categorised into non-smoking, light smoking (fewer than 9 cigarettes per day), heavy smoking (more than 9 cigarettes per day), and missing information. Finally, we included information about presence of *hypertension during pregnancy* (yes vs. no),^{15,17} and *maternal origin*, classified as being born in Sweden or not.⁴⁴

Statistical methods

To examine the average association between, on the one hand, the categorical variables mentioned above, and on the other, being SGA, we simply calculated odds ratios and 95% confidence intervals obtained from logistic regression analyses.

The discriminatory accuracy of a risk factor is better appraised by measuring *the true positive fraction* (TPF) and *the false positive fraction* (FPF). For a dichotomous risk factor, the TPF expresses the probability of being exposed to the risk factor when the SGA occurs (ie, cases that are exposed to the risk factor), and the FPF indicates the probability of being exposed to the risk factor. In the ideal scenario the TPF should be 1 and the FPF should be 0, even if a lower TPF or a higher FPF. For instance, if the identification of the risk factors conveys pharmacological treatment, we should try to keep

the FPF as low as possible.

For the evaluation of the discriminatory accuracy of the combination of risk factors within a risk score (ie, predicted probability) we obtained *the receiver operating characteristic* (ROC) *curve*. The ROC curve is constructed by plotting the TPF against the FPF for different risk score thresholds.^{16,45,46} A traditional measure of discriminatory accuracy is the *area under the ROC curve* (AU-ROC) or C statistic.^{16,45,47-49} The AU-ROC extends from 0.5 to 1.0. An AU-ROC=0.5 means that the discriminatory accuracy of the candidate risk factor is similar to that obtained by flipping a coin. That is, a risk factor with an AU-ROC=0.5 is useless. An AU-ROC=1.0 means complete accuracy.

In series of simple logistic regression models we identified the single variables with the highest discriminatory accuracy. Using this information, thereafter, we performed two models. Model A only with the two variables with the higher discriminatory accuracy (i.e., smoking and birth order) and model B which adds the rest of covariates to the initial model A. We ran this second model in order to assess the change of discriminatory accuracy when adding the rest of information to a simpler model. We appraised the incremental value of a model by the difference between AU-ROCs. Due to a problem of collinearity, Stata automatically deleted the two categories in common (i.e., 'only child' and first child) shared by the variables of birth interval and previous child with SGA, keeping them only in the former. All models were stratified by gestational age (i.e., preterm, term and post-term) because it is has been suggested that SGA at term and at preterm may have been driven by a different etiology⁵⁰. We included post-term in order to complete the classification.

We performed the analyses in the whole population, stratifying by gestational age (ie, preterm, term, and post-term). We performed the statistical analyses using STATA 12.0 (College Station, TX: StataCorp LP) and SPSS 21.0 (Armonk, NY: IBM Corp.).

RESULTS

Table 1 shows the maternal and individual characteristics of the population of newborns by SGA status. We see that SGA is much more prevalent among preterm babies (10.14%) than among term (1.87%) and post-term (3.03%) babies. Females show higher prevalence of SGA than males among preterm, and slightly lower prevalence among those born post-term. Regardless of gestational age, firstborns had a higher risk of SGA than their siblings. SGA is more prevalent
Page 8 of 46

BMJ Open

among children who had a previous sibling during the same year, except among those babies born at preterm, but this may be due to the larger amount of missing information about the previous siblings (11%). Mothers who had a previous child with SGA are more likely to have a current SGA baby regardless of gestational age.

SGA was more frequent in mothers younger than 20 years of age; among divorced, widowed, and single women; and among those who were born outside Sweden and those with low educational achievement. In babies born SGA, hypertension was more frequent among preterm than among post-term.

Table 1. Prevalence of SGA in the whole population of babies and in strata of gestational age in Sweden 1987–1993

				SGA st	ratified by	gestatio	nal age	
	SGA		Term		Preterm		Post-ter	m
	768 059		676 961		36 080		55 018	
	(2.3%)		(1.9%)		(10.1%)		(3.0%)	
	Ν	%	Ν	%	Ν	%	Ν	%
Newborn sex								
Male	393 538	2.33	343 549	1.86	19 635	9.20	30 354	3.24
Female	374 521	2.34	333 412	1.87	16 445	11.26	24 664	2.78
Birth order								
First	325 326	3.30	279 426	2.66	18 796	11.67	27 104	4.10
Second	268 998	1.63	242 276	1.31	9 774	8.79	16 948	2.01
<u>></u> Third	173 735	1.63	155 259	1.30	7,510	8.07	10 866	1.97
Birth intervals								
< 1 year	4557	3.55	3841	2.76	415	6.99	301	8.97
1-2 years	77 922	2.00	69 232	1.61	3585	8.23	5105	2.86
>2 years	147 354	1.97	130 979	1.57	5970	9.45	10 405	2.72
'Only child'	341 906	2.77	298 264	2.20	17 789	11.63	25 853	3.29
First child	196 320	1.97	174 645	1.60	8321	8.45	13 354	2.70
Previous child SGA								
Yes	3371	9.58	2913	7.69	259	30.12	196	9.18
No	226 462	1.90	201 139	1.52	9711	8.34	15 615	2.80
'Only child'	341 906	2.77	298 264	2.20	17 789	11.63	25 853	3.29
First child	196 320	1.97	174 645	1.60	8321	8.45	13 354	2.70
Maternal age								
<20	40 735	3.06	35 367	2.49	2485	9.93	2883	4.30
20–24	156 736	2.39	138 232	1.91	7842	9.82	10 662	3.24
25–34	481 689	2.19	425 980	1.76	20 796	9.59	34 913	2.89
>35	88 899	2.69	77 382	2.08	4957	12.04	6 560	2.90

Hypertension	3166	8.69	2648	5.44	383	32.11	135	5.93
Smoking								
Non-smoker	541 962	1.77	479 851	1.37	22 603	9.14	39 508	2.41
Light smoker	109 799	3.62	76 327	3.05	5884	11.86	7588	4.40
Heavy smoker	67 454	4.41	58 790	3.79	4234	11.76	4430	5.64
Missing	48 844	2.90	41 993	2.12	3359	11.79	3492	3.78
Marital status								
Married or								
cohabiting	385 173	2.02	342 330	1.61	16,50	9.36	26 343	2.69
Divorced or								
widowed	27 151	3.03	23 475	2.43	1706	10.67	1970	3.55
Single	355 735	2.63	311 156	2.10	17 874	10.81	26 705	3.33
Maternal origin								
Born in Sweden	663 043	2.25	584 683	1.79	30 745	10.19	47 615	2.79
Not born in								
Sweden	95 393	2.83	84 022	2.30	4749	10.07	6 622	4.27
Missing	9623	2.67	8256	3.02	586	8.02	781	7.30
Maternal education	1							
University	181 083	1.95	160,102	1.52	7367	10.26	13 614	2.57
Secondary	400 363	2.22	353 813	1.77	18 457	9.96	28 093	2.74
Primary	140 186	3.21	122 919	2.31	7631	10.78	9636	3.84
Missing	46 427	2.34	40 127	2.67	2625	9.22	3675	4.82

Table 2 indicates that the risk for being SGA was similar in boys and girls. However, as expected, not being a firstborn reduced the risk of being SGA. With respect to maternal characteristics, mothers younger than 20 years and those 35 years and older had higher risk of delivering an SGA baby compared to 20- to 24-year-old mothers. Mothers who had a previous child during the same year have a higher risk of having a SGA baby as well as those who had a previous child with SGA. Mothers who experienced hypertension during pregnancy had a higher risk of delivering SGA babies. Compared to non-smoker mothers, light and heavy smoker mothers had higher risk of delivering an SGA baby. Divorced and widowed mothers as well as single mothers were more likely to deliver an SGA baby than married and cohabiting mothers. Mothers with primary and secondary education had higher risk of delivering SGA babies than mothers with a university degree. Similarly, mothers who were not born in Sweden were at higher risk of delivering an SGA baby.

BMJ Open: first published as 10.1136/bmjopen-2014-005388 on 30 July 2014. Downloaded from http://bmjopen.bmj.com/ on June 12, 2025 at Department GEZ-LTA Erasmushogeschool .

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

Table 2. Measures of association between offspring and maternal characteristics, and being small for gestational age (SGA), in the whole population of babies and in strata of gestational age in Sweden 1987–1993. Values are odds ratios (OR) and [95% confidence intervals (IC)]. Crude models

		SGA		SGA 1	FERM		SGA F	PRETER	М	SGA F	POST-TE	RM
	Unad	justed r	nodel	Unad	justed r	nodel	Unad	justed ı	nodel	Unad	justed n	nodel
	OR	IC-95%	6	OR	IC-95%	6	OR	IC-95%	6	OR	IC-95%	6
Newborn sex												
Male (ref)												
Female	1.01	[0.98	1.04]	1.01	[0.97	1.04]	0.86	[0.84	0.88]	0.85	[0.77	0.94]
Birth order												
First (ref)												
Second	0.48	[0.47	0.50]	0.49	[0.47	0.51]	0.60	[0.59	0.62]	0.48	[0.42	0.54]
<u>></u> Third	0.49	[0.47	0.51]	0.48	[0.46	0.50]	0.72	[0.70	0.74]	0.47	[0.41	0.54]
Gestational age	0.80	[0.79	0.81]	0.87	[0.85	0.88]	0.86	[0.85	0.87]	1.77	[1.61	1.95]
Birth intervals												
1-2 years (ref)												
<1 year	1.81	[1.53	2.13]	1.74	[1.42	2.12]	0.84	[0.56	1.24]	3.35	[2.18	5.13]
>2 years	0.98	[0.92	1.05]	0.97	[0.90	1.05]	1.16	[1.00	1.35]	0.95	[0.77	1.16]
'Only child'	1.40	[1.32	1.05]	1.37	[1.29	1.46]	1.47	[1.29	1.67]	1.16	[0.97	1.38]
First child	0.99	[0.93	1.05]	0.99	[0.93	1.07]	1.03	[0.89	1.19]	0.94	[0.78	1.15]
Previous child SGA												
No (ref)												
Yes	5.48	[4.87	6.17]	5.45	[4.74	6.28]	4.75	[3.61	6.24]	3.95	[2.49	6.42]
'Only child'	1.47	[1.42	1.53]	1.46	[1.40	1.52]	1.45	[1.33	1.58]	1.18	[1.05	1.33]
First child	1.04	- [0.99	1.09]	1.06	[1.01	1.12]	1.01	- [0.91	1.13]	0.97	[0.84	1.11]
Maternal age		-	-		-	-		-	-		-	-
25–34 (ref)												
<20	1.41	[1.33	1.49]	1.42	[1.33	1.53]	1.44	[1.38	1.50]	1.51	[1.25	1.83]
20–24	1.09	[1.05	1.13]	1.09	[1.04	1.14]	1.16	[1.13	1.19]	1.12	[0.99	1.27]
>35	1.23	- [1.18	1.29]	1.18	[1.12	1.25]	1.31	[1.27	1.36]	1.00	[0.86	1.17]
Hypertension		-	-		-	-		-	_		-	-
No (ref) vs. ves	4.02	[3.55	4.55]	3.05	[2.58	3.61]	2.73	[2.45	3.04]	2.02	[0.99	4.13]
Smoking		-	-		-	-		-	-		-	-
Non-smoker (ref)												
Light smoker	2.08	[2.01	2.16]	2.27	[2.17	2.371	1.30	[1.26	1.33]	1.86	[1.64	2.12]
Heavy smoker	2.56	[2.46	2.67]	2.84	[2.70	2.98]	1.53	[1.48	1.58]	2.42	[2.10	2.791
Missing	1.66	[1.57	1.76]	1.56	[1.45	1.67]	1.70	[1.64	1.76]	1.59	[1.32	1.92]
Marital status		1			L	,		1			1	,
Married or cohabit	ing											
Divorced or widowed	1.52	[1.41	1.63]	1.52	[1.40	1.66]	1.51	[1.43	1.59]	1.33	[1.04	1.71]

BMJ Open

2													
3	Single	1.31	[1.27	1.35]	1.31	[1.27	1.36]	1.19	[1.17	1.22]	1.25	[1.13	1.38]
4 5	Maternal origin												
6	Born in Sweden												
7	(ref)												
8 9	Not born in Sweden	1.27	[1.21	1.32]	1.30	[1.23	1.36]	1.07	[1.04	1.11]	1.56	[1.37	1.77]
10 11	Missing	1.66	[1.49	1.84]	1.71	[1.51	1.94]	1.35	[1.24	1.47]	2.74	[2.08	3.61]
12	Maternal												
13	education												
14	University (ref)												
15	Secondary	1.14	[1.10	1.19]	1.17	[1.12	1.23]	1.13	[1.10	1.17]	1.07	[0.94	1.22]
10	Primary	1.49	[1.42	1.56]	1.53	[1.45	1.62]	1.35	[1.31	1.39]	1.51	[1.30	1.76]
18	Missing	1.67	[1.57	1.77]	1.78	[1.66	1.92]	1.42	[1.36	1.49]	1.92	[1.59	2.31]

Figure 2 shows the values for the AU-ROC of the variables included in Table 2. Overall, their discriminatory accuracy was rather low. Newborn sex had the lower discriminatory accuracy. Having a SGA child and hypertension, despite being the risk factors most strongly associated with SGA (OR 5.48 and 4.02, respectively), had a very low discriminatory accuracy (AU-ROC 0.54 and 0.51). Birth order and smoking were the variables with highest accuracy (AU-ROC 0.59).

[Figure 2 around here]

Figure 3 shows the AU-ROC for SGA of different risk factors after stratification by preterm, term, and post-term. As in the non-stratified analysis, the discriminatory accuracy of the variables was low. Smoking at term showed the highest discrimination (AU-ROC 0.60).

[Figure 3 around here]

Figure 4 shows that the discriminatory accuracy of the general model A, including only birth order and smoking, was slightly improved (just 0.05 proportion units), when all variables were included in the full model B. Among preterm babies model B improved the discriminatory

accuracy of the model by 0.1 proportion units, while this improvement was much lower among SGA term and SGA post-term babies.

[Figure 4 around here]

DISCUSSIONS

We were able to verify a number of recognised maternal and newborn risk factors for SGA. For instance, we found that smoking (OR 2.56), and especially having had a previous SGA baby (OR 5.48) and maternal hypertension (OR 4.02), were 'strongly' associated with being SGA. However, even if the magnitude of the ORs was of a size normally considered as undoubtedly relevant in epidemiology, none of those traditional risk factors for SGA provided enough accuracy to discriminate babies with SGA from other babies. In fact, the AU-ROC for having had a previous SGA child and maternal hypertension was slightly higher than 0.5, which means that the accuracy of this variable for discriminating babies with SGA from those without SGA was rather similar to that obtained by flipping a coin. That is, we need to recognise that, although on average, mothers with hypertension were four times more likely to have an SGA baby, many mothers with hypertension delivered babies without SGA, and many SGA babies were born to mothers without hypertension. Our findings, therefore, seriously question the utility of maternal hypertension during pregnancy for planning strategies of prevention against SGA. This statement, however, does not mean that hypertension during pregnancy is irrelevant to understanding the origin of SGA, but rather that we need to determine who among hypertensive mothers is actually prone to deliver an SGA baby.

There is a tacit but fallacious belief that the discriminatory accuracy of a risk factor is high when it is supported by a 'strong' association (e.g, an OR of 4, as in the case of maternal hypertension). However, for an association to be an accurate instrument for discrimination, it must be of a magnitude rarely identified in epidemiologic studies.^{16,51-53} Following our example, a low discriminatory accuracy only indicates that any attempt of intervention based on the existence of the risk factor will be inefficient and even inappropriate, because health professionals will unnecessarily treat many mothers. The decision to start an intervention should seriously take into account the existence of important (physical or emotional) side effects in the false-positive women. That is, it is always important to consider the principle of *primum non nocere*.²²

Compared to the other variables studied, birth order and smoking presented a higher discriminatory capacity. However, their discriminatory accuracy was still very low in absolute terms (AU-ROC \approx 0.59). Also, combining all the variables in the same model did not substantially increase the discriminatory accuracy (AU-ROC=0.69). In other words, our results indicate that we actually do not know so much about what determines being SGA.

The existence of a low discriminatory accuracy suggests that around the population average risk there is considerable individual heterogeneity. Therefore, a logical consequence should be to identify which women are most susceptible to the risk factors. Hence, we explored the discriminatory accuracy of the chosen risk factors in different strata of gestational age at birth. We found that the combination of all variables in the same model had a minor improvement for discriminatory accuracy among those born at term or preterm as well as post-term.

Our finding suggests the existence of individual heterogeneity of responses to some specific variables, so the discriminatory accuracy depended, for instance, on whether the baby was preterm, term, or post-term. In fact, smoking, birth order, maternal origin, and marital status had a lower discriminatory capacity among preterm than among babies at term. On the contrary, newborn sex and hypertension had a higher discriminatory accuracy among preterm than among term babies. In the same way, newborn sex and maternal origin had a lower discriminatory capacity among post-term babies, while for smoking and having had a previous SGA baby we found the opposite relation. The variation of the magnitude of the discriminatory accuracy by gestational age at birth expresses the existence of individual heterogeneity.

In addition, the definition of SGA may also actively contribute to reducing the discriminatory accuracy of the traditional risk factors since discrimination depends both on the outcome and as on the exposure. Thus, low discrimination can result from the fact that SGA fails to distinguish between pathological and constitutionally small babies, that is, to properly capture the health dimension that it is supposed to be a proxy for (IUGR)¹. In order to address this shortcoming, we stratified SGA by gestational age as this has been identified as a good strategy to distinguish between these two ⁵⁰. However, we do not find support for this approach since we found a lower discriminatory accuracy among preterm (presumably pathological small) than among term SGA babies. In this regard, our findings show awareness of the caveats pointed out by previous studies

on the use of SGA as a proxy for IUGR¹, and encourage further research aiming to better capture IUGR.

Our findings have important research and policymaking implications. A possible reason for the low discriminatory accuracy of many average associations is that average effects are a mixture of individual level effects and therefore mix inter individual heterogeneity (ie, some individuals respond intensively to the exposure, while others are resilient or might even respond in the opposite direction). The approach based on discriminatory accuracy understands average effects as an idealised mean value that does not necessarily represent the heterogeneity of individual effects.²² Some scholars prefer to conceive individual outcome as the expression of a stochastic phenomenon that is best estimated by the average risk using a probabilistic approach.⁵⁴ Our understanding instead is that individual outcome reflects the interindividual heterogeneity of responses that can be potentially determined; lack of knowledge could be amended by a better understanding of individual responses.⁵⁵ See elsewhere for a longer explanation of these ideas.^{22,56,57} From this perspective, reducing exposure to a risk factor would only be effective when acting on the susceptible, but not on the resilient, individuals. For instance, we need to better capture babies who suffer from IUGR, since, so far, we are incapable of distinguishing between babies who are constitutionally small from those who are pathologically growth restricted.¹ By stratifying between preterm, term, and post-term, we might be able to better approach the underlying heterogeneity.

From the policymaking perspective, our findings suggest that hitherto there has not been enough knowledge to identify any specific risk factor or combination of them that could discriminate with accuracy children with and without SGA status. Our findings support policymaking oriented to lifestyle modification, as according to the principle of *primum non nocere*,²² they have mostly positive consequences, even for 'false-positive' mothers. For instance, persuading women to quit smoking reduces the risk of SGA in some babies, but it improves general wellbeing in everyone. However, other risk factors with low discriminatory accuracy that lead to pharmacological treatment or screening might result in unnecessary side effects and cost. In the long run, an uncritical use of variables with low discriminatory accuracy may hinder the identification of pertinent risk factors and susceptible individuals and damage the scientific credibility of modern epidemiology.^{22,56,57}

Our conclusions are based on classical measures of discriminatory accuracy such as the AU-ROC curve. These measures have been criticised as insensitive to small changes in predicted individual risk.⁵⁸ Some authors propose more specific measures of reclassification, like the net reclassification improvement (NRI), and the integrated discrimination improvement (IDI).⁵⁹⁻⁶² We applied NRI and IDI in a sensitivity analysis (results not shown in tables). For example, using NRI, we observed a reclassification of 30%. However, this figure does not add substantial information to our results, since NRI (as well as IDI) refers to the misclassification occurring all along the risk scale, instead of capturing the misclassification which takes place around the fixed threshold. Furthermore, the new NRI and IDI measures have also been criticised,⁶³ and some authors⁶⁴ have explicitly advised against their use in common epidemiological practice because, unlike IDI and NRI, traditional measures of discrimination like the AU-ROC curve have the advantage that prognostic performance cannot be manipulated.⁶⁴ Therefore, we preferred to quantify discriminatory accuracy by analysing ROC curves and AU-ROCs.

Our analyses are based on a national medical registry covering almost the entire population of residents in Sweden. Nearly all births are registered in the MBR, because giving birth at home is very unusual in Sweden. In addition, estimation of SGA is routinely calculated at the MBR following standard intrauterine growth curves.²⁷ However, our study also has a number of limitations. Because of lack of data, we could not assess many other variables identified in the literature as 'risk factors', such as genetic or nutritional factors.^{15,65} In spite of the quality of the MBR, the information regarding smoking is based on a self-reported questionnaire (anamnesis) administered by the midwife at the first antenatal visit (ie, between 10 and 12 gestational weeks), which to some extent might bias the result by including misclassification of exposure.⁶⁶ However, a study conducted in Sweden comparing self-reported nicotine exposure and plasma levels of cotinine in early and late pregnancy concluded that self-reported smoking information had acceptable validity.⁶⁷

Unfortunately, we could not identify those mothers who suffered from preeclampsia, for which the discriminatory accuracy concerning SGA may be higher than for hypertension in our model. Further analysis on this aspect is required. Another limitation of our study is that we calculated the discriminatory accuracy in the same sample used for constructing the predicted model. This procedure, however, might overestimate the discriminatory accuracy of the models, so the low discriminatory accuracy found may be an underestimation.

BMJ Open: first published as 10.1136/bmjopen-2014-005388 on 30 July 2014. Downloaded from http://bmjopen.bmj.com/ on June 12, 2025 at Department GEZ-LTA Erasmushogeschool . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

Since our study has been carried out with data from 1987-1993, we performed a sensitivity analysis to check possible differences in current years (2000-2010) but the results remain the same. Given the consistence of the results, we preferred to maintain the results for years 1987-1993 in order to use a more accurate outcome since the definition of SGA is based on standard curves estimated for Sweden with data from 1985-1989²⁸. Moreover, the period we cover is of relevance to our study since most of the risk factors which are discussed in our paper were mainly identified in that period and, in the case of Sweden, with the data we use.

In conclusion, applying measures of discriminatory accuracy rather than measures of association only, our study revisits known risk factors of SGA and discusses their role from a public health perspective. We found that neither models including simple variables nor models including several variables at the same time have a good discriminatory accuracy to discriminate babies with SGA from those without SGA. This finding is of fundamental relevance to address future research and to design policymaking recommendations in a more informed way.

As noted elsewhere,^{56,57} there is need of a new epidemiological approach that systematically provides information on the discriminatory accuracy and interindividual heterogeneity of effects and does not rely only on average measures of association.⁶⁸ In this line, new statistical methods like logic regression seem promising.^{69,70} A fundamental change is needed in the way traditional risk factors are currently interpreted in public health epidemiology. If the *discriminatory accuracy* of most classical risk factors is very low, what happens with the vast majority of recommendations given so far in epidemiology and public health? Are health professionals misleading the community by raising the alarm about risks that may be harmless for most individuals? What are the ethical repercussions of using risk factors with low discriminatory accuracy? Are there problems of inefficiency, medicalisation, and stigmatisation? We believe that these questions have a high significance for both the community and the future of public health research.

Acknowledgments

We would like to thank Dr. Karin Kallén for her very helping comments on a previous version of this paper. This work was supported by the Swedish Research Council (VR) (Dnr #2013-2484, PI Juan Merlo), the Centre for Economic Demography and the SIMSAM early life Lund (Dnr #2013-5474). The funders had no role in study design, data collection, and analysis; decision to publish; or preparation of the manuscript.

Authors' contributions

Merlo had the original idea of applying measures of discriminatory accuracy for the interpretation of risk factors and discussed it with Wagner and Juarez. Merlo and Juárez initiated the study. Merlo, Juárez, and Wagner contributed to the design of the study; Juárez performed the analyses under the supervision of Merlo and Wagner. Juárez wrote the first draft of the manuscript, and Merlo contributed to the writing of the final version. All authors made substantial contributions to the interpretation of the results and manuscript revision. All authors approved the final version of the manuscript.

Competing interest statement

The authors have declared that no competing interests exist.

Data sharing

No additional data available

Figure legends

Figure 1. Flow diagram showing the individuals excluded from the study population.

Figure 2. Area under the ROC curve (AU-ROC) to compare the discriminatory accuracy of different models to distinguish between SGA and non-SGA babies.

Figure 3. Area under the ROC curve (AU-ROC) for specific maternal and newborn characteristics.

Figure 4. Area under the ROC curve (AU-ROC) for specific maternal and newborn characteristics after stratifying by gestational age (preterm, term, and post-term).

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

REFERENCES

- **1.** Urquia ML, Ray JG. Seven caveats on the use of low birthweight and related indicators in health research. *J Epidemiol Community Health*. 2012;66(11):971-975.
- Grisaru-Granovsky S, Reichman B, Lerner-Geva L, et al. Mortality and morbidity in preterm small-for-gestational-age infants: a population-based study. *Am J Obstet Gynecol.* . 2012;206(2):150e151-e156.
- **3.** Streimish IG, Ehrenkranz RA, Allred EN, et al. Birth weight- and fetal weight-growth restriction: impact on neurodevelopment. *Early Hum Dev.* 2012;88(9):765-771.
- **4.** Eikenes L, Martinussen MP, Lund LK, et al. Being born small for gestational age reduces white matter integrity in adulthood: a prospective cohort study. *Pediatr Res.* . 2012;72(6):649-654.
- 5. Moore GS, Kneitel AW, Walker CK, et al. Autism risk in small- and large-for-gestational-age infants. *Am J Obstet Gynecol.* 2012;206(4):314.
- 6. Nielsen PR, Mortensen PB, Dalman C, et al. Fetal growth and schizophrenia: A nested casecontrol and case-sibling study. *Schizophr Bull.* 2012;12(): .
- **7.** Løhaugen GC, Ostgård HF, Andreassen S, et al. Small for gestational age and intrauterine growth restriction decreases cognitive function in young adults. *J. Pediatr.* 2013; (): .
- **8.** Wingren CJ, Agardh D, Merlo J. Revisiting the risk of celiac disease in children born small for gestational age: a sibling design perspective. *Scand J Gastroenterol.* 2012;47(6):632-639.
- **9.** van de Lagemaat M, Rotteveel J, van Weissenbruch MM, et al. Small-for-gestational-age preterm-born infants already have lower bone mass during early infancy. *Bone.* 2012;51(3):441-446.
- **10.** Forssell L, Cnattingius S, Bottai M, et al. Increased risk of Barrett's esophagus among Individuals born preterm or small for gestational age. *Clin Gastroenterol Hepatol.* . 2013; (): .
- **11.** Forssell L, Cnattingius S, Bottai M, et al. Risk of esophagitis among individuals born preterm or small for gestational age. *Clin Gastroenterol Hepatol.* . 2012;10(12):1369-1375.
- **12.** Dalla Pozza RD, Bechtold S, Putzker S, et al. Young adults born small for gestational age: is reduced baroreceptor sensitivity a risk factor for hypertension? *Clin Cardiol.* . 2006;29(5):215-218.
- **13.** Crispi F, Figueras F, Cruz-Lemini M, et al. Cardiovascular programming in children born small for gestational age and relationship with prenatal signs of severity. *Am J Obstet Gynecol.* 2012;207(2):121.e121–121.e129.
- **14.** Kramer MS. Determinants of low birth weight: methodological assessment and meta-analysis. *Bull. World Health Organ.* 1987;65(5):663-737.
- **15.** McCowan L, Horgan RP. Risk factors for small for gestational age infants. *Best Pract Res Clin Obstet Gynaecol.* . 2009;23(6):779-793.
- **16.** Pepe MS, Janes H, Longton G, et al. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *Am J Epidemiol.* May 1 2004;159(9):882-890.
- **17.** Catov JM, Nohr EA, Olsen J, et al. Chronic hypertension related to risk for preterm and term small for gestational age births. *Obstet Gynecol.* 2008;112(2):290-296.
- Zethelius B, Berglund L, Sundstrom J, et al. Use of multiple biomarkers to improve the prediction of death from cardiovascular causes. *New England Journal of Medicine*. May 15 2008;358(20):2107-2116.
- **19.** Cooney MT, Dudina AL, Graham IM. Value and limitations of existing scores for the assessment of cardiovascular risk: a review for clinicians. *J Am Coll Cardiol.* Sep 29 2009;54(14):1209-1227.

2		
3	20.	Law MR Wald NL Morris IK. The performance of blood pressure and other cardiovascular risk
4	20.	factors as screening tests for ischaemic heart disease and stroke. J Med Screen. 2004:11(1):3-7.
5	21.	Melander O, Newton-Cheh C, Almgren P, et al. Novel and conventional biomarkers for
0 7		prediction of incident cardiovascular events in the community. JAMA : the journal of the
8		American Medical Association. Jul 1 2009;302(1):49-57.
9	22.	Merlo J. Multilevel analysis of individual heterogeneity: a fundamental critique of the current
10		probabilistic risk factor epidemiology (invited commentary). American Journal of Epidemiology.
11		2014;In press.
12	23.	Facts about mothers, childbirth and newborn children. Medical Birth Registry, 1973 to 2000.
14		Stockholm: Centre of epidemiology;2002.
15	24.	Cnattingius S, Ericson A, Gunnarskog J, Kallen B. A quality study of a medical birth registry. Scand
16		<i>J Soc Med</i> . Jun 1990;18(2):143-148.
17	25.	Cunningham FG. Williams Obstetrics. Third ed: McGraw-Hill; 2005.
18 10	26.	Källén B. A birth weight for gestational age standard based on data in the Swedish Medical Birth
20		Registry, 1985-1989. European Journal of Epidemiology. 1995;11(5):601-606.
21	27.	Marsál K, Persson PH, Larsen T, et al. Intrauterine growth curves based on ultrasonically
22		estimated foetal weights. Acta Paediatr. 1996;85(7):843-848.
23	28.	Alexander GR, Kogan MD, Himes JH. 1994-1996 U.S. singleton birth weight percentiles for
24		gestational age by race, Hispanic origin, and gender. <i>Matern. Child Health J.</i> 1999;3(4):225-232.
25 26	29.	Swamy GK, Edwards MJ, Gelfand A, et al. Maternal age, birth order, and race: differential effects
27		on birthweight. J. Epidemiol. Community Health. 2012;66(2):136-142.
28	30.	Elshibly EM, Schmalisch G. The effect of maternal anthropometric characteristics and social
29		factors on gestational age and birth weight in Sudanese newborn infants. BMC Public Health.
30		2008;8(244): .
31	31.	Grisaru Granovsky S, Gordon ES, Haklai Z, et al. Effect of interpregnancy interval on adverse
32 33		perinatal outcomesa national study. <i>Contraception</i> . 2009;80(6):512-518.
34	32.	Voskamp B, Kazemier BM, Ravelli AC, et al. Recurrence of small-for-gestational-age pregnancy:
35		analysis of first and subsequent singleton pregnancies in The Netherlands. Am J Obstet Gynecol.
36		2013;208(5):3/4.e3/1-3/6.
37	33.	Kramer MS, Seguin L, Lydos J, et al. Socio-economic disparities in pregnancy outcome: why do
38	24	the poor fare so poorly? Paealatric and Perinatal Epidemiology. 2000;14:194-210.
39 40	34.	Auger N, Park AL, Harper S, et al. Educational inequalities in preterm and term small-tor-
41	25	gestational-age birth over time. Ann Epidemiol 2012;22(3):100-107.
42	35.	shall PS, Zao J, Ali S, Dirths. KSGODOPL. Maternal marital status and Dirth outcomes: a systematic
43	26	Castro T. Single motherhood and low birthweight in Spain: Narrowing social inequalities in
44	50.	health? Demographic Research 2010:22(27):863-890
45 46	37	Fraser AM Brockert IF Ward RH Association of young maternal age with adverse repreductive
47	57.	outcomes The New England Journal of Medicine 1995:332(17):1113-1118
48	38	Odibo A Nelson D Stamilo DM et al. Advanced maternal age is an independent risk factor for
49	50.	intrautering growth restriction American Journal of Perinatology 2006;23(5):325-328
50	39.	Campbell MK Cartier S Xie B et al. Determinants of small for gestational age birth at term
51 52	••••	Paediatr Perinat Enidemiol. 2012:26(6):525-533
52 53	40.	Rogers JM. Tobacco and pregnancy: overview of exposures and effects. <i>Birth defects research</i>
54		2008:Part C(84):1-15.
55	41.	Kabir Z, Daly S, Clarke V, et al. Smoking Ban and Small-For-Gestational Age Births in Ireland. PLoS
56		ONE. 2013;8(3):e57441.
57 59		
50 59		
60		19
		For near review only them is non-him com/site/shout/audalines white

- inced maternal age is an independent risk factor for lournal of Perinatology. 2006;23(5):325-328.
- minants of small for gestational age birth at term. 5-533.
- iew of exposures and effects. Birth defects research
- n and Small-For-Gestational Age Births in Ireland. PLoS

- 42. Mitchell EA, Thompson JM, Robinson E, et al. Smoking, nicotine and tar and risk of small for gestational age babies. Acta Paediatr. 2002;91(3):323-328. 43. Juárez SP, Merlo J. Revisiting the Effect of Maternal smoking during pregnancy on offspring birthweight: a quasi-experimental sibling analysis in Sweden. PLoS ONE. 2013;8(4):e61734. 44. Li X, Sundquist K, Sundquist J. Risks of small-for-gestational-age births in immigrants: A nationwide epidemiological study in Sweden. Scand. J. Public Health. 2012;40:634-640. Pepe MS. The statistical evaluation of medical tests for classification and prediction. Oxford ; 45. New York: Oxford University Press; 2003. 46. Zweig MH, Broste SK, Reinhart RA. ROC curve analysis: an example showing the relationships among serum lipid and apolipoprotein concentrations in identifying patients with coronary artery disease. Clin Chem. Aug 1992;38(8 Pt 1):1425-1428. 47. Pepe MS, Janes H, Gu JW. Letter by Pepe et al regarding article, "Use and misuse of the receiver reply e134. 48. Statistics in Medicine. Oct 30 2010;29(24):2508-2520. 49. Biometrische Zeitschrift. Aug 2008;50(4):457-479. 50. age births in population-based studies. Early Hum Dev. 2009;85(10):653-658. 51.
 - operating characteristic curve in risk prediction". Circulation. Aug 7 2007;116(6):e132; author
 - Royston P, Altman DG. Visualizing and assessing discrimination in the logistic regression model.
 - Gerds TA, Cai T, Schumacher M. The performance of risk prediction models. Biometrical journal.
 - Ananth CV, Vintzileos AM. Distinguishing pathological from constitutional small for gestational
 - Boyko EJ, Alderman BW. The use of risk factors in medical diagnosis: opportunities and cautions. Journal of clinical epidemiology. 1990;43(9):851-858.
 - 52. Wald NJ, Hackshaw AK, Frost CD. When can a risk factor be used as a worthwhile screening test? BMJ. Dec 11 1999;319(7224):1562-1565.
 - 53. Khoury MJ, Newill CA, Chase GA. Epidemiologic evaluation of screening for risk factors: application to genetic screening. American journal of public health. Oct 1985;75(10):1204-1208.
 - 54. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation. Feb 20 2007;115(7):928-935.
 - Zernicka-Goetz M, Huang S. Stochasticity versus determinism in development: a false 55. dichotomy? Nature reviews. Genetics. Nov 2010;11(11):743-744.
 - 56. Merlo J, Wagner P. The tyranny of the averages and the indiscriminate use of risk factors in public health: a call for revolution. European journal of epidemiology. 2013;28(1, Supplement):148.
 - 57. Merlo J, Wagner P, Juarez S, et al. The tyranny of the averages and the indiscriminate use of risk factors and population attributable fractions in Public Health: the case of coronary heart disease. Working paper version 2013-09-26. Unit for Social Epidemiology, Department of Clinical Sciences, Faculty of Medicine, Lund University. 2013;http://www.med.lu.se/english/klinvetmalmo/unit for social epidemiology/abstracts c working papers c.
 - 58. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation. Feb 20 2007;115(7):928-935.
 - 59. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. Statistics in Medicine. Jan 30 2008;27(2):157-172.
 - 60. Pencina MJ, D'Agostino RB, Pencina KM, et al. Interpreting incremental value of markers added to risk prediction models. Am J Epidemiol. Sep 15 2012;176(6):473-481.
 - 61. Pencina MJ, D'Agostino RB, Sr., Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. Stat Med. Jan 15 2011;30(1):11-21.

1

2 3

4

5

6 7

8

9

10

11

12

13

14 15

16

17

18

19

20

21

22 23

24

25

26

27

28

29

30 31

32

33

34

35

36

37

38 39

40

41

42

43

44

45

46 47

48

49

50

51

52

53 54

55

56

57

58 59

60

BMJ Open

- **62.** Hilden J. On NRI, IDI, and "good-looking" statistics with nothing underneath. *Epidemiology*. 2014;25(2):265-267.
 - **63.** Pepe MS. Problems with risk reclassification methods for evaluating prediction models. *Am J Epidemiol.* Jun 1 2011;173(11):1327-1335.
 - **64.** Hidden J, Gerds T. Evaluating the impact of novel biomarkers: Do not rely on IDI and NRI. Department of Biostatistics, University of Copenhagen, Denmark. 2012;Research Report 12/08.
 - **65.** Kramer MS. Determinants of low birth weight: methodological assessment and meta-analysis. *Bulletin of World Health Organization.* 1984;65(5):663-737.
 - **66.** Frisell T, Öberg S, Kuja-Halkola R, et al.Sibling comparison designs: bias from non-shared confounders and measurement error. *Epidemiology*. 2012;23(5):713-720.
 - **67.** Cnattingius S, Ericson A, Gunnarskog J, et al. A quality study of a medical birth registry. *Scand. J. Soc. Med.* 1990;18:143-148.
- **68.** Grove A. Rethinking clinical trials. *Science*. Sep 23 2011;333(6050):1679.
- **69.** Kooperberg C, Ruczinski I, LeBlanc ML, et al. Sequence analysis using logic regression. *Genet Epidemiol.* 2001;21(1):s626-631.
- **70.** Janes H, Pepe M, Kooperberg C, et al. Identifying target populations for screening or not screening using logic regression. *Stat Med.* . 2005;15(24):1321-1338.

BMJ Open: first published as 10.1136/bmjopen-2014-005388 on 30 July 2014. Downloaded from http://bmjopen.bmj.com/ on June 12, 2025 at Department GEZ-LTA Erasmushogeschool .

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

Applying measures of discriminatory accuracy to revisit traditional risk factors for being small for gestational age in Sweden: a national crosssectional study

Sol Pía Juárez

Centre for Economic Demography, Lund University Unit of Social Epidemiology, Department of Clinical Sciences, Lund University, Malmö P.O. Box 7083 SE-220 07 Lund, Sweden Tel. +46 46-222 00 00 Sol.juarez@ekh.lu.se

Phillip Wagner

Unit of Social Epidemiology, Department of Clinical Sciences, Lund University, Malmö Skåne University Hospital (SUS Malmö) Jan Waldenströms gata 35 SE-214 21 Malmö, Sweden

Juan Merlo

Unit of Social Epidemiology, Department of Clinical Sciences, Lund University, Malmö Skåne University Hospital (SUS Malmö) Jan Waldenströms gata 35 SE- 214 21 Malmö, Sweden

Keywords: risk factors, SGA, discriminatory accuracy

Abstract

Objectives Small for gestational age (SGA) is considered an indicator of intrauterine growth restriction, and multiple maternal and newborn characteristics have been identified as risk factors for SGA. This knowledge is mainly based on measures of average association—(i.e. odds ratio (OR)—that quantify differences in average risk between exposed and unexposed groups. Nevertheless, average associations do not assess the discriminatory accuracy of the risk factors (i.e., its ability to discriminate the babies who will develop SGA from those that will not). Therefore, applying measures of discriminatory accuracy rather than measures of association only, our study revisits known risk factors of SGA and discusses their role from a public health perspective.

Design Cross-sectional study. We measured maternal (ie, smoking, hypertension, age, marital status, education) and delivery (ie, sex, gestational age, birth order) characteristics and performed logistic regression models to estimate both ORs and measures of discriminatory accuracy, like area under the receiver operating characteristic curve (AU-ROC) and the net reclassification improvement.

Setting Data were obtained from the Swedish Medical Birth Registry.

Participants Our sample included 731 989 babies born during 1987–1993.

Results We replicated the expected associations. For instance, smoking (OR=2.57), having had a previous SGA baby (OR=5.48) and hypertension (OR=4.02) were strongly associated with SGA. However, they show a very small discriminatory accuracy (AU-ROC \approx 0.5). The discriminatory accuracy increased, but remained unsatisfactorily low (AU-ROC=0.6), when including all variables studied in the same model.

Conclusions Traditional risk factors for SGA alone or in combination have a low accuracy for discriminating babies with SGA from those without SGA. A proper understanding of these findings is of fundamental relevance to address future research and to design policymaking recommendations in a more informed way.

BMJ Open: first published as 10.1136/bmjopen-2014-005388 on 30 July 2014. Downloaded from http://bmjopen.bmj.com/ on June 12, 2025 at Department GEZ-LTA Erasmushogeschool .

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

BMJ Open: first published as 10.1136/bmjopen-2014-005388 on 30 July 2014. Downloaded from http://bmjopen.bmj.com/ on June 12, 2025 at Department GEZ-LTA Erasmushogeschool .

Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies

Strengths and limitations of this study

- Our study emphasises the use and interpretation of measures of discriminatory accuracy (ie, capacity to distinguish between SGA and non-SGA babies) when evaluating risk factors.
- We confirm statistical associations between maternal and newborn characteristics and risk for SGA, but we underline that the discriminatory capacity of all the risk factors studied was very low.
- This low discriminatory capacity suggests that we know very little about the determinants of SGA in the population and that more effort should be paid to understanding individual heterogeneity of effects.
- Our finding is of fundamental relevance to address future research and to design policymaking recommendations in a more informed way.

INTRODUCTION

Small for gestational age (SGA) is commonly identified as a proxy of intrauterine growth restriction (IUGR).¹ This disorder has been associated with neonatal mortality and morbidities² as well as with major medical problems across the life course, such as a higher risk of neurodevelopmental impairments,^{3,4} autism,⁵ schizophrenia,⁶ impaired cognitive function,⁷ coeliac disease in boys,⁸ and reduced bone mass during early infancy,⁹ as well as Barrett's oesophagus and oesophagitis^{10,11} and others.^{12,13} Therefore, the identification of maternal and newborn characteristics (denominated as 'risk factors' in the rest of this work) associated with an increased risk for SGA is of obvious relevance in public health and preventive medicine.

Two reviews, one from 1987¹⁴ and other from 2009,¹⁵ pointed out that SGA is associated with a broad number of genetic, obstetric, demographic, and socioeconomic factors as well as maternal morbidities and toxic exposures before and during pregnancy. However, the identification of these risk factors has been exclusively based on measures of average association (eg, odds ratio) but without considering their accuracy for discriminating babies with, from those without, SGA. Indeed, it is a common practice to use measures of average association to gauge the ability of a factor to discriminate future cases of disease.¹⁶ For example, it is known that maternal hypertension during pregnancy gives a 5.5-fold increased risk of delivering an SGA baby¹⁷. Therefore, this variable is implicitly used as a predictive test to classify who will and who will not deliver an SGA baby. However, in spite of this popular belief, measures of association alone are inappropriate for this discriminatory purpose insofar as there are different scenarios of sensitivity/specificity for a given OR.^{16–23}

Although measures of discriminatory accuracy are extensively applied in other fields of epidemiology like, for instance, the identification of new biomarkers for cardiovascular diseases,¹⁸⁻²¹ these measures are still unusual in public health and epidemiology.²² In fact, as far we know, they have never been explicitly used to formally revisit established maternal and newborn risk factors for SGA.

With this background our study aims to revisit the role of current risk factors for SGA in public health. We do it in two steps. Firstly, using measures of average association, we aim to replicate previous findings and identify maternal and newborn risk factors for SGA. Secondly, we apply measures of discriminatory accuracy to assess the ability of those risk factors (alone or in

combination) to discriminate babies with, from those without, SGA in the whole population and in different subgroups according to gestational age.

DATA AND METHODS

Study design, setting, and participants

This is a cross-sectional study based on a population-based register. We identified all the 811 599 babies born alive and recorded at the Swedish Medical Birth Registry (MBR) between 1 January 1987 and 31 December 1993. The MBR collects detailed and standardised information on nearly all pregnancies in Sweden culminating in delivery.^{23,24} Using a unique personal identification number, the Swedish authorities (National Board of Health and Welfare and Statistics Sweden) linked the MBR to the Register of the Total Population and the Swedish 1990 population census and created a research database. This database was delivered to us without the personal identification numbers to protect the anonymity of the subjects. The Regional Ethics Review Board in southern Sweden approved the construction of the database.

For the purposes of our study, we selected singletons, because it is known that multiple births (n=19 167) have a different intrauterine growth pattern from gestational weeks 28-30.²⁵ We excluded 13 539 babies born with significant congenital anomalies according to the MBR. Following previously established criteria,²⁶ we also excluded babies with inconsistent information on birthweight according to gestational age (n=9195) and babies weighing less than 500 g (n=51) as well as 15 observations with missing information on maternal age and birth order. The final sample contained 768 059 babies. Thereafter, we stratified the population by gestational age into pre-term (<37 gestational weeks) term (\geq 37 and <42 gestational weeks) and post-term babies (>42 gestational weeks) (Figure 1).

[Figure 1 about here]

Variables

The outcome variable combined birthweight and gestational age to dichotomise as being SGA or not, and using the last category as the reference. This variable was available at the MBR, where

it is routinely calculated following standard intrauterine growth curves.²⁷ Infants were defined as SGA if they weighed less than 2 standard deviations below the expected birthweight for gestational age and gender, according to a Swedish intrauterine growth curve.³¹

In our analyses we included child and maternal characteristics that are known to be associated with low birthweight and SGA.

As child characteristics we used $sex^{14,28}$ and *birth order*^{29,30} classified into three categories (ie, firstborn, second, and third or more). Among maternal characteristics we included *birth interval between newborns*^{14,31}, categorised into <1 year, 1-2 years, >2 years, 'only child' (i.e., when we know that the newborns have previous siblings but we do not have their information in our setting) and first child (i.e., we know the newborn has not a previous sibling); whether the mother has a *previous child with SGA*³² categorised into yes, no, 'only child' and first child; *education*,^{33,34} categorised into low (primary education or less), middle (secondary school), and high education (graduate and PhD); *marital status*,^{35,36} categorised into single, widowed, or divorced, and married or cohabiting; and *maternal age at delivery*,³⁷⁻³⁹ categorised into four groups (ie, <20, 20–24, 25–34, and >35 years old), as well as information on *smoking habits*,⁴⁰⁻⁴³ categorised into non-smoking, light smoking (fewer than 9 cigarettes per day), heavy smoking (more than 9 cigarettes per day), and missing information. Finally, we included information about presence of *hypertension during pregnancy* (yes vs. no),^{15,17} and *maternal origin*, classified as being born in Sweden or not.⁴⁴

Statistical methods

To examine the average association between, on the one hand, the categorical variables mentioned above, and on the other, being SGA, we simply calculated odds ratios and 95% confidence intervals obtained from logistic regression analyses.

The discriminatory accuracy of a risk factor is better appraised by measuring *the true positive fraction* (TPF) and *the false positive fraction* (FPF). For a dichotomous risk factor, the TPF expresses the probability of being exposed to the risk factor when the SGA occurs (ie, cases that are exposed to the risk factor), and the FPF indicates the probability of being exposed to the risk factor. In the ideal scenario the TPF should be 1 and the FPF should be 0, even if a lower TPF or a higher FPF. For instance, if the identification of the risk factors conveys pharmacological treatment, we should try to keep

BMJ Open: first published as 10.1136/bmjopen-2014-005388 on 30 July 2014. Downloaded from http://bmjopen.bmj.com/ on June 12, 2025 at Department GEZ-LTA Erasmushogeschool . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

the FPF as low as possible.

For the evaluation of the discriminatory accuracy of the combination of risk factors within a risk score (ie, predicted probability) we obtained *the receiver operating characteristic* (ROC) *curve*. The ROC curve is constructed by plotting the TPF against the FPF for different risk score thresholds.^{16,45,46} A traditional measure of discriminatory accuracy is the *area under the ROC curve* (AU-ROC) or C statistic.^{16,45,47-49} The AU-ROC extends from 0.5 to 1.0. An AU-ROC=0.5 means that the discriminatory accuracy of the candidate risk factor is similar to that obtained by flipping a coin. That is, a risk factor with an AU-ROC=0.5 is useless. An AU-ROC=1.0 means complete accuracy.

In series of simple logistic regression models we identified the single variables with the highest discriminatory accuracy. Using this information, thereafter, we performed two models. Model A only with the two variables with the higher discriminatory accuracy (i.e., smoking and birth order) and model B which adds the rest of covariates to the initial model A. We ran this second model in order to assess the change of discriminatory accuracy when adding the rest of information to a simpler model. We appraised the incremental value of a model by the difference between AU-ROCs. Due to a problem of collinearity, Stata automatically deleted the two categories in common (i.e., 'only child' and first child) shared by the variables of birth interval and previous child with SGA, keeping them only in the former. All models were stratified by gestational age (i.e., preterm, term and post-term) because it is has been suggested that SGA at term and at preterm may have been driven by a different etiology⁵⁰. We included post-term in order to complete the classification.

We performed the analyses in the whole population, stratifying by gestational age (ie, preterm, term, and post-term). We performed the statistical analyses using STATA 12.0 (College Station, TX: StataCorp LP) and SPSS 21.0 (Armonk, NY: IBM Corp.).

RESULTS

Table 1 shows the maternal and individual characteristics of the population of newborns by SGA status. We see that SGA is much more prevalent among preterm babies (10.14%) than among term (1.87%) and post-term (3.03%) babies. Females show higher prevalence of SGA than males among preterm, and slightly lower prevalence among those born post-term. Regardless of gestational age, firstborns had a higher risk of SGA than their siblings. SGA is more prevalent

among children who had a previous sibling during the same year, except among those babies born at preterm, but this may be due to the larger amount of missing information about the previous siblings (11%). Mothers who had a previous child with SGA are more likely to have a current SGA baby regardless of gestational age.

SGA was more frequent in mothers younger than 20 years of age; among divorced, widowed, and single women; and among those who were born outside Sweden and those with low educational achievement. In babies born SGA, hypertension was more frequent among preterm than among post-term.

Table 1. Prevalence of SGA in the whole population of babies and in strata of gestational age in Sweden 1987–1993

				SGA	stratified by	gestatio	nal age	
	SGA		Term		Preterm		Post-tei	rm
	768 059		676 961		36 080		55 018	
	(2.3%)		(1.9%)		(10.1%)		(3.0%)	
	N	%	Ν	%	Ν	%	Ν	%
Newborn sex								
Male	393 538	2.33	343 549	1.86	19 635	9.20	30 354	3.24
Female	374 521	2.34	333 412	1.87	16 445	11.26	24 664	2.78
Birth order								
First	325 326	3.30	279 426	2.66	18 796	11.67	27 104	4.10
Second	268 998	1.63	242 276	1.31	9 774	8.79	16 948	2.01
<u>></u> Third	173 735	1.63	155 259	1.30	7,510	8.07	10 866	1.97
Birth intervals								
< 1 year	4557	3.55	3841	2.76	415	6.99	301	8.97
1-2 years	77 922	2.00	69 232	1.61	3585	8.23	5105	2.86
>2 years	147 354	1.97	130 979	1.57	5970	9.45	10 405	2.72
'Only child'	341 906	2.77	298 264	2.20	17 789	11.63	25 853	3.29
First child	196 320	1.97	174 645	1.60	8321	8.45	13 354	2.70
Previous child SG	Α							
Yes	3371	9.58	2913	7.69	259	30.12	196	9.18
No	226 462	1.90	201 139	1.52	9711	8.34	15 615	2.80
'Only child'	341 906	2.77	298 264	2.20	17 789	11.63	25 853	3.29
First child	196 320	1.97	174 645	1.60	8321	8.45	13 354	2.70
Maternal age								
<20	40 735	3.06	35 367	2.49	2485	9.93	2883	4.30
20–24	156 736	2.39	138 232	1.91	7842	9.82	10 662	3.24
25–34	481 689	2.19	425 980	1.76	20 796	9.59	34 913	2.89
>35	88 899	2.69	77 382	2.08	4957	12.04	6 560	2.90

Hypertension	3166	8.69	2648	5.44	383	32.11	135	5.93
Smoking								
Non-smoker	541 962	1.77	479 851	1.37	22 603	9.14	39 508	2.41
Light smoker	109 799	3.62	76 327	3.05	5884	11.86	7588	4.40
Heavy smoker	67 454	4.41	58 790	3.79	4234	11.76	4430	5.64
Missing	48 844	2.90	41 993	2.12	3359	11.79	3492	3.78
Marital status								
Married or								
cohabiting	385 173	2.02	342 330	1.61	16,50	9.36	26 343	2.69
Divorced or								
widowed	27 151	3.03	23 475	2.43	1706	10.67	1970	3.55
Single	355 735	2.63	311 156	2.10	17 874	10.81	26 705	3.33
Maternal origin								
Born in Sweden	663 043	2.25	584 683	1.79	30 745	10.19	47 615	2.79
Not born in								
Sweden	95 393	2.83	84 022	2.30	4749	10.07	6 622	4.27
Missing	9623	2.67	8256	3.02	586	8.02	781	7.30
Maternal education	on							
University	181 083	1.95	160,102	1.52	7367	10.26	13 614	2.57
Secondary	400 363	2.22	353 813	1.77	18 457	9.96	28 093	2.74
Primary	140 186	3.21	122 919	2.31	7631	10.78	9636	3.84
Missing	46 427	2.34	40 127	2.67	2625	9.22	3675	4.82

Table 2 indicates that the risk for being SGA was similar in boys and girls. However, as expected, not being a firstborn reduced the risk of being SGA. With respect to maternal characteristics, mothers younger than 20 years and those 35 years and older had higher risk of delivering an SGA baby compared to 20- to 24-year-old mothers. Mothers who had a previous child during the same year have a higher risk of having a SGA baby as well as those who had a previous child with SGA. Mothers who experienced hypertension during pregnancy had a higher risk of delivering SGA babies. Compared to non-smoker mothers, light and heavy smoker mothers had higher risk of delivering an SGA baby. Divorced and widowed mothers as well as single mothers were more likely to deliver an SGA baby than married and cohabiting mothers. Mothers with primary and secondary education had higher risk of delivering SGA babies than mothers with a university degree. Similarly, mothers who were not born in Sweden were at higher risk of delivering an SGA baby.

BMJ Open

Table 2. Measures of association between offspring and maternal characteristics, and being small for gestational age (SGA), in the whole population of babies and in strata of gestational age in Sweden 1987–1993. Values are odds ratios (OR) and [95% confidence intervals (IC)]. Crude models

	linad	SGA	model	SGA T	FERM	model	SGA I	PRETER	M model	SGA F	POST-TE	RM
	OR	IC-95%	6	OR	IC-95%	6	OR	IC-95%	6	OR	IC-95%	6
Newborn sex			-	•		-	•		-	•		-
Male (ref)												
Female	1.01	[0.98	1.04]	1.01	[0.97	1.04]	0.86	[0.84	0.88]	0.85	[0.77	0.94]
Birth order												
First (ref)												
Second	0.48	[0.47	0.50]	0.49	[0.47	0.51]	0.60	[0.59	0.62]	0.48	[0.42	0.54
<u>></u> Third	0.49	[0.47	0.51]	0.48	[0.46	0.50]	0.72	[0.70	0.74]	0.47	[0.41	0.54]
Gestational age	0.80	[0.79	0.81]	0.87	[0.85	0.88]	0.86	[0.85	0.87]	1.77	[1.61	1.95]
Birth intervals												
1-2 years (ref)												
<1 year	1.81	[1.53	2.13]	1.74	[1.42	2.12]	0.84	[0.56	1.24]	3.35	[2.18	5.13]
>2 years	0.98	[0.92	1.05]	0.97	[0.90	1.05]	1.16	[1.00	1.35]	0.95	[0.77	1.16]
'Only child'	1.40	[1.32	1.05]	1.37	[1.29	1.46]	1.47	[1.29	1.67]	1.16	[0.97	1.38]
First child	0.99	[0.93	1.05]	0.99	[0.93	1.07]	1.03	[0.89	1.19]	0.94	[0.78	1.15]
Previous child SGA												
No (ref)												
Yes	5.48	[4.87	6.17]	5.45	[4.74	6.28]	4.75	[3.61	6.24]	3.95	[2.49	6.42]
'Only child'	1.47	[1.42	1.53]	1.46	[1.40	1.52]	1.45	[1.33	1.58]	1.18	[1.05	1.33]
First child	1.04	[0.99	1.09]	1.06	[1.01	1.12]	1.01	[0.91	1.13]	0.97	[0.84	1.11]
Maternal age												
25–34 (ref)												
<20	1.41	[1.33	1.49]	1.42	[1.33	1.53]	1.44	[1.38	1.50]	1.51	[1.25	1.83]
20–24	1.09	[1.05	1.13]	1.09	[1.04	1.14]	1.16	[1.13	1.19]	1.12	[0.99	1.27]
>35	1.23	[1.18	1.29]	1.18	[1.12	1.25]	1.31	[1.27	1.36]	1.00	[0.86	1.17]
Hypertension												
No (ref) vs. yes	4.02	[3.55	4.55]	3.05	[2.58	3.61]	2.73	[2.45	3.04]	2.02	[0.99	4.13]
Smoking												
Non-smoker (ref)												
Light smoker	2.08	[2.01	2.16]	2.27	[2.17	2.37]	1.30	[1.26	1.33]	1.86	[1.64	2.12
Heavy smoker	2.56	[2.46	2.67]	2.84	[2.70	2.98]	1.53	[1.48	1.58]	2.42	[2.10	2.79
Missing	1.66	[1.57	1.76]	1.56	[1.45	1.67]	1.70	[1.64	1.76]	1.59	[1.32	1.92
Marital status												
Married or cohabit (ref)	ing											
Divorced or widowed	1.52	[1.41	1.63]	1.52	[1.40	1.66]	1.51	[1.43	1.59]	1.33	[1.04	1.71

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19	Sing Mat Borr (ref) Not Swe Miss Mat edu Univ Secc Prim Miss
20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36	Figur discr Havi with 0.54 0.59)
36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52	Figu: term varia
53	order

Single	1.31	[1.27	1.35]	1.31	[1.27	1.36]	1.19	[1.17	1.22]	1.25	[1.13	1.38]
Maternal origin												
Born in Sweden												
(ref)												
Not born in	1 27	[1 21	1 2 2 1	1 20	[1 22	1 261	1 07	[1 0/	1 1 1 1	1 56	[1 27	1 771
Sweden	1.27	[1.21	1.52]	1.50	[1.23	1.30]	1.07	[1.04	1.11]	1.50	[1.37	1.//]
Missing	1.66	[1.49	1.84]	1.71	[1.51	1.94]	1.35	[1.24	1.47]	2.74	[2.08	3.61]
Maternal												
education												
University (ref)												
Secondary	1.14	[1.10	1.19]	1.17	[1.12	1.23]	1.13	[1.10	1.17]	1.07	[0.94	1.22]
Primary	1.49	[1.42	1.56]	1.53	[1.45	1.62]	1.35	[1.31	1.39]	1.51	[1.30	1.76]
Missing	1.67	[1.57	1.77]	1.78	[1.66	1.92]	1.42	[1.36	1.49]	1.92	[1.59	2.31]

Figure 2 shows the values for the AU-ROC of the variables included in Table 2. Overall, their discriminatory accuracy was rather low. Newborn sex had the lower discriminatory accuracy. Having a SGA child and hypertension, despite being the risk factors most strongly associated with SGA (OR 5.48 and 4.02, respectively), had a very low discriminatory accuracy (AU-ROC 0.54 and 0.51). Birth order and smoking were the variables with highest accuracy (AU-ROC 0.59).

[Figure 2 around here]

Figure 3 shows the AU-ROC for SGA of different risk factors after stratification by preterm, term, and post-term. As in the non-stratified analysis, the discriminatory accuracy of the variables was low. Smoking at term showed the highest discrimination (AU-ROC 0.60).

[Figure 3 around here]

Figure 4 shows that the discriminatory accuracy of the general model A, including only birth order and smoking, was slightly improved (just 0.05 proportion units), when all variables were included in the full model B. Among preterm babies model B improved the discriminatory

accuracy of the model by 0.1 proportion units, while this improvement was much lower among SGA term and SGA post-term babies.

[Figure 4 around here]

DISCUSSIONS

We were able to verify a number of recognised maternal and newborn risk factors for SGA. For instance, we found that smoking (OR 2.56), and especially having had a previous SGA baby (OR 5.48) and maternal hypertension (OR 4.02), were 'strongly' associated with being SGA. However, even if the magnitude of the ORs was of a size normally considered as undoubtedly relevant in epidemiology, none of those traditional risk factors for SGA provided enough accuracy to discriminate babies with SGA from other babies. In fact, the AU-ROC for having had a previous SGA child and maternal hypertension was slightly higher than 0.5, which means that the accuracy of this variable for discriminating babies with SGA from those without SGA was rather similar to that obtained by flipping a coin. That is, we need to recognise that, although on average, mothers with hypertension were four times more likely to have an SGA baby, many mothers with hypertension delivered babies without SGA, and many SGA babies were born to mothers without hypertension. Our findings, therefore, seriously question the utility of maternal hypertension during pregnancy for planning strategies of prevention against SGA. This statement, however, does not mean that hypertension during pregnancy is irrelevant to understanding the origin of SGA, but rather that we need to determine who among hypertensive mothers is actually prone to deliver an SGA baby.

There is a tacit but fallacious belief that the discriminatory accuracy of a risk factor is high when it is supported by a 'strong' association (e.g, an OR of 4, as in the case of maternal hypertension). However, for an association to be an accurate instrument for discrimination, it must be of a magnitude rarely identified in epidemiologic studies.^{16,51-53} Following our example, a low discriminatory accuracy only indicates that any attempt of intervention based on the existence of the risk factor will be inefficient and even inappropriate, because health professionals will unnecessarily treat many mothers. The decision to start an intervention should seriously take into account the existence of important (physical or emotional) side effects in the false-positive women. That is, it is always important to consider the principle of *primum non*

BMJ Open: first published as 10.1136/bmjopen-2014-005388 on 30 July 2014. Downloaded from http://bmjopen.bmj.com/ on June 12, 2025 at Department GEZ-LTA Erasmushogeschool . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

Compared to the other variables studied, birth order and smoking presented a higher discriminatory capacity. However, their discriminatory accuracy was still very low in absolute terms (AU-ROC \approx 0.59). Also, combining all the variables in the same model did not substantially increase the discriminatory accuracy (AU-ROC=0.69). In other words, our results indicate that we actually do not know so much about what determines being SGA.

The existence of a low discriminatory accuracy suggests that around the population average risk there is considerable individual heterogeneity. Therefore, a logical consequence should be to identify which women are most susceptible to the risk factors. Hence, we explored the discriminatory accuracy of the chosen risk factors in different strata of gestational age at birth. We found that the combination of all variables in the same model had a minor improvement for discriminatory accuracy among those born at term or preterm as well as post-term.

Our finding suggests the existence of individual heterogeneity of responses to some specific variables, so the discriminatory accuracy depended, for instance, on whether the baby was preterm, term, or post-term. In fact, smoking, birth order, maternal origin, and marital status had a lower discriminatory capacity among preterm than among babies at term. On the contrary, newborn sex and hypertension had a higher discriminatory accuracy among preterm than among term babies. In the same way, newborn sex and maternal origin had a lower discriminatory capacity among post-term babies, while for smoking and having had a previous SGA baby we found the opposite relation. The variation of the magnitude of the discriminatory accuracy by gestational age at birth expresses the existence of individual heterogeneity.

In addition, the definition of SGA may also actively contribute to reducing the discriminatory accuracy of the traditional risk factors since discrimination depends both on the outcome and as on the exposure. Thus, low discrimination can result from the fact that SGA fails to distinguish between pathological and constitutionally small babies, that is, to properly capture the health dimension that it is supposed to be a proxy for (IUGR)¹. In order to address this shortcoming, we stratified SGA by gestational age as this has been identified as a good strategy to distinguish between these two ⁵⁰. However, we do not find support for this approach since we found a lower discriminatory accuracy among preterm (presumably pathological small) than among term SGA babies. In this regard, our findings show awareness of the caveats pointed out by previous studies

on the use of SGA as a proxy for IUGR¹, and encourage further research aiming to better capture IUGR.

Our findings have important research and policymaking implications. A possible reason for the low discriminatory accuracy of many average associations is that average effects are a mixture of individual level effects and therefore mix inter individual heterogeneity (ie, some individuals respond intensively to the exposure, while others are resilient or might even respond in the opposite direction). The approach based on discriminatory accuracy understands average effects as an idealised mean value that does not necessarily represent the heterogeneity of individual effects.²² Some scholars prefer to conceive individual outcome as the expression of a stochastic phenomenon that is best estimated by the average risk using a probabilistic approach.⁵⁴ Our understanding instead is that individual outcome reflects the interindividual heterogeneity of responses that can be potentially determined; lack of knowledge could be amended by a better understanding of individual responses.⁵⁵ See elsewhere for a longer explanation of these ideas.^{22,56,57} From this perspective, reducing exposure to a risk factor would only be effective when acting on the susceptible, but not on the resilient, individuals. For instance, we need to better capture babies who suffer from IUGR, since, so far, we are incapable of distinguishing between babies who are constitutionally small from those who are pathologically growth restricted.¹ By stratifying between preterm, term, and post-term, we might be able to better approach the underlying heterogeneity.

From the policymaking perspective, our findings suggest that hitherto there has not been enough knowledge to identify any specific risk factor or combination of them that could discriminate with accuracy children with and without SGA status. Our findings support policymaking oriented to lifestyle modification, as according to the principle of *primum non nocere*,²² they have mostly positive consequences, even for 'false-positive' mothers. For instance, persuading women to quit smoking reduces the risk of SGA in some babies, but it improves general wellbeing in everyone. However, other risk factors with low discriminatory accuracy that lead to pharmacological treatment or screening might result in unnecessary side effects and cost. In the long run, an uncritical use of variables with low discriminatory accuracy may hinder the identification of pertinent risk factors and susceptible individuals and damage the scientific credibility of modern epidemiology.^{22,56,57}

 Our conclusions are based on classical measures of discriminatory accuracy such as the AU-ROC curve. These measures have been criticised as insensitive to small changes in predicted individual risk.⁵⁸ Some authors propose more specific measures of reclassification, like the net reclassification improvement (NRI), and the integrated discrimination improvement (IDI).⁵⁹⁻⁶² We applied NRI and IDI in a sensitivity analysis (results not shown in tables). For example, using NRI, we observed a reclassification of 30%. However, this figure does not add substantial information to our results, since NRI (as well as IDI) refers to the misclassification occurring all along the risk scale, instead of capturing the misclassification which takes place around the fixed threshold. Furthermore, the new NRI and IDI measures have also been criticised,⁶³ and some authors⁶⁴ have explicitly advised against their use in common epidemiological practice because, unlike IDI and NRI, traditional measures of discrimination like the AU-ROC curve have the advantage that prognostic performance cannot be manipulated.⁶⁴ Therefore, we preferred to quantify discriminatory accuracy by analysing ROC curves and AU-ROCs.

Our analyses are based on a national medical registry covering almost the entire population of residents in Sweden. Nearly all births are registered in the MBR, because giving birth at home is very unusual in Sweden. In addition, estimation of SGA is routinely calculated at the MBR following standard intrauterine growth curves.²⁷ However, our study also has a number of limitations. Because of lack of data, we could not assess many other variables identified in the literature as 'risk factors', such as genetic or nutritional factors.^{15,65} In spite of the quality of the MBR, the information regarding smoking is based on a self-reported questionnaire (anamnesis) administered by the midwife at the first antenatal visit (ie, between 10 and 12 gestational weeks), which to some extent might bias the result by including misclassification of exposure.⁶⁶ However, a study conducted in Sweden comparing self-reported nicotine exposure and plasma levels of cotinine in early and late pregnancy concluded that self-reported smoking information had acceptable validity.⁶⁷

Unfortunately, we could not identify those mothers who suffered from preeclampsia, for which the discriminatory accuracy concerning SGA may be higher than for hypertension in our model. Further analysis on this aspect is required. Another limitation of our study is that we calculated the discriminatory accuracy in the same sample used for constructing the predicted model. This procedure, however, might overestimate the discriminatory accuracy of the models, so the low discriminatory accuracy found may be an underestimation.

In conclusion, applying measures of discriminatory accuracy rather than measures of association only, our study revisits known risk factors of SGA and discusses their role from a public health perspective. We found that neither models including simple variables nor models including several variables at the same time have a good discriminatory accuracy to discriminate babies with SGA from those without SGA. This finding is of fundamental relevance to address future research and to design policymaking recommendations in a more informed way.

As noted elsewhere,^{56,57} there is need of a new epidemiological approach that systematically provides information on the discriminatory accuracy and interindividual heterogeneity of effects and does not rely only on average measures of association.⁶⁸ In this line, new statistical methods like logic regression seem promising.^{69,70} A fundamental change is needed in the way traditional risk factors are currently interpreted in public health epidemiology. If the *discriminatory accuracy* of most classical risk factors is very low, what happens with the vast majority of recommendations given so far in epidemiology and public health? Are health professionals misleading the community by raising the alarm about risks that may be harmless for most individuals? What are the ethical repercussions of using risk factors with low discriminatory accuracy? Are there problems of inefficiency, medicalisation, and stigmatisation? We believe that these questions have a high significance for both the community and the future of public health research.

BMJ Open: first published as 10.1136/bmjopen-2014-005388 on 30 July 2014. Downloaded from http://bmjopen.bmj.com/ on June 12, 2025 at Department GEZ-LTA Erasmushogeschool .

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

Acknowledgments

We would like to thank Dr. Karin Kallén for her very helping comments on a previous version of this paper. This work was supported by the Swedish Research Council (VR) (Dnr #2013-2484, PI Juan Merlo), the Centre for Economic Demography and the SIMSAM early life Lund (Dnr #2013-5474). The funders had no role in study design, data collection, and analysis; decision to publish; or preparation of the manuscript.

Authors' contributions

Merlo had the original idea of applying measures of discriminatory accuracy for the interpretation of risk factors and discussed it with Wagner and Juarez. Merlo and Juárez initiated the study. Merlo, Juárez, and Wagner contributed to the design of the study; Juárez performed the analyses under the supervision of Merlo and Wagner. Juárez wrote the first draft of the manuscript, and Merlo contributed to the writing of the final version. All authors made substantial contributions to the interpretation of the results and manuscript revision. All authors approved the final version of the manuscript.

Competing interest statement

No additional data available

Figure 1. Flow diagram showing the individuals excluded from the study population.

Figure 2. Area under the ROC curve (AU-ROC) to compare the discriminatory accuracy of different models to distinguish between SGA and non-SGA babies.

Figure 3. Area under the ROC curve (AU-ROC) for specific maternal and newborn characteristics.

Figure 4. Area under the ROC curve (AU-ROC) for specific maternal and newborn characteristics after stratifying by gestational age (preterm, term, and post-term).

The authors have declared that no competing interests exist.

REFERENCES

- **1.** Urquia ML, Ray JG. Seven caveats on the use of low birthweight and related indicators in health research. *J Epidemiol Community Health.* 2012;66(11):971-975.
- Grisaru-Granovsky S, Reichman B, Lerner-Geva L, et al. Mortality and morbidity in preterm small-for-gestational-age infants: a population-based study. *Am J Obstet Gynecol.* . 2012;206(2):150e151-e156.
- **3.** Streimish IG, Ehrenkranz RA, Allred EN, et al. Birth weight- and fetal weight-growth restriction: impact on neurodevelopment. *Early Hum Dev.* 2012;88(9):765-771.
- **4.** Eikenes L, Martinussen MP, Lund LK, et al. Being born small for gestational age reduces white matter integrity in adulthood: a prospective cohort study. *Pediatr Res.* . 2012;72(6):649-654.
- **5.** Moore GS, Kneitel AW, Walker CK, Gilbert WM, Xing G. Autism risk in small- and large-for-gestational-age infants. *Am J Obstet Gynecol.* 2012;206(4):314.
- **6.** Nielsen PR, Mortensen PB, Dalman C, et al. Fetal growth and schizophrenia: A nested case-control and case-sibling study. *Schizophr Bull.* 2012;12(): .
- **7.** Løhaugen GC, Ostgård HF, Andreassen S, et al. Small for gestational age and intrauterine growth restriction decreases cognitive function in young adults. *J. Pediatr.* 2013; (): .
- **8.** Wingren CJ, Agardh D, Merlo J. Revisiting the risk of celiac disease in children born small for gestational age: a sibling design perspective. *Scand J Gastroenterol.* 2012;47(6):632-639.
- **9.** van de Lagemaat M, Rotteveel J, van Weissenbruch MM, Lafeber HN. Small-for-gestational-age preterm-born infants already have lower bone mass during early infancy. *Bone.* 2012;51(3):441-446.
- **10.** Forssell L, Cnattingius S, Bottai M, et al. Increased risk of Barrett's esophagus among Individuals born preterm or small for gestational age. *Clin Gastroenterol Hepatol.* . 2013; (): .
- **11.** Forssell L, Cnattingius S, Bottai M, Lagergren J, Ekbom A, Akre O. Risk of esophagitis among individuals born preterm or small for gestational age. *Clin Gastroenterol Hepatol.* . 2012;10(12):1369-1375.
- Dalla Pozza RD, Bechtold S, Putzker S, Bonfig W, Netz H, Schwarz HP. Young adults born small for gestational age: is reduced baroreceptor sensitivity a risk factor for hypertension? *Clin Cardiol.* . 2006;29(5):215-218.
- **13.** Crispi F, Figueras F, Cruz-Lemini M, Bartrons J, Bijnens B, Gratacos E. Cardiovascular programming in children born small for gestational age and relationship with prenatal signs of severity. *Am J Obstet Gynecol.* 2012;207(2):121.e121–121.e129.
- **14.** Kramer MS. Determinants of low birth weight: methodological assessment and meta-analysis. *Bull. World Health Organ.* 1987;65(5):663-737.
- **15.** McCowan L, Horgan RP. Risk factors for small for gestational age infants. *Best Pract Res Clin Obstet Gynaecol.* 2009;23(6):779-793.
- **16.** Pepe MS, Janes H, Longton G, Leisenring W, Newcomb P. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *Am J Epidemiol.* May 1 2004;159(9):882-890.
- **17.** Catov JM, Nohr EA, Olsen J, Ness RB. Chronic hypertension related to risk for preterm and term small for gestational age births. *Obstet Gynecol.* 2008;112(2):290-296.
- Zethelius B, Berglund L, Sundstrom J, et al. Use of multiple biomarkers to improve the prediction of death from cardiovascular causes. *New England Journal of Medicine*. May 15 2008;358(20):2107-2116.
- **19.** Cooney MT, Dudina AL, Graham IM. Value and limitations of existing scores for the assessment of cardiovascular risk: a review for clinicians. *J Am Coll Cardiol*. Sep 29 2009;54(14):1209-1227.

BMJ Open: first published as 10.1136/bmjopen-2014-005388 on 30 July 2014. Downloaded from http://bmjopen.bmj.com/ on June 12, 2025 at Department GEZ-LTA Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

Law MR, Wald NJ, Morris JK. The performance of blood pressure and other cardiovascular risk

1

2 3

4

5

6 7

8

9

10

11

12

13

14 15

16

17

18

19

20

21

22 23

24

25

26

27

28

29

30 31

32

33

34

35

36

37

38 39

40

41

42

43

44

45 46

47

48

49

50

51

52

53 54

55

56

57

58 59 60 20.

factors as screening tests for ischaemic heart disease and stroke. J Med Screen. 2004;11(1):3-7. 21. Melander O, Newton-Cheh C, Almgren P, et al. Novel and conventional biomarkers for prediction of incident cardiovascular events in the community. JAMA : the journal of the American Medical Association. Jul 1 2009;302(1):49-57. 22. Merlo J. Multilevel analysis of individual heterogeneity: a fundamental critique of the current probabilistic risk factor epidemiology (invited commentary). American Journal of Epidemiology. 2014; In press. 23. Facts about mothers, childbirth and newborn children. Medical Birth Registry, 1973 to 2000. Stockholm: Centre of epidemiology;2002. 24. Cnattingius S, Ericson A, Gunnarskog J, Kallen B. A quality study of a medical birth registry. Scand J Soc Med. Jun 1990;18(2):143-148. Cunningham FG. Williams Obstetrics. Third ed: McGraw-Hill; 2005. 25. 26. Källén B. A birth weight for gestational age standard based on data in the Swedish Medical Birth Registry, 1985-1989. European Journal of Epidemiology. 1995;11(5):601-606. 27. Marsál K, Persson PH, Larsen T, Lilja H, Selbing A, Sultan B. Intrauterine growth curves based on ultrasonically estimated foetal weights. Acta Paediatr. 1996;85(7):843-848. 28. Alexander GR, Kogan MD, Himes JH. 1994-1996 U.S. singleton birth weight percentiles for gestational age by race, Hispanic origin, and gender. Matern. Child Health J. 1999;3(4):225-232. 29. Swamy GK, Edwards MJ, Gelfand A, James SA, Miranda ML. Maternal age, birth order, and race: differential effects on birthweight. J. Epidemiol. Community Health. 2012;66(2):136-142. 30. Elshibly EM, Schmalisch G. The effect of maternal anthropometric characteristics and social factors on gestational age and birth weight in Sudanese newborn infants. BMC Public Health. 2008;8(244): . 31. Grisaru Granovsky S, Gordon ES, Haklai Z, Samueloff A, Schimmel MM. Effect of interpregnancy interval on adverse perinatal outcomes--a national study. Contraception. 2009;80(6):512-518. 32. Voskamp B, Kazemier BM, Ravelli AC, Schaaf J, Mol BW, Pajkrt E. Recurrence of small-forgestational-age pregnancy: analysis of first and subsequent singleton pregnancies in The Netherlands. Am J Obstet Gynecol. 2013;208(5):374.e371-376. 33. Kramer MS, Séguin L, Lydos J, Goulet L. Socio-economic disparities in pregnancy outcome: why do the poor fare so poorly? Paediatric and Perinatal Epidemiology. 2000;14:194-210. 34. Auger N, Park AL, Harper S, Daniel M, Roncarolo F, Platt RW. Educational inequalities in preterm and term small-for-gestational-age birth over time. Ann Epidemiol. 2012;22(3):160-167. 35. Shah PS, Zao J, Ali S, births. KSGoDopL. Maternal marital status and birth outcomes: a systematic review and meta-analyses. Matern. Child Health J. 2011;15(7):1097-1109. 36. Castro T. Single motherhood and low birthweight in Spain: Narrowing social inequalitues in health? Demographic Research. 2010;22(27):863-890. 37. Fraser AM, Brockert JE, Ward RH. Association of young maternal age with adverse repreductive outcomes. The New England Journal of Medicine. 1995;332(17):1113-1118. 38. Odibo A, Nelson D, Stamilo DM, Sehdev HM, Macones GA. Advanced maternal age is an independent risk factor for intrauterine growth restriction. American Journal of Perinatology. 2006;23(5):325-328. 39. Campbell MK, Cartier S, Xie B, Kouniakis G, Huang W, Han V. Determinants of small for gestational age birth at term. Paediatr Perinat Epidemiol. 2012;26(6):525-533. 40. Rogers JM. Tobacco and pregnancy: overview of exposures and effects. Birth defects research 2008;Part C(84):1-15. 41. Kabir Z, Daly S, Clarke V, Keogan S, Clancy L. Smoking Ban and Small-For-Gestational Age Births in Ireland. PLoS ONE. 2013;8(3):e57441. . For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

BMJ Open

42.	Mitchell EA, Thompson JM, Robinson E, et al. Smoking, nicotine and tar and risk of small for
43.	Juárez SP, Merlo J. Revisiting the Effect of Maternal smoking during pregnancy on offspring
44.	birthweight: a quasi-experimental sibling analysis in Sweden. <i>PLoS ONE</i> . 2013;8(4):e61734. Li X, Sundquist K, Sundquist J. Risks of small-for-gestational-age births in immigrants: A
45	nationwide epidemiological study in Sweden. <i>Scand. J. Public Health.</i> 2012;40:634-640. Pene MS. The statistical evaluation of medical tests for classification and prediction. Oxford :
43.	New York: Oxford University Press; 2003.
46.	Zweig MH, Broste SK, Reinhart RA. ROC curve analysis: an example showing the relationships among serum lipid and apolipoprotein concentrations in identifying patients with coronary artery disease. <i>Clin Chem.</i> Aug 1992;38(8 Pt 1):1425-1428.
47.	Pepe MS, Janes H, Gu JW. Letter by Pepe et al regarding article, "Use and misuse of the receiver operating characteristic curve in risk prediction". <i>Circulation</i> . Aug 7 2007;116(6):e132; author reply e134.
48.	Royston P, Altman DG. Visualizing and assessing discrimination in the logistic regression model. <i>Statistics in Medicine</i> . Oct 30 2010;29(24):2508-2520.
49.	Gerds TA, Cai T, Schumacher M. The performance of risk prediction models. <i>Biometrical journal. Biometrische Zeitschrift</i> . Aug 2008;50(4):457-479.
50.	Ananth CV, Vintzileos AM. Distinguishing pathological from constitutional small for gestational age births in population-based studies. <i>Early Hum Dev.</i> 2009;85(10):653-658.
51.	Boyko EJ, Alderman BW. The use of risk factors in medical diagnosis: opportunities and cautions. <i>Journal of clinical epidemiology</i> . 1990;43(9):851-858.
52.	Wald NJ, Hackshaw AK, Frost CD. When can a risk factor be used as a worthwhile screening test? <i>BMJ.</i> Dec 11 1999;319(7224):1562-1565.
53.	Khoury MJ, Newill CA, Chase GA. Epidemiologic evaluation of screening for risk factors: application to genetic screening. <i>American journal of public health.</i> Oct 1985;75(10):1204-1208.
54.	Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. <i>Circulation.</i> Feb 20 2007;115(7):928-935.
55.	Zernicka-Goetz M, Huang S. Stochasticity versus determinism in development: a false dichotomy? <i>Nature reviews. Genetics.</i> Nov 2010;11(11):743-744.
56.	Merlo J, Wagner P. The tyranny of the averages and the indiscriminate use of risk factors in public health: a call for revolution. <i>European journal of epidemiology</i> . 2013;28(1, Supplement):148.
57.	 Merlo J, Wagner P, Juarez S, Mulinari S, Hedblad B. The tyranny of the averages and the indiscriminate use of risk factors and population attributable fractions in Public Health: the case of coronary heart disease. Working paper version 2013-09-26. Unit for Social Epidemiology, Department of Clinical Sciences, Faculty of Medicine, Lund University. 2013;http://www.med.lu.se/english/klinvetmalmo/unit_for_social_epidemiology/abstracts_c_
58.	working papers c. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. <i>Circulation</i> , Feb 20 2007:115(7):928-935.
59.	Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. <i>Statistics in Medicine</i> Jan 30 2008;27(2):157-172
60.	Pencina MJ, D'Agostino RB, Pencina KM, Janssens AC, Greenland P. Interpreting incremental
61.	Pencina MJ, D'Agostino RB, Sr., Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. <i>Stat Med.</i> Jan 15 2011;30(1):11-21

- **62.** Hilden J. On NRI, IDI, and "good-looking" statistics with nothing underneath. *Epidemiology*. 2014;25(2):265-267.
- **63.** Pepe MS. Problems with risk reclassification methods for evaluating prediction models. *Am J Epidemiol.* Jun 1 2011;173(11):1327-1335.
- **64.** Hidden J, Gerds T. Evaluating the impact of novel biomarkers: Do not rely on IDI and NRI. Department of Biostatistics, University of Copenhagen, Denmark. 2012;Research Report 12/08.
- **65.** Kramer MS. Determinants of low birth weight: methodological assessment and meta-analysis. *Bulletin of World Health Organization.* 1984;65(5):663-737.
- **66.** Frisell T, Öberg S, Kuja-Halkola R, Sjölander A. Sibling comparison designs: bias from non-shared confounders and measurement error. *Epidemiology*. 2012;23(5):713-720.
- **67.** Cnattingius S, Ericson A, Gunnarskog J, Källen K. A quality study of a medical birth registry. *Scand. J. Soc. Med.* 1990;18:143-148.
- **68.** Grove A. Rethinking clinical trials. *Science.* Sep 23 2011;333(6050):1679.

- **69.** Kooperberg C, Ruczinski I, LeBlanc ML, Hsu L. Sequence analysis using logic regression. *Genet Epidemiol.* 2001;21(1):s626-631.
- **70.** Janes H, Pepe M, Kooperberg C, Newcomb P. Identifying target populations for screening or not screening using logic regression. *Stat Med.* . 2005;15(24):1321-1338.




297x420mm (300 x 300 DPI)

- 7
- 8
- 9 10
- 11
- 12 13
- 14
- 15 16
- 17 18
- 19 20
- 21 22 23
- 24 25 26
- 28 29 30

27

- 31 32
- 33 34 35
- 36 37 38
- 39 40
- 41 42 43
- 44 45 46
- 47 48 49 50
- 51 52 53 54 55

0.9 0.8 0.7 0.6 JO 0.4 0.3 0.2 0.1 and the set and the set and the ten and Birth AU-RO 0.53 0.50 CI-95% [0.519 [0.498 AU-ROC CI-95% 0.51 [0.502 0.55 [0.541 Newborn sex Pret Term 0.537] Maternal education Pret 0.506] Term 0.55 0.55 0.54 0.54 Post Pret 0.52 [0.507 [0.509 0.531] 0.516] Post Pret [0.534 [0.535 **Birth interval** Hypertension [0.537 Term 0.50 [0.503 0.504] Term 0.53 0.55 0.55 [0.537 [0.516 [0.543 [0.549 0.50 [0.499 [0.508 0.503] Post Post Pret

Previous SGA

Birth orde

Smoke

Pret

Term Post

Pret

Term Post

Pret

Term

Post

0.53 [0.545 [0.515 [0.537

0.53 0.59 0.59 0.53 0.60 0.58

Maternal age

Marital status

ternal origin

0.52 0.52 0.52 0.50

0.52 0.54 0.52

0.53

Term

Post

Pret

Term Post

Pret

Term 0.54

Post

[0.507 [0.507 [0.497 0.525] 0.532] 0.532]

[0.518 [0.527

[0.511

[0.532 [0.516

0.524] 0.546] 0.528]

0.540

0.541]

0.521]

0.549]

0.563

0.553]

0.547]

0.543]

0.5571

0.540]

0.596] 0.602]

0.543]

0.593]

[0.537 [0.587 [0.578 [0.526

[0.596 0.605]

[0.568

297x420mm (300 x 300 DPI)

BMJ Open



297x420mm (300 x 300 DPI)