Author(s): K Hemming, JL Hutton, SV Glinianaia, SN Jarvis and MJ Platt

Article Title: Differences between European birthweight standards: impact on classification of 'small for gestational age'

Year of publication: 2006

Link to published version: http://dx.doi.org/10.1017/S0012162206001988

Publisher statement: None

Differences between European birthweight standards: impact on classification of 'small for gestational age'

K Hemming PhD;
J L Hutton* PhD CStat, Department of Statistics, University of Warwick, Coventry;
S V Glinianaia MD, School of Population and Health Sciences;
S N Jarvis MD, Institute of Child Health, University of Newcastle, Newcastle upon Tyne;
M J Platt MB BS MPH MFPH FRCPCH, Department of Public Health, The University of Liverpool, Liverpool, UK.

**Correspondence to second author at* Department of Statistics, University of Warwick, Coventry CV4 7AL, UK.
E-mail: J.L.Hutton@warwick.ac.uk

We describe a quantitative and comparative review of a selection of European birthweight standards for gestational age for singletons, to enable appropriate choices to be made for clinical and research use. Differences between median values at term across standards in 10 regions and misclassification of 'small for gestational age' (SGA), were studied. Sex and parity differences, exclusion criteria, and methods of construction were considered. There was wide variation between countries in exclusion criteria, methods of calculating standards, and median birthweight at term. The lightest standards (e.g. France's medians are 255g lower than Norway's medians) were associated with fewer exclusion criteria. Up to 20% of the population used in the construction of the Scottish standard would be classified as SGA using the Norwegian standard. Substantial misclassification of SGA is possible. Assumptions about variation used in the construction of some standards were not justified. It is not possible to conclude that there are real differences in birthweight standards between European countries. Country-based standards control for some population features but add misclassification due to the differing ways in which standards are derived. Standards should be chosen to reflect clinical or research need. If standards stratified by sex or parity are not available, adjustments should be made. In multinational studies, comparisons should be made between results using both a common standard and country-based standards.

Birthweight is often used as an outcome measure or indicator of perinatal and infant health in both clinical practice and perinatal research. Infants described as 'small for gestational age' (SGA; e.g. 10th centile) are recognized to be at higher risk of poor perinatal and infant outcomes. A qualitative review of birthweight standards included a discussion of differing clinical expectations of birthweight.[1] However, the birthweight standards that are used to determine SGA are often applied without consideration of their appropriateness to the population under study. Birthweight for gestational age depends on many factors, both intrinsic (e.g. fetal, sex, ethnicity, maternal parity, and plurality of pregnancy) and extrinsic (e.g. maternal smoking and obesity). Further, the inclusion and exclusion criteria used in constructing birthweight standards vary, with some standards including all live-born infants, and others having more specific inclusion and exclusion criteria. Standards that are most applicable to the infants under study (i.e. reflecting the population from which the infant comes) are often not available, leaving both clinicians and researchers to use less specific standards. The different methods used in constructing standards may result in misclassification and affect the conclusions that clinicians and researchers draw. Thus, to assess the robustness of conclusions about rates of morbidity (e.g. cerebral palsy in above-average- and below-average-weight European infants,[2] childhood diabetes,[3] or perinatal mortality[4]) it is important to consider both the populations included and differences in methods of creating standards.

Previous studies have highlighted observed differences in birthweights between and within countries that might be attributable to population characteristics,[5] and variations have been found when comparing international standards, particularly in the preterm period.[6] When measuring optimum birthweight in terms of lowest risk of perinatal mortality, other authors have reported variations among Western European countries,[7] demonstrating that an inappropriate choice of birthweight standard may lead to misclassification. However, clinicians and researchers may not have access to robust standards that apply to their particular population.

This paper presents a comparison of a selection of birthweight standards known to be in current use. In differentiating between the variations brought about by different inclusion criteria, rather than by presumed country variability, we discuss the implications of the choice of birthweight standard on clinical practice and suggest assumptions and methods that may be used to overcome the problems identified.

Method

SELECTION OF STANDARDS

There are many published birthweight standards derived for European national and subnational populations. We focused our assessment of variability on 10 standards drawn from several countries and used by one European collaboration, the Surveillance of Cerebral Palsy in Europe (SCPE).[8] Information on the standards was extracted from published papers, and summaries of the content and methods of construction (such as inclusion criteria) were compiled (Table I).

We used singleton standards stratified by sex but not parity, because the majority of the standards did not include parity. We explored differences by parity where available. Two of the standards (North of England[9] and Aberdeen[10]), provided separate male and female singleton standards for primiparous and multiparous births. Male and female standards were produced for these two standards by combining the multiparous and primiparous standards, assuming equal numbers of first and subsequent births, because only a small proportion of mothers have more than two children. One of the standards (Aberdeen) provided an empirical and Gaussian version; we used the Gaussian version because estimates based on the empirical standard can be less stable as a result of the small cohort size.

We focused on comparisons of the medians – a measure of central tendency – and conducted only limited comparisons of the 10th centile, one of the cut-offs used for SGA. For standards that did not provide the median, the mean or mode were used instead. Two standards, Sweden A[11] and Sweden B,[12] did not provide the 10th or 90th centiles and so approximate centiles were constructed from the mean and standard deviations (SD), assuming normality. France provides

Table I: Composition of singleton European birthweight standards

Area	*Reference*	*Gestation (wks)*	*Birth years*	*Number[a]*	*Exclusions*
Norway	16	20–44	1987–1998[b]	104 730[c]	SB, SM, CS, EO
Sweden A	11	28–42	1977–1981	362 280	SB, SM, MC, EO
Sweden B	12	25–45	1985–1989	484 176	SB, RD, EO
North of England	9	22–42	1983–1991	118 284	ASB, SM, UGA
Oxford, UK	18	24–42	1978–1984[d]	20 487	SB, SM, ED, EO
Scotland	17	24–42	1975–1989	877 061	SB
Aberdeen, UK	10	32–42	1979–1983	14 219	SB
Denmark	13	28–40	1983	N	N
Sassari, Italy	15	34–42	1985–1985	4638	SB, SM, R, UGA
France	14	28–42	1984–1988	88 706	SB, EO

[a]Number of births after exclusions; [b]for preterm infants, years 1967–98; [c]number before exclusions; [d]1978–1981 for all gestational ages, 1982–1984 for gestational age less than 35 weeks. SB, stillbirths; SM, severe malformations; CS, caesarean sections; EO, extreme outliers; MC, maternal complications; RD, residual distribution; ASB, antepartum stillbirths; UGA, uncertain gestational age; ED, elective deliveries; R, race exclusions; N, not stated.

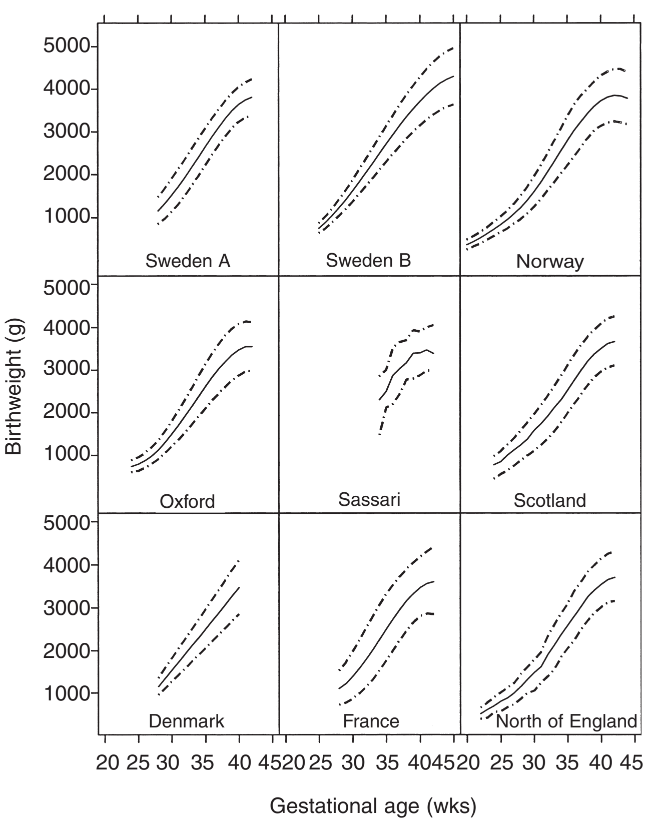

Figure 1: *Male birthweights (g) by gestational age. Solid lines, medians; dot-dashed lines, 10th and 90th centiles (except for France, for which these indicate 5th and 95th centiles respectively).*

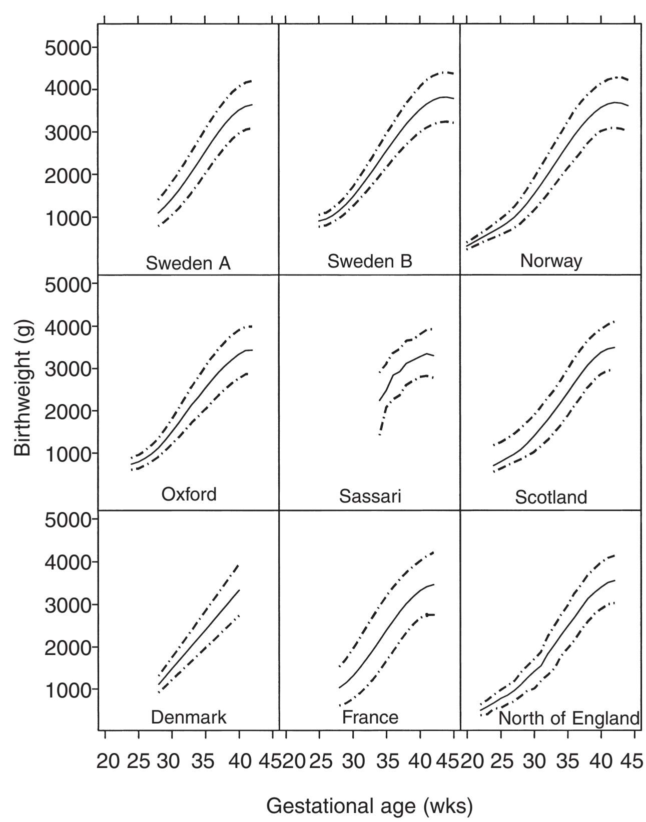

Figure 2: *Female birthweights (g) by gestational age. Solid lines, medians; dot-dashed lines, 10th and 90th centiles (except for France, for which these indicate 5th and 95th centiles respectively).*

5th and 95th centiles. Pairwise comparisons were all made at term (40wks), and other comparisons were made over weeks 28 to 42 (or, if unavailable in the standard, part of this period).

STATISTICAL METHODS

Simple numerical summaries of the male and female standards (Figs 1 and 2) were provided by plotting median birthweight, 10th centiles, and 90th centiles against gestational age. We evaluated the magnitude of differences between male and female standards (Table II), and multiparous and primiparous standards (Table III) for both median and 10th centiles, using those standards that provided the necessary stratification.

Differences between pairs of standards were summarized by mean differences, at term, in median birthweight, for males (Table IV) and females (Table V). The tabulated values were the median birthweight of the standard in the row minus that of the standard in the column. A negative value means that the row standard was lighter than the column standard. For example, the median for Sweden A males was 77g less than that for Sweden B. The average mean difference for each standard in comparison with all other standards was also calculated. We presented observed differences only, and avoided multiple comparison testing. To assess the effect of the published standards on classifications of SGA, we calculated, under the

Table II: Differences in birthweight between male and female standards (males minus females)

Gestational age (wks)	*Male minus female, median*		*Male minus female, 10th centile*	
	Diff. (g)	*Diff./av. (%)*	*Diff. (g)*	*Diff./av. (%)*
28	62	2.7	56	5.3
29	69	2.7	64	5.2
30	76	2.6	70	5.0
31	79	2.4	76	4.7
32	86	2.3	82	4.5
33	90	2.1	85	4.2
34	96	2.1	92	4.1
35	98	1.9	100	3.8
36	106	1.9	94	3.7
37	120	2.0	116	3.9
38	122	1.9	132	3.8
39	146	2.2	135	4.3
40	144	2.1	137	4.1
41	151	2.1	155	4.1
42	155	2.1	168	4.2
Mean	107	2.2	104	4.3

Percentages are male minus female difference as a percentage of average birthweight over all standards. Diff, difference; av, average.

Table III: Differences in European multiparous and primiparous birthweight standards

Gestational age (wks)	*Male infants: multiparous minus primiparous*				*Female infants: multiparous minus primiparous*			
	Median		*10th centile*		*Median*		*10th centile*	
	Diff. (g)	*Diff./av. (%)*	*Diff. (g)*	*Diff./av. (%)*	*Diff. (g)*	*Diff./av. (%)*	*Diff. (g)*	*Diff./av. (%)*
28	110	9.4	230	24.1	50	4.5	90	10.7
29	120	9.2	255	23.9	70	5.6	110	11.8
30	120	8.2	280	23.7	70	4.9	120	11.5
31	120	7.3	265	20.2	80	5.0	120	10.3
32	143	7.5	256	17.8	221	12.8	120	9.3
33	35	1.8	63	4.3	47	2.4	102	7.0
34	85	3.7	258	15.1	50	2.4	–4	–0.3
35	80	3.2	148	7.6	58	2.4	20	1.1
36	69	2.5	130	6.1	0	0.0	–39	–1.8
37	103	3.5	194	8.2	58	2.0	81	3.4
38	125	3.9	102	3.9	90	2.9	34	1.3
39	115	3.4	96	3.4	118	3.7	55	2.0
40	130	3.7	109	3.7	100	3.0	75	2.6
41	126	3.5	105	3.5	100	2.9	74	2.5
42	83	1.7	–13	–0.4	53	1.5	87	2.9

These results are based on three standards, Scotland, Aberdeen, and the North of England, which were stratified by parity. Percentages for each gestational age are difference between multiparous and primiparous as a percentage of average birthweight over all standards. Diff, difference; av, average.

assumption that term birthweights follow a normal distribution, the proportion of term female infants who would be classified as SGA when SGA was defined as a birthweight below the 10th centile, for a subset of standards (Table VI).

Two standards were constructed by assuming that the SD is a constant proportion of the mean across all gestational ages.[12,13] To assess this assumption, the coefficient of variation (CoV), which is the SD divided by the mean and multiplied by 100, was plotted against gestational age (Fig. 3).

Results

STANDARDS INCLUDED

We included the following 10 European standards: France,[14] Sassari (Sardinia, Italy),[15] Denmark,[13] Norway,[16] Sweden (A, years of birth 1977–1981),[11] Sweden (B, years of birth 1985–1989),[12] Aberdeen (Scotland),[10] Scotland,[17] North of England,[9] and Oxford (UK).[18] Table I summarizes the standards. No composition information was available for the Danish standard, so it is excluded from the discussions immediately following but is included in the analysis.

The North of England, Aberdeen, and Sassari standards had a regional basis, whereas the others were drawn from the national population. All standards were based on a geographical population except those for Oxford, which used regional hospital-based populations, and France, which used maternity units from various regions across France. All were derived from routine population data, except the North of England standard, which was based on routine data combined with data from a special study for gestational ages 22 to 31 weeks. The numbers of births on which the standards were based, after exclusions, ranged from about 5000 (Sassari), to nearly 1 million (Scotland).

All standards excluded cases with incomplete records on gestational age, birthweight, sex, plurality, parity (where applicable), and infants born at gestational ages outside those included in the standard. We considered only standards for singletons; by definition, singleton standards excluded multiple births. None of the standards considered race as a stratifying factor, although the Sassari standard excluded births from non-Sardinian parents. There were large variations in the proportion of births excluded from the populations from which the standards were constructed: from less than 1% (Scotland, only stillbirths) to 21% (Sweden A). A brief summary of the details of exclusions was made (Table I), and full details can be found in the relevant references for the particular standard. For example, the 'extreme outliers' excluded might be 4SD or 5SD from the mean,[11,12,14,16,18] or births might be excluded to achieve a normal 'residual distribution'[12] about the mean.

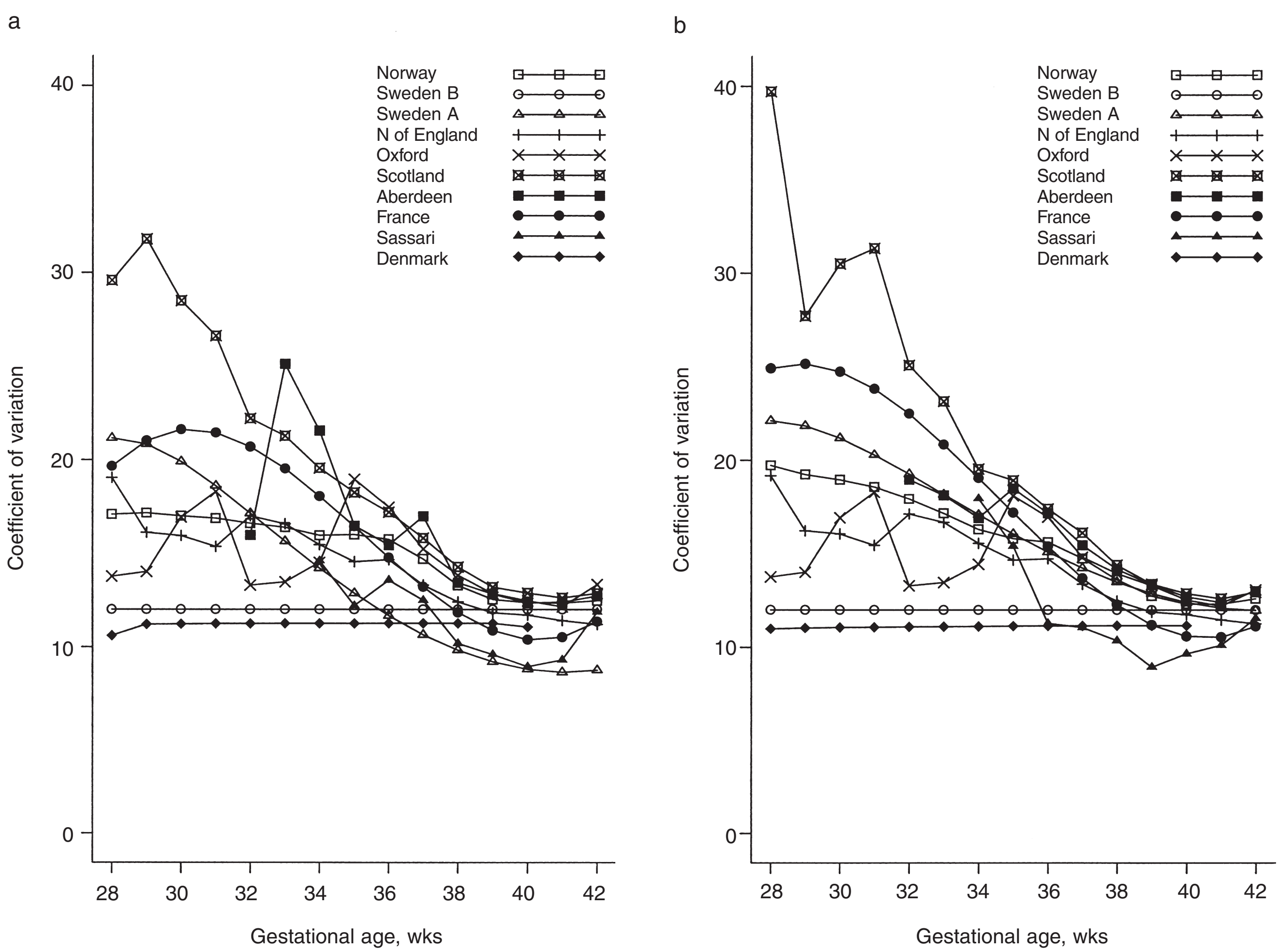

Figure 3: *Coefficients of variation (SD divided by mean and multiplied by 100) in (a) male and (b) female standards.*

All except three of the standards provided medians. For Norway and Sweden A, the mean was used in the comparative analysis; for Sweden B the mode was used: at term, mean, median, and mode are essentially equal. Most standards included gestations in the range 28 to 42 weeks.

COMPARISON OF STANDARDS

Figures 1 and 2 show the median birthweight and the 10th and 90th centiles by gestational age. The extent to which growth slows near term varied between the standards. Sweden B, for example, showed a continuation in growth after 40 weeks, whereas the Sassari standard showed a marked downwards turn. The range covered by the 10th and 90th centile bands varied considerably between the standards at low gestational ages. For example, the difference between the 90th and 10th centile bands at low gestational ages for the Scotland and Sweden A standards were large (600–900g), whereas the North of England and Sweden B standards differed by only 300 to 400g.

A comparison of male and female standards (Table II) confirmed that males were consistently heavier than females: at term, males had a median birthweight on average 144g (2%) heavier than that of females, and 137g (4%) heavier at the 10th centile. Although absolute differences were larger nearer term, differences as a percentage of average birthweight for the gestational age were steady across all gestational ages. For the two standards that stratify by parity (Scotland and Aberdeen), average, median, and 10th centile weights for multiparous infants were greater than those for primiparous infants (Table III), as expected. The multiparous to primiparous differences were larger between the 10th centiles and were most marked at early gestational ages for males. At 28 weeks

Table IV: Differences at term in median birthweight (g) for males (row minus column)

	Sweden A	*Sweden B*	*NoE*	*Oxford*	*Scotland*	*Aberdeen*	*France*	*Sassari*	*Denmark*	*Mean*
Norway	87	10	187	255	195	259	256	325	255	183
Sweden A	0	–77	100	168	108	172	169	238	168	96
Sweden B	_	0	177	245	185	248	246	315	245	173
NoE	_	_	0	68	8	72	69	138	68	–4
Oxford	_	_	_	0	–60	4	1	70	0	–72
Scotland	_	_	_	_	0	64	61	130	60	–12
Aberdeen	_	_	_	_	_	0	–3	67	–4	–76
France	_	_	_	_	_	_	0	69	–1	–73
Sassari	_	_	_	_	_	_	_	0	–70	–142
Denmark	_	_	_	_	_	_	_	_	0	–72

Values are positive if standard tabulated in row is greater than that in column; thus median birthweight for Norway is 87g greater than median for Sweden A. Average differences for each standard in comparison with all other standards are also presented. NoE, North of England.

Table V: Differences at term in median birthweight (g) for females (row minus column)

	Sweden A	*Sweden B*	*NoE*	*Oxford*	*Scotland*	*Aberdeen*	*France*	*Sassari*	*Denmark*	*Mean*
Norway	64	49	175	245	195	245	255	295	245	177
Sweden A	0	–15	112	181	131	181	191	231	181	113
Sweden B	–	0	127	196	146	196	206	246	196	128
NoE	–	–	0	70	20	70	80	120	70	1
Oxford	–	–	–	0	–50	0	10	50	0	–68
Scotland	–	–	–	–	0	50	60	100	50	–18
Aberdeen	–	–	–	–	–	0	10	50	0	–68
France	–	–	–	–	–	–	0	40	–10	–78
Sassari	–	–	–	–	–	–	–	0	–50	–118
Denmark	–	–	–	–	–	–	–	–	0	–68

Values are positive if standard tabulated in row is greater than that in column; thus median birthweight for Norway is 64g greater than median for Sweden A. Average differences for each standard in comparison with all other standards are also presented. NoE, North of England.

Table VI: Misclassification of term females as small for gestational age (SGA) at 10th centile

Standard used for classification	*Approximate percentage (assuming normality) classified as SGA in standard populations*			
	Norway	*Sweden A*	*North of England*	*Scotland*
Norway	10.2	12.7	16.7	20.0
Sweden A	9.0	11.3	14.9	18.2
North of England	6.0	7.7	10.1	13.0
Scotland	5.2	6.7	8.8	11.6

the 10th centiles for males differed by 230g (24%), but only by 109g (4%) at term. This decrease in percentage difference between parity with increasing gestational age was evident for both sexes, for both median and 10th centile.

Pairwise differences between median values at term show that Norway's standard and the two Swedish standards were the heaviest for both males and females (Tables IV and V). Norway had an average birthweight at term that was 183g heavier than that in other standards for males, and 177g heavier for females. The lightest standard was Sassari: median birthweights averaged 142 and 118g lighter than the other standards for males and females respectively. These differences were of the same magnitude as observed differences between sex and parity. The standards for Oxford, Aberdeen, Denmark, and France were also light, with term birthweights about 70g lighter than the others. The North of England standard was at the overall average.

The effect of the differences in methods of construction of standards on the centiles is illustrated in Table VI. With SGA defined as being below the 10th centile, the proportion of term female infants classified as SGA varied considerably as standards from one country were applied to other countries.

The Sweden B standard assumed that the SD was 12% of the mean, which was a CoV of 12 for all gestational ages (Fig. 3); Denmark assumed a CoV of 11. However, the other standards suggest that this might not be a sensible assumption. Preterm infants have higher CoVs, many above 20 for infants of 28 to 32 weeks' gestational age, particularly females. The standard for Scotland, a growth reference with no exclusions, has early CoVs of about 30: the SD is about 30% of the mean. Even the Norway standard, which has many exclusions, shows a decline in the CoV with increasing gestational age.

Discussion

Adjusting birthweight by gestational age for factors such as sex, parity, and plurality allows for known sources of variation. The intention is to improve the focus on other associations. However, if the standards used to make adjustments differ, then associations might be hidden or reversed, rather than revealed. Within-standard sex and parity differences provide a comparison for the size of differences between countries. We confirmed from singleton standards, a small percentage excess (2–3%) of male median and 10th centile birthweights relative to females. Higher parity was associated with larger birthweight, as expected, with the difference being more marked and consistent for male infants. Other factors, such as smoking or maternal height, also affect birthweight,[3] but standards adjusting for smoking or maternal height are not readily available.

The observed differences in standards between and within countries are not necessarily due to any real differences in birthweight distributions. Differences in exclusions, methods of dating pregnancy, or delivery and weighing of infants might explain some or most of the apparent differences. Sweden A excluded 21% of births. If the lowest 20% of term birthweights were excluded from the Scottish singleton males, the mean would increase by 127g, bringing it much closer to Sweden A.

Norway, which has the heaviest standard, excluded caesarean births (potentially lighter births[16]), consisted of the most recent cohort of births, and did not use ultrasound dating. Two of the light standards, Oxford and France, were hospital based, which might introduce biases. For example, these centres might have a larger proportion of preterm deliveries with complex problems because they might be centres of excellence and receive referrals from the surrounding area. Scotland excluded all stillbirths, whereas the North of England, for example, excluded severe malformations and uncertain gestational ages in addition to antepartum stillbirths. Because infants with a severe malformation are more likely to be SGA, it is not surprising that Scotland has a lower mean than the North of England.

Statistical methods used within the construction of the standard can also have far-reaching implications. The distribution of birthweights at short gestation is often observed to be skewed,[4] and differences from the Gaussian distribution have been found even at term.[12] Excluding certain birth types from the standard, in an attempt to extract only 'healthy' infants, removes some of this non-normality.[18] However, the data may need to be transformed before normality is attained,[11] or such assumptions may be avoided altogether by providing empirical standards only.[17] The standards compared here each took a different view on these issues. Sweden A, for example, first transformed the data before assuming normality. Sweden B assumed that there existed a constant CoV across all gestational ages, i.e. that the SD was 12% of the mean. Data from other standards show that this assumption is not justified. The assumption by Tin et al.[9] of a constant percentage difference with gestational age for sex and parity is not confirmed.

To illustrate the research implications of the assumption of a constant CoV of 12, consider a case–control study that evaluated the association between intrauterine growth and risk of childhood-onset diabetes.[3] The gestational ages of the infants ranged from 32 to 45 weeks, and the Sweden B standard[12] was used to define deviation in birthweight by expressing the difference from the expected birthweight for gestational age in SDs. For gestational ages 32 to 34 weeks, the SD is more likely to be at least 18% of the mean (Fig. 3). Hence the deviations will be overestimated by 50%, e.g. the value '3SD' should be '2SD'. This leads to an increase in the proportion of infants defined as SGA. If the Sweden A standard[11] had been used, which is reasonable because the diabetes study considered births from 1973, the conclusions would probably have been stronger.[3]

There are also clinical implications of the assumption that the SD is always 12% of the mean. As the usual variation is greater than this, too many fetuses and preterm infants might be classified as growth restricted or heavy. This could lead to inappropriate management choices, or incorrect conclusions about the rates of SGA.

Appropriate classification of SGA depends on accurate and relevant tables. With the definition of SGA as centile, differences between standards could lead to only 5%, or as many as 20%, of infants being classified as SGA. Our emphasis has been on differences at term, because most infants are born at term. Obstetricians might be less interested in SGA as a clinical measure because they can follow intrauterine growth restriction on serial ultrasound during pregnancy. Neonatologists and epidemiologists need to consider misclassification of SGA (because it could affect postnatal care as well as epidemiological studies), including those evaluating antenatal care. The choice of standards can affect estimates of magnitude of risk, and the power of studies. There has been less focus on heavy infants, but similar considerations apply.

Differing exclusion criteria and methods of construction prevent conclusions from being drawn as to whether infants differ in median birthweight across Europe,[19] but do not prevent us from concluding that standards themselves do differ. The inclusion of further standards, comparisons of preterm births, and comparisons of 10th centiles, could only increase the variation seen. Average differences of about 100g between standards may not seem to be clinically significant, but they are the same magnitude as differences by sex and parity. Maximum observed differences in excess of 300g are important.

The use of a single country-based standard within an international collaborative study will be subject to some misclassification due to local population differences (such as maternal size and differing rates of smoking), whereas the use of local-based standards, although controlling for local population features will add other misclassifications of their own, due to differing methods of construction, exclusion criteria, and secular trends. Study designers should be aware of the limitations and investigate the robustness of conclusions by comparing results obtained using several different country-based standards.

Conclusion

Any international study that uses growth standards should decide what level of adjustment is relevant to the study aims. If country-based standards are used, they should be as similar as possible in terms of methods of construction, time period, and births excluded, so as to minimize misclassification due to methods and selection criteria.

We make the following recommendations. Use country-based standards stratified by sex and parity for singletons. Consider what biases might arise from the exclusions and methods of construction. If data on sex are not available, adjust the standard at each gestational age by following the estimated differences between standards by sex, as reported in Table II. If data on parity are not available, similarly adjust the standard at each gestational age by following the estimated differences between standards by parity, as reported in Table III. If no country-based standard is available, use a standard for a similar country.

DOI: 10.1017/S0012162206001988

Accepted for publication 24th April 2006.

Acknowledgements
JLH, SNJ, and MJP are grateful for the encouragement of SCPE. SCPE was supported by European Commission funds: DGXII-BIOMED2-Contract NBMH4-983701.The authors also thank Edmund Hey for reading and commenting on draft versions of this paper. KH was supported by a Medical Research Council grant awarded to JLH.

References

1. Keirse MJNC. (2000) International variations in intrauterine growth. *Eur J Obstet Gynecol Reprod Biol* **92:** 21–28.
2. Jarvis SN, Glinianaia SV, Torrioli MG, Platt MJ, Miceli M, Jouk PS, Johnson A, Hutton J, Hemming K, Hagberg G, et al. (2003) Cerebral palsy and intrauterine growth in single births: a European collaborative study. *Lancet* **362:** 1106–1111.
3. Dahlquist G, Bennich SS, Kallen B. (1996) Intrauterine growth pattern and risk of childhood onset insulin dependent (type I) diabetes: population based case-control study. *Br Med J* **313:** 1174–1177.
4. Wilcox A, Skjaerven R. (1992) Birthweight and perinatal mortality: the effect of gestational age. *Am J Publ Health* **82:** 378–382.
5. Wilcox AJ. (2001) On the importance – and the unimportance – of birthweight. *Int J Epidemiol* **30:** 1233–1241.
6. Forbes JF, Smalls MJ. (1983) A comparative analysis of birthweight for gestational age standards. *Br J Obstet Gynaecol* **99:** 297–303.
7. Graafmans WC, Richardus JH, MacFarlane A, Rebagliato M, Blondel B, Verloove-Vanhorick SP, Mackenbach JP. (2001) Comparability of published perinatal mortality rates in Western Europe: the quantitative impact of differences in gestational age and birthweight criteria. *Br J Obstet Gynaecol* **109:** 1237–1245.
8. Surveillance of Cerebral Palsy in Europe (SCPE) Collaborative Group. (2000) Surveillance of Cerebral Palsy in Europe: a collaboration of cerebral palsy surveys and registers. *Dev Med Child Neurol* **42:** 816–824.
9. Tin W, Wariyar UK, Hey EN. (1997) Selection biases invalidate current low birthweight for gestation standards. *Br J Obstet Gynaecol* **104:** 180–185.
10. Campbell D, Hall M, Lemon J, Carr-Hill R, Pritchard C, Samphier M. (1993) Clinical birthweight standards for a total population in the 1980s. *Br J Obstet Gynaecol* **100:** 436–445.
11. Niklasson A, Ericson A, Fryer JG, Karlberg J, Lawrence C, Karlberg P. (1991) An update on the Swedish reference standards for weight, length and head circumference at birth for given gestational age (1977–1981). *Acta Paediatr Scand* **80:** 756–762.
12. Kallen B. (1995) A birthweight for gestational age standard based on data in the Swedish Medical Birth Registry, 1985–1989. *Eur J Epidemiol* **11:** 601–606.
13. Greisen G, Michaelsen KF. (1989) Perinatal vaekst. *Ugeskr Laeger* **151:** 1813–1815. (In Danish)
14. Mamelle N, Munoz F, Grandjean H. (1996) Fetal growth from the AUDIPOG study. I. Establishment of reference curves. *J Gynecol Obstet Biol Reprod (Paris)* **25:** 61–70.
15. Spano B, Grimaldi MA, Serra GC, Sanna MCA, Sardu MF, Manca MA, Madeddu R, Orzalesi M. (1988) Curve di accrescimento endouterino da 34 a 42 settimane di gravidanza nella provincia di Sassari. *Minerva Pediatr* **40:** 315–319. (In Italian)
16. Skjaerven R, Gjessing HK, Bakketeig LS. (2000) Birthweight by gestational age in Norway. *Acta Obstet Gynecol Scand* **79:** 440–449.
17. Scottish Health Services Common Services Agency. (1990) *Scotland: Birthweight, Head Circumference and Length for Gestational Age.* Edinburgh: Scottish Health Services Common Services Agency.
18. Yudkin PL, Aboualfa M, Eyre JA, Redman CWG, Wilkinson AR. (1987) New birthweight and head circumference centiles for gestational ages 24 to 42 weeks. *Early Hum Dev* **15:** 45–52.
19. Graafmans WC, Richardus JH, Borsboom GJ, Bakketeig L, Langhoff-Roos J, Bergsjo P, Macfarlaine A, Verloove-Vanhorick SP, Mackenbach JP. (2002) Birthweight and perinatal mortality: comparison of optimal birthweight in seven Western European countries. *Epidemiology* **5:** 569–574.