European Journal of Obstetrics & Gynecology and Reproductive Biology 206 (2016) 184–193

Contents lists available at ScienceDirect

European Journal of Obstetrics & Gynecology and Reproductive Biology

journal homepage: www.elsevier.com/locate/ejogrb

Development of a birthweight standard and comparison with currently used standards. What is a 10th centile?

Ricardo F. Sousa-Santos[a,b,*], Rui F. Miguelote[a,c], Ricardo J. Cruz-Correia[b,d], Cristina C. Santos[b,d], João F.M.A.L. Bernardes[b,d]

[a] Gynecology and Obstetrics Department, Hospital Senhora da Oliveira, Rua dos Cutileiros, Creixomil, 4835-044 Guimarães, Portugal
[b] Center for Research in Health Technologies and Information Systems (CINTESIS), Al. Prof. Hernâni Monteiro, 4200-319 Porto, Portugal
[c] Life and Health Sciences Research Institute (ICVS), Universidade do Minho, Campus de Gualtar, 4710-057 Braga, Portugal
[d] Faculty of Medicine, University of Porto, Al. Prof. Hernâni Monteiro, 4200-319 Porto, Portugal

ARTICLE INFO

Article history:
Received 30 November 2015
Received in revised form 18 September 2016
Accepted 22 September 2016

Keywords:
Birthweight
Fetal development
Prenatal
Growth charts

ABSTRACT

Introduction: Fetal growth charts are often used in clinical practice. It is important to understand the usefulness and the pitfalls associated with these tools. Without validation, it is difficult to ascertain if the cutoffs we intend are the ones we actually select. We developed a national standard for birthweight (BW) and compared it with other published reference values.
Study design: Multicenter retrospective study. We collected data on live births, including first trimester ultrasound and pathology, from 23 to 42 weeks' gestational age (GA). We used a variation of the lambda (λ), mu (μ), and sigma (σ) method (LMS) to construct and smooth predicted centiles. GA data was plotted and modeled in days from 24 to 42 weeks. Resulting centiles were validated and compared with other published and widely used reference values. Data from both BW and estimated fetal weight was used to validate the model.
Results: Data on 661,338 births were collected from 22 institutions, including 71,515 cases with first trimester ultrasound. We excluded preterm cesarean section from analysis, because of a significant bias (up to 18%) on BW and used exclusively first trimester ultrasound dates from 34 to 42 weeks. The standard compares favorably with tables currently in use, both ultrasound and birthweight based.
Conclusion: The use of first trimester ultrasound limits variability by minimizing some random error sources, such as data introduction and GA errors, while allowing better precision (GA in days). This results in a narrower range in the extreme centiles than other charts. Validation with estimates of fetal weight are sound in second and early third trimester fetuses, because that will be a "real world" usage of this standard. While there are similarities between our series and some international/foreign growth charts, other are unfit to characterize our population. This reinforces the need for validation of standards, and sound methodological practices when doing so.

Introduction

Obstetricians and neonatologists worldwide are familiar with the concept of fetal growth charts, arguably introduced by Lubchenco et al. in 1963 [1], in which birthweight is plotted as a function of gestational age (GA). As growth centiles correlate with fetal and neonatal outcomes, these are often used to produce clinical judgements. While several limitations and pitfalls are apparent, these are readily available tools, easy to understand and widely adopted [2,3]. Given the large number of such charts published over the years, it is up to the clinician to choose one to which compare its population of fetuses/newborns. Such choice and its implications are not always well understood [4] and even professional societies sometimes do not agree on recommendations based on available evidence [5].

Whatever the tool, validation is of paramount importance, to ensure that, when the clinician selects the 10th centile, he is indeed referring to 10% of his intended population (e.g., a healthy local woman). We aim to describe how a population standard relates to others, and if differences are clinically significant, given the regional, ethnical and geographical differences previously identified [6–8]. We describe a birthweight (BW) standard for Portugal,

* Corresponding author at: CIDES, Faculdade de Medicina da Universidade do Porto, Al. Prof. Hernâni Monteiro, 4200-319 Porto, Portugal.
E-mail address: ricardosantos@fastmail.com (R.F. Sousa-Santos).

http://dx.doi.org/10.1016/j.ejogrb.2016.09.028
0301-2115/

from 24 to 42 weeks and compare it with other commonly used charts.

Methods

Participants

Singleton neonates from gestational age (GA) 168–300 days (23–42 completed weeks), live born between January 2004 and June 2014. Mothers with medical pathology were excluded from analysis (hypertensive diseases and diabetes, gestational or otherwise; autoimmune diseases; epilepsy; chronic medications). Newborn malformations, including chromosomal abnormalities, and hemolytic diseases were also excluded.

Study design

Cross sectional study. We identified and sought institutional authorizations for data retrieval regarding births in state financed maternity hospitals with over 1500 births per year.

Data collection

Computerized records of births and first trimester ultrasound were collected. Standardized record search strategies were developed in the main Portuguese Electronic Patient Record (SONHO, currently maintained by state owned Serviços Partilhados do Ministério da Saúde, Portugal) and the fetal ultrasound software Astraia (Astraia software GmbH, Munich, Germany). Databases were joined by using unique anonymized identifiers: institutional episode, patient record for ultrasound data. Additional validations were: dates of ultrasound and birth, calculated (ultrasound) and recorded GA. Ultrasound based estimates of fetal weight (EFW) were available in 17 institutions and used for validation purposes. A validation sample of BW from June 2014 to August 2016 was additionally collected from one of the participant hospitals.

Variables description

We collected institution name, mother's age, parity, medical and pregnancy related conditions, last menstrual period (LMP), date and Crown Rump Length (CRL) of first trimester ultrasound, EFW(s), newborn diagnosis, birth date, newborn sex, recorded gestational age, birthweight (BW), and type of delivery. Birth and ultrasound queries included, respectively, episode and patient diagnosis. We calculated GA with Hadlock's CRL formula [9]. Validation data for EFW in low GA fetuses was based in Hadlock's 4 parameter formula [10].

Statistical analysis

GA was calculated in days by ultrasound (CRL up to 85 mm). Given the scarcity of premature births, we decided to use the full dataset, regardless of ultrasound dating, in newborns with GA <34 weeks, while exclusively using complete (with CRL) data from 34 weeks (238 days) onward.

An obstetrician (RS) scanned the birthweight dataset for probable data errors, with the aid of the Tukey method for outlier identification [11]. We explored the previously reported associations of cesarean delivery with observed birthweight [12], as well

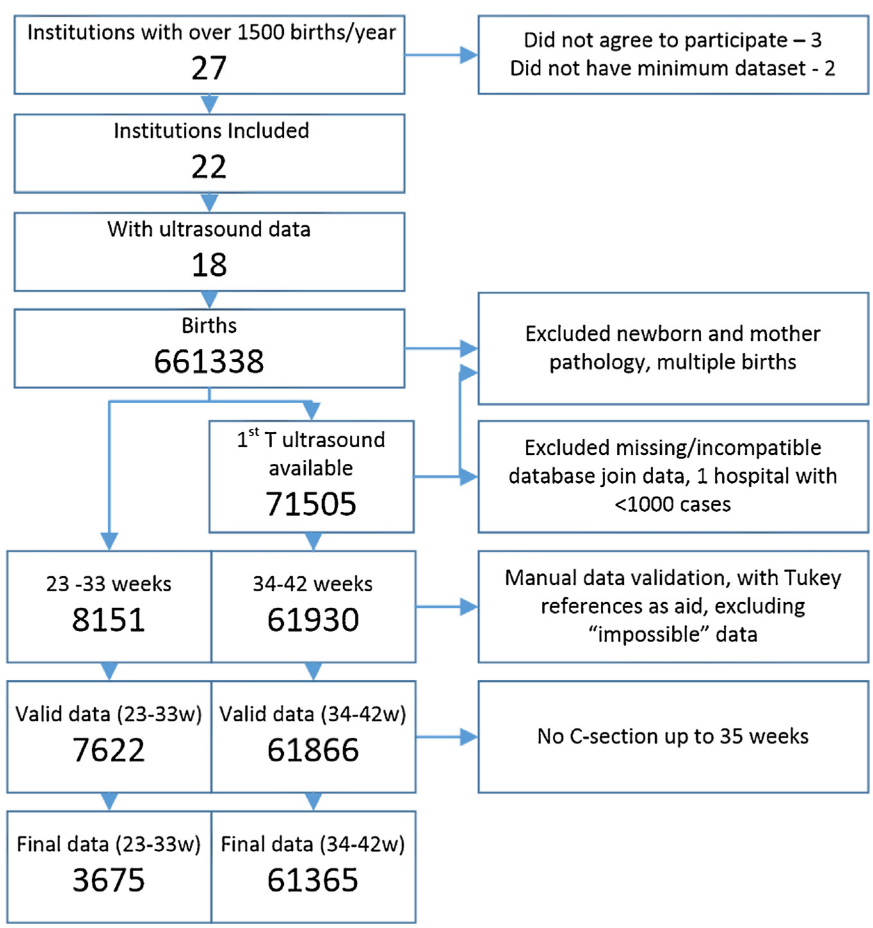

Fig. 1. Flowchart for data collection/validation.

as newborn sex weight differences. Comparisons were performed with *Student's t-test* in SPSS v23 (IBM Corp., Armonk, NY). We excluded Cesarean section births under 36 weeks (positively skewed), and calculated centiles separately for boys and girls. Given the lack of significant birthweight differences by sex below 29 weeks, both sexes were used to fit the model, augmenting the number of individuals available for analysis and enhancing its robustness in these GA's. We considered data corresponding to complete weeks a "middle of the week" estimate while data in days was analyzed in each gestational day.

To model growth, we compared the lambda (λ), mu (μ), and sigma (σ), or LMS method by Cole and Green [13] and two extended methods, LMSP and LMST, proposed by Rigby and Stasinopoulos [14], which adjust for kurtosis (τ). A power transformation for x (age) was allowed, and local maximum likelihood was used for smoothing parameters, with GAMLSS [15] package in *R* 3.2.1 software (R Foundation for Statistical Computing, Vienna, Austria). LMSP provided the best fit, evaluated by visual inspection, worm plots and Q-Statistics. Predicted centiles were calculated for each completed day from 168 to 300 days (24–42 weeks). To compare with other authors and create reference tables, we used values from middle of the week, as a proxy to "completed weeks" data.

The birthweight validation samples were used without manipulation. The birth data was plotted raw, with no exclusions of outliers, diseases or type of birth. For EFW (ultrasound) data, we excluded from the validation sample all fetuses in the model fit and randomly selected one observation per fetus if multiple were available. "N − 1" chi-squared test was used to compare proportions.

Results

Twenty-seven maternity hospitals had more than 1500 births per year (roughly 70% of maternities and 90% of births in the country), of which 24 authorized data collection. Computerized records were retrieved from 22 of these institutions. In the two remaining institutions, it was not possible to retrieve the required minimum set of data (Fig. 1).

Data was not available throughout the whole time frame in some institutions. Ultrasound data was available in 18 institutions. A total of 661,338 live births were identified. First trimester ultrasound was available in 71,505 cases.

8151 newborns between 23 and 33 weeks were included from the main database, while data from ultrasound dated pregnancies accounted for 61,930 cases between 34 and 42 weeks (238–300 days). A total of 593 cases were further excluded because of missing or wrong/impossible data, 79% of which from the main database (up to 33 weeks). Table 1 summarizes the final dataset.

Newborns delivered by C-section were consistently smaller up to 35 weeks (up to 18% of mean BW for GA), and heavier from 37 weeks onward (1,4–2,6% of mean BW for GA). Differences in BW by sex were small and not consistent up to 29 weeks, but girls weighted less in all GA's (up to 5% of mean BW).

The sample contains more boys (50,9%), in line with national statistics. There were consistently more premature boys, regardless of delivery mode ($p < 0{,}001$).

The predicted centiles were fitted with LMSP, using a Box-Cox Power Exponential distribution. After visual analysis, Q–Q statistics and worm plots comparison between models, this was agreed to be the best overall fit, followed by LMST. The model's global deviance was 977,196 for both sexes, 469,295 for girls and 496,793 for boys. Fig. 2a and b depicts some of these statistics for the full dataset.

The validation dataset comprised 5267 newborns, of which 5172 from 34 to 41 weeks and 64,919 estimates of fetal weight from 24 to 40 weeks, each from a unique fetus, not used from the main dataset (model creation). Fig. 2c and d shows the plotted centiles (unfiltered, uncorrected data, with no GA correction). Predicted and validation raw centiles have a good fit throughout GA. The large validation sample of EFW has noticeably higher 3rd and 10th centiles between 28 and 32 weeks. Table 4 compares the raw 10th and 90th centiles of validation data with cutoffs provided by selected charts. The number of fetuses under or above the represented centiles are shown, as well as sensitivity and

Table 1
Dataset summary.

GA	GA precision	Total analyzed				Sex differences				Cesarean section differences		
		Initial	After data validation	Cesareans eliminated	Included	% girls	Mean diff. (g)	p value	Analysis	Mean diff. (g)	p value	Analysis
23	Weeks	91	83	8	75	44	−29	0,11	Aggregated	30	0,305	Excluded
24	Weeks	213	201	58	143	44	−17	0,316	Aggregated	−7	0,795	Excluded
25	Weeks	267	252	103	149	48	−54	0,011	Aggregated	−76	<0,001	Excluded
26	Weeks	323	302	162	140	44	−32	0,191	Aggregated	−110	<0,001	Excluded
27	Weeks	398	383	220	163	40	−62	0,016	Aggregated	−143	<0,001	Excluded
28	Weeks	542	486	306	180	42	−45	0,152	Aggregated	−209	<0,001	Excluded
29	Weeks	547	500	301	199	42	−81	0,007	Separated	−203	<0,001	Excluded
30	Weeks	892	769	468	301	45	−58	0,041	Separated	−233	<0,001	Excluded
31	Weeks	967	884	490	394	42	−53	0,036	Separated	−241	<0,001	Excluded
32	Weeks	1653	1559	847	712	41	−104	<0,001	Separated	−239	<0,001	Excluded
33	Weeks	2258	2203	984	1219	45	−43	0,03	Separated	−214	<0,001	Excluded
34	Days	478	472	209	263	43	−124	0,003	Separated	−193	<0,001	Excluded
35	Days	893	879	292	587	41	−91	0,003	Separated	−163	<0,001	Excluded
36	Days	2066	2053	–	2053	47	−134	<0,001	Separated	−8	0,712	Included
37	Days	5356	5348	–	5348	47	−104	<0,001	Separated	42	0,001	Included
38	Days	12,083	12,078	–	12,078	48	−119	<0,001	Separated	47	<0,001	Included
39	Days	18,624	18,616	–	18,616	49	−139	<0,001	Separated	67	<0,001	Included
40	Days	15,418	15,413	–	15,413	50	−138	<0,001	Separated	77	<0,001	Included
41	Days	6368	6365	–	6365	50	−124	<0,001	Separated	89	<0,001	Included
42	Days	644	642	–	642	48	−127	<0,001	Separated	77	0,018	Included
Total	Days	71,879	71,253	–	65,229							

Aggregated: both sexes were included in the analysis of the gender specific tables; *separated*: sexes were analyzed separately in these gestational ages; excluded: deliveries by cesarean section were excluded from the centile analysis; included: deliveries by cesarean section were included in the centile analysis.

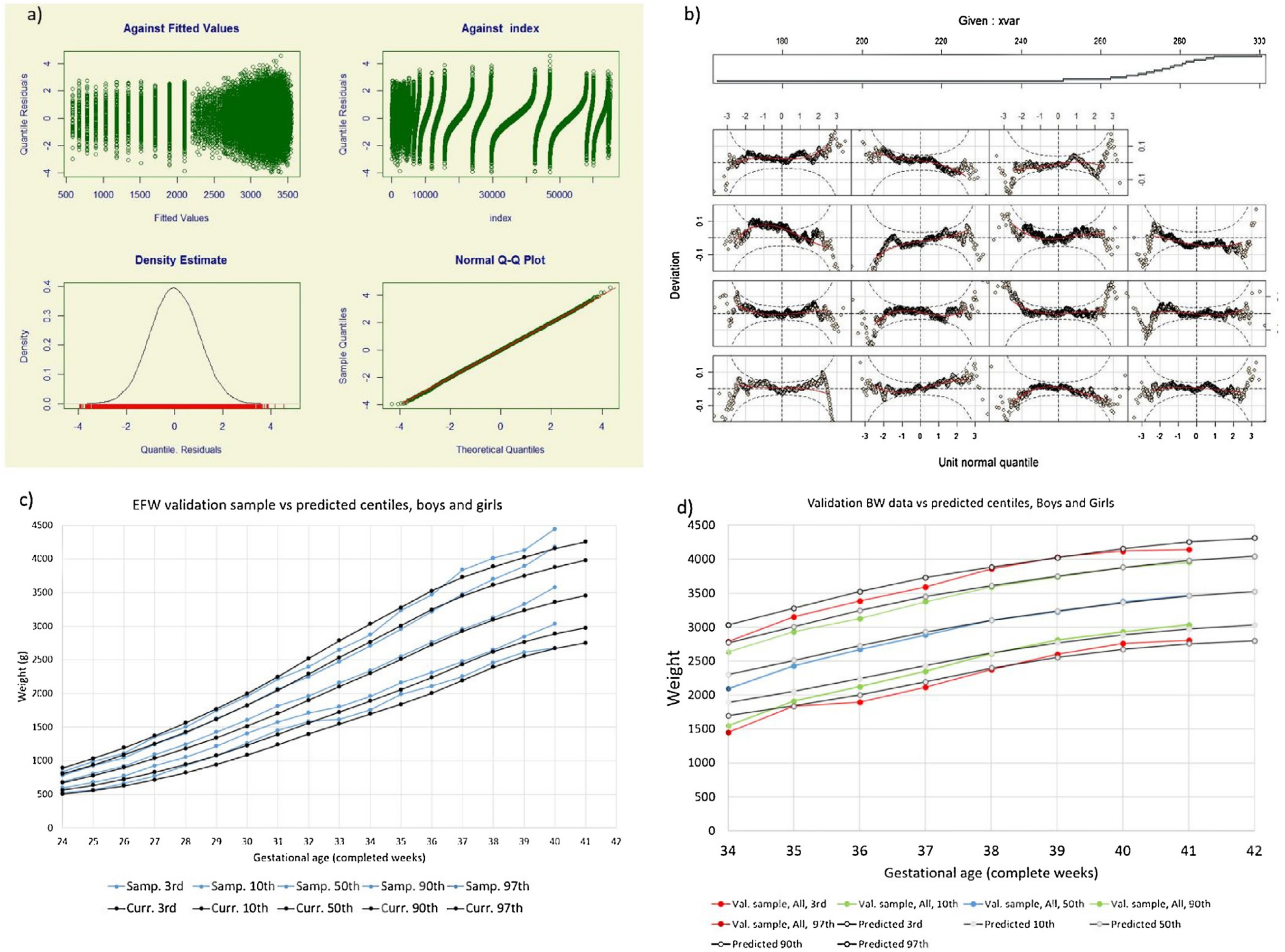

Fig. 2. (a) Model fit, boys and girls; (b) wormplots for the model fit, boys and girls; (c) validation sample from estimated fetal weight (raw data) vs predicted centiles; (d) validation sample from birthweight (raw data) vs predicted centiles.

specificity at each cutoff. The current series has the lowest number of statistically different proportions from the validation samples, and the overall highest sensitivity, while maintain high specificity.

Predicted (smoothed) centiles, in completed weeks are presented in Tables 2 and 3. The mean and standard deviation (SD) were calculated from the samples. Figs. 3 and 4 depict fitted

Table 2
Data summary and predicted Centiles for each complete week, both sexes.

Boys and girls												
GA	N	Mean	SD	3rd	5th	10th	25th	50th	75th	90th	95th	97th
24	143	678	99	502	525	560	617	675	737	806	855	889
25	149	785	130	558	588	634	706	779	855	938	994	1033
26	140	895	144	628	666	723	812	900	991	1086	1150	1193
27	163	1045	161	716	761	828	932	1036	1141	1250	1322	1371
28	180	1194	205	820	870	945	1064	1182	1303	1426	1508	1563
29	199	1333	211	942	996	1078	1207	1339	1475	1614	1706	1768
30	301	1521	245	1085	1141	1228	1368	1513	1665	1821	1925	1996
31	394	1717	248	1239	1299	1391	1542	1703	1873	2050	2167	2248
32	712	1918	294	1396	1460	1558	1722	1901	2092	2292	2426	2519
33	1219	2128	334	1549	1617	1724	1903	2100	2312	2535	2684	2787
34	263	2321	336	1694	1770	1887	2084	2300	2530	2768	2925	3033
35	587	2523	368	1838	1924	2055	2273	2509	2756	3006	3167	3277
36	2053	2752	406	2003	2095	2237	2472	2725	2987	3246	3412	3523
37	5348	2947	410	2196	2289	2433	2671	2927	3193	3453	3619	3730
38	12,078	3115	394	2391	2480	2618	2846	3097	3357	3613	3777	3886
39	18,616	3252	387	2552	2637	2768	2990	3236	3495	3750	3914	4024
40	15,413	3368	396	2670	2755	2886	3108	3357	3619	3878	4044	4155
41	6365	3457	394	2752	2839	2974	3201	3455	3722	3981	4145	4255
42	642	3497	398	2803	2893	3032	3264	3523	3789	4044	4203	4309

Table 3
Data summary and predicted centiles for each complete week, separated by gender.

GA	N	Mean	SD	3rd	5th	10th	25th	50th	75th	90th	95th	97th
Boys												
24	80[a]	678	99	504	526	560	616	675	737	806	854	887
25	78[a]	785	130	556	586	631	704	778	856	939	995	1034
26	78[a]	895	144	626	664	720	809	899	991	1086	1148	1191
27	97[a]	1045	161	717	761	828	932	1038	1144	1252	1322	1369
28	105[a]	1194	205	824	875	952	1072	1193	1315	1438	1517	1570
29	116	1367	220	959	1014	1096	1228	1362	1498	1635	1723	1783
30	166	1547	228	1116	1173	1259	1399	1544	1693	1844	1942	2008
31	228	1740	248	1281	1340	1431	1579	1737	1901	2069	2179	2253
32	419	1961	280	1442	1504	1601	1761	1935	2118	2307	2431	2516
33	666	2147	318	1593	1660	1765	1941	2133	2337	2547	2685	2780
34	149	2375	354	1736	1813	1931	2127	2341	2565	2791	2939	3039
35	347	2560	359	1880	1968	2102	2323	2561	2806	3049	3205	3309
36	1084	2815	411	2037	2134	2282	2523	2782	3046	3305	3468	3578
37	2841	2996	415	2239	2335	2481	2722	2982	3248	3507	3671	3781
38	6254	3172	390	2452	2541	2678	2907	3157	3416	3671	3833	3943
39	9508	3320	389	2616	2702	2834	3057	3305	3564	3821	3985	4096
40	7648	3438	404	2725	2811	2946	3173	3427	3692	3954	4122	4234
41	3190	3532	398	2813	2902	3041	3273	3530	3796	4055	4219	4328
42	333	3557	393	2885	2977	3117	3351	3606	3867	4115	4269	4372
Girls												
24	63[a]	678	99	493	519	557	617	675	735	803	850	883
25	71[a]	785	130	557	588	635	707	779	854	937	995	1036
26	62[a]	895	144	637	673	728	814	900	990	1090	1159	1207
27	66[a]	1045	161	730	771	834	934	1035	1141	1257	1336	1392
28	75[a]	1194	205	829	875	946	1059	1174	1296	1427	1517	1580
29	83	1286	189	932	983	1061	1187	1317	1453	1600	1699	1769
30	135	1489	260	1050	1106	1192	1331	1477	1629	1793	1902	1979
31	166	1686	245	1183	1245	1340	1496	1659	1831	2013	2135	2219
32	293	1856	303	1327	1396	1501	1674	1857	2051	2254	2389	2482
33	553	2104	351	1478	1553	1668	1858	2061	2275	2497	2645	2747
34	114	2250	299	1632	1712	1834	2037	2255	2486	2723	2880	2988
35	240	2470	376	1794	1877	2006	2219	2450	2693	2942	3105	3217
36	969	2681	389	1978	2064	2197	2419	2659	2912	3169	3337	3451
37	2507	2891	397	2173	2260	2394	2619	2864	3121	3380	3548	3662
38	5824	3053	389	2350	2436	2568	2789	3031	3285	3538	3700	3810
39	9108	3180	372	2501	2585	2714	2930	3167	3415	3660	3817	3923
40	7765	3300	377	2627	2710	2838	3054	3292	3541	3784	3940	4044
41	3175	3381	376	2709	2793	2922	3140	3383	3636	3882	4038	4143
42	309	3433	394	2760	2845	2976	3197	3445	3703	3953	4111	4217

[a] Summary data calculated with both sexes in these GA.

(predicted) centiles for boys and girls, for each gestational day from 24 to 42 weeks. These charts are available for download at www.fetalgrowth.med.up.pt.

Fig. 5 compares the current series with Yudkin's (boys), used in almost all obstetrical ultrasound departments in Portugal. Fig. 6 shows the 10th and 90th centiles across some widely used tables in Portugal (Hadlock's and Lubchenco's) [4], a table from a large United States sample from Alexander et al. [16] and an international standard—Intergrowth (only 33–42 weeks) [17]. Of note the wide disparity with Alexander's and Lubchenco data, as well as the proximity with data from Hadlock's and Intergrowth newborn data, and the subtle differences with Yudkin's.

Discussion

To our knowledge, the number of first trimester ultrasound dated pregnancies is one of the largest published to date in a growth chart, especially considering the use of ultrasound databases, which permitted GA in days from 34 weeks onward. This precision led to a model with more information in the temporal relation between GA and BW. The multiple sources of data may be an issue, but the few database providers, standardized queries and the easily identifiable variables, eased the process of data collection. Institutions were of, at least, moderate size (1500 or more births per year) in order to provide homogeneity in the standard of care and maximize chances of data on premature newborns (small maternities do not deliver fetuses under 34 weeks). The number of contributing hospitals also improved the sample's representativity.

The decision to exclude C-section, below 36 w from the model is explained by the systematic and clinically important bias in the BW of neonates delivered by this route (Table 1), which is better explained by elective deliveries with pathology. This agrees with previous findings [12], that C-section elimination further improves detection of appropriate BW. From 37 weeks onward this bias is inverted and small. Given the relatively high percentage of C-section in the sample and the higher chances of it being performed in large, otherwise healthy fetuses, we decided against exclusion in this group. Since all term neonates had data simultaneously extracted from two databases the chances of pathology identification were also higher.

The low numbers of premature ultrasound dated neonates led us to use the full database up to 33 weeks, because the much larger numbers would improve the model fit to a greater extent than the precise identification of GA. This subset of data had a much larger percentage of wrong/impossible data, which further validates the decision to use ultrasound dating whenever possible. Clerical errors clearly account for some of these errors, as we often see a bimodal distribution with a "hump" on BW of premature births. The premature outliers peak near 30 weeks [16,18], which may be

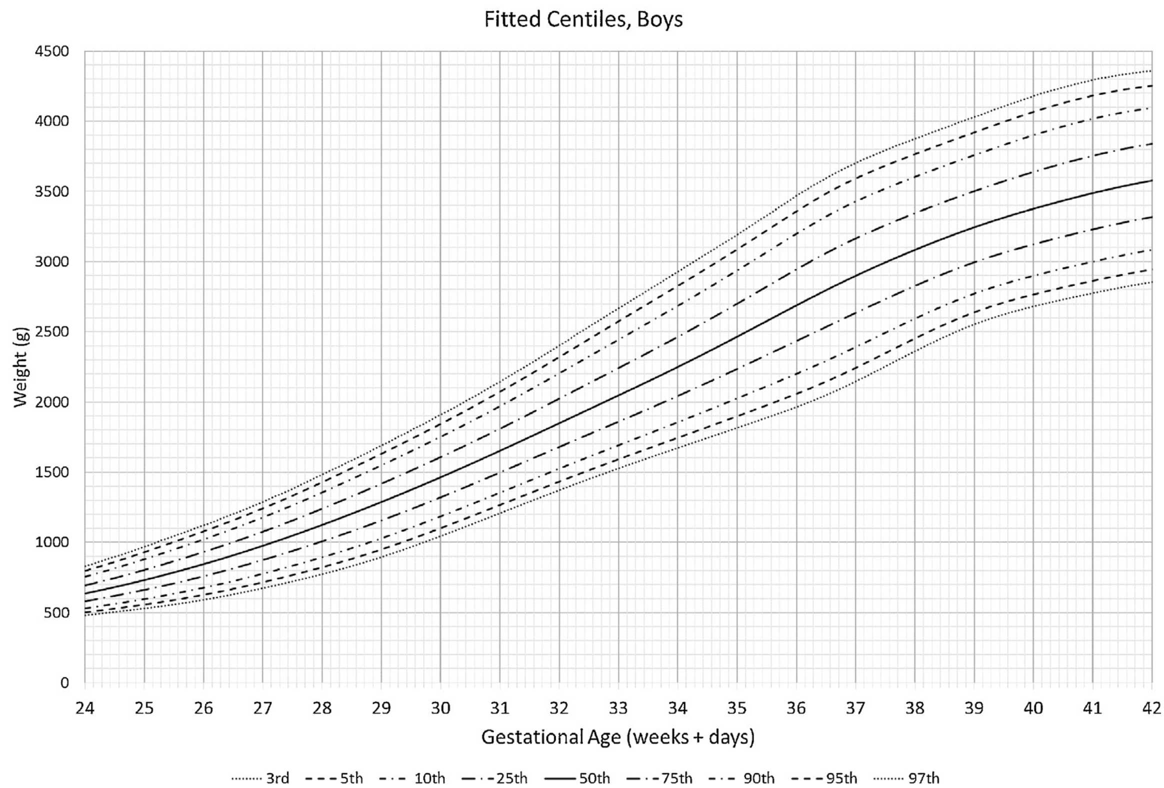

Fig. 3. Centiles for boys, for each completed day, 24–42 weeks.

Fig. 4. Centiles for girls, for each completed day, 24–42 weeks.

explained by a character typo (random interchange of 2, 3 or 4 in the first character). A small percentage of, e.g., 39 weeks' newborns erroneously classified as 29 weeks would have a large effect on centiles whereas the opposite would have a negligible effect, because of the much larger sample size at term [18]. We combined a mathematical method (Tukey's) with expert opinion for outlier

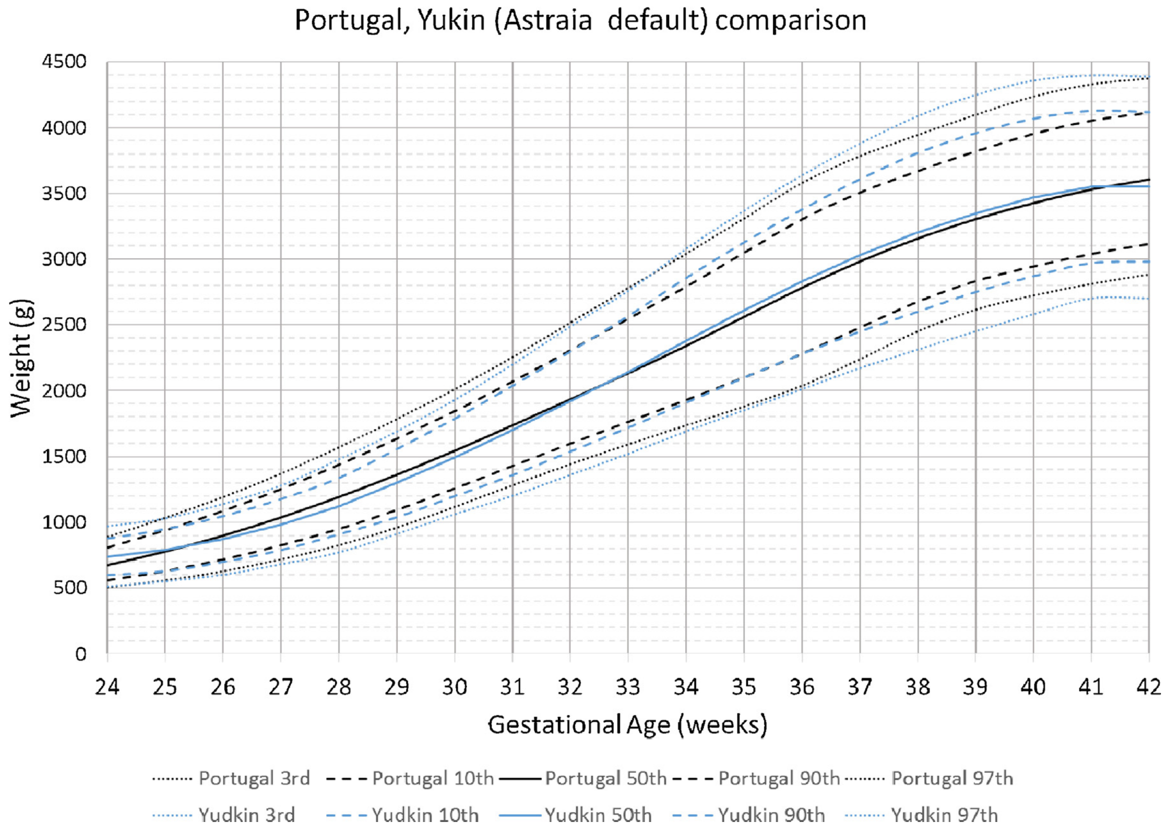

Fig. 5. Current series comparison with Yudkin's [12] (boys).

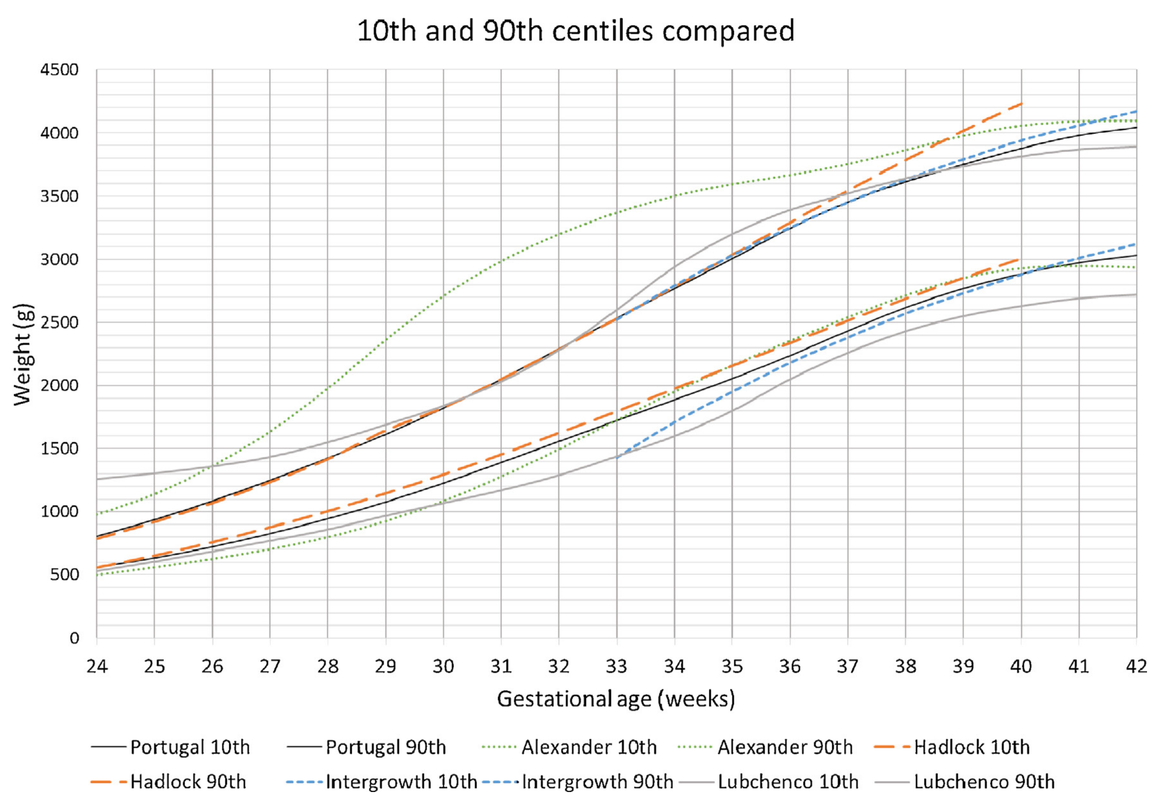

Fig. 6. 10th and 90th centile comparison with other tables (boys and girls). Portugal: current series; Alexander: Alexander et al. [16]; Hadlock: Hadlock et al. [20]; Intergrowth: INTERGROWTH-21st Project [17]; Lubchenco: Lubchenco et al. [1].

detection, which allowed for interpretation, rather than cutoff based exclusion.

Older charts, with GA based on LMP, show a steep decrease in growth velocity after 40 weeks, with diminishing mean BW after 41 weeks [1,12,16]. We agree that this may be the result of incorrect dating which would over represent in later GA's younger, incorrectly dated fetuses, and is not apparent in data derived from ultrasound dating based charts [17,19]. Even in prematurity, our series do not deviate greatly from Hadlock's [20] ultrasound fetal weight estimate based chart, which is the basis for the customized antenatal growth charts by Gardosi et al. [21], attesting the methods used in our series to minimize the birthweight bias linked to pathology.

The proportion of premature boys was significantly higher throughout GA (Table 1). At term, a much minor gap was found.

Table 4
Comparison of estimated fetal weights (EFW) and birthweight (BW) independent validation samples to 10th and 90th centiles as reported in selected tables (boys and girls).

		Val. sample	Current series	Yudkin et al. [12]	Lubchenco et al. [1]	Alexander et al. [16]	Intergrowth [17]	Hadlock et al. [20]
24 w EFW n = 1350	10th (%)	10	4,4**	9,3	3,2**	1,9**	NA	4,1**
	n SGA	135	59	126	43	26	NA	55
	Sens. (%)	100	44	93	32	19	NA	41
	Spec. (%)	100	100	100	100	100	NA	100
	90th (%)	90,1	94**	98,8**	100**	99,9**	NA	91
	n LGA	134	81	16	0	1	NA	100
	Sens. (%)	100	60	12	0	1	NA	75
	Spec. (%)	100	100	100	100	100	NA	100
28 w EFW n = 2739	10th (%)	10	1,9**	1,2**	1**	0,6**	NA	4**
	n SGA	274	52	33	27	16	NA	110
	Sens. (%)	100	19	12	10	12	NA	82
	Spec. (%)	100	100	100	100	100	NA	100
	90th (%)	90,1	91,1	78,8**	98,2**	99,9**	NA	90
	n LGA	271	244	581	49	3	NA	274
	Sens. (%)	100	90	100	18	1	NA	100
	Spec. (%)	100	100	87	100	100	NA	100
32 w EFW n = 14,997	10th (%)	10	1,7**	1,5**	0,2**	0,8**	NA	3,4**
	N SGA	1500	255	225	30	120	NA	510
	Sens. (%)	100	17	15	2	8	NA	34
	Spec. (%)	100	100	100	100	100	NA	100
	90th (%)	90,1	92,2**	92,7**	91,6**	100**	NA	91,8**
	n LGA	1485	1170	1095	1260	0	NA	1230
	Sens. (%)	100	79	74	85	0	NA	83
	Spec. (%)	100	100	100	100	100	NA	100
35 w BW n = 100	10th (%)	10	16	17	2*	22*	11*	22*
	n SGA	10	16	17	2	22	11	22
	Sens. (%)	100	100	100	20	100	100	100
	Spec. (%)	100	93	92	100	87	99	87
	90th	90	93	96	98*	100*	93	93
	n LGA	10	7	4	2	0	7	7
	Sens. (%)	100	70	40	20	0	70	70
	Spec. (%)	100	100	100	100	100	100	100
37 w BW n = 471	10th (%)	10,2	13,8	14,4*	6,8	17,8**	10,8	16,1*
	n SGA	48	65	68	32	84	51	76
	Sens. (%)	100	100	100	66,66667	100	100	100
	Spec. (%)	100	96	95	100	91	99	93
	90th (%)	90,2	93,8*	97**	95,5*	98,7**	93,8*	95,8**
	n LGA	46	29	14	21	6	29	20
	Sens. (%)	100	63,04348	30,43478	45,65217	13,04348	63,04348	43,47826
	Spec. (%)	100	100	100	100	100	100	106
39 w BW n = 1811	10th (%)	10,2	8,6	7,8*	2**	12,7*	7,3*	12,7*
	n SGA	185	156	141	36	230	132	230
	Sens. (%)	100	84	76	20	100	71	100
	Spec. (%)	100	100	100	100	97	100	97
	90th (%)	90,2	90,4	96,2**	90	96,5**	92	96,9**
	n LGA	88	86	34	90	32	72	28
	Sens. (%)	100	98	39	100	36	82	32
	Spec. (%)	100	100	100	100	100	100	100
Summary	Σ sign. diferences	NA	5/12	9/12	10/12	12/12	3/6	9/12
	EFW mean Se.		51	51	24	7	NA	69
	EFW mean Sp.		100	98	100	100	NA	100
	BW mean Se.		86	64	45	58	81	74
	BW mean Sp.		98	98	100	96	100	97

For charts without values for both sexes we considered the boys predicted values (Yudkin's from 34 w and Intergrowth's); BW—birthweight; EFW—estimated fetal weight; Se—sensitivity; Sp—specificity; SGA—small for gestational age; LGA—large for gestational age; NA—not applicable; Val.—validation.
* $p < 0{,}05$ for the difference to the proportion of the validation sample ("N – 1" chi-squared test).
** $p < 0{,}001$ for the difference to the proportion of the validation sample ("N – 1" chi-squared test).

There were relatively more girls born at 40–41 weeks, given the overall 50,9% of boys.

The large number of ultrasound dated pregnancies in our series is the result of universal first trimester screening for aneuploidy, mainly performed in National Health Service hospitals including the first of three recommend ultrasounds [22]. For this reason, we also believe that, in the subset in which we did not have access to the ultrasound report (up to 33 completed weeks), registered GA in completed weeks had this exam in account.

The BW validation data represents a use case scenario for a single hospital, fits the model without obvious bias (Fig. 2d). The very large EFW sample, which also represents everyday usage of the charts created, with ultrasound estimates, fits the model remarkably well (Fig. 2c), even comparing with an EFW based chart (Table 4). The very large number of unique observations per week and the fact that it is a multicenter sample makes for a robust validation sample. Fetuses were only included once to prevent bias, as serial measurements are not truly independent [23].

If we compare the 3rd and 97th centiles at 40 weeks (Fig. 5) for boys, a non-ultrasound dated chart widely used in the country [4] (Yudkin's [12]) ranges 1780 g between these centiles, while the current series ranges 1510 g, which more closely resembles Intergrowth's [17] range of 1580 g (also ultrasound dated). The narrower range is caused by precise dates which limit a large source of random variability [19], but also the geographically, socially and ethnically homogenous Portuguese population, compared with intergrowth's [17] sample. The smaller range enhances the classification of extreme centiles. At 39 weeks, the Yudkin table has 10th and 3rd centiles that are, respectively, 84 g and 166 g lower than the current series (Fig. 5) which causes lower sensitivity for SGA (Table 4). It also diagnoses 3,8% of 39 w newborns as LGA, compared with 9,6% for the current series. The methodology used in Yudkin et al. [12] was different, by considering every GA as normally distributed, and applying a multiple of the standard deviation to the mean to derive a centile, which was then smoothed, rather than modeling sample centiles.

Lubchenco's [1] is still one of the most used charts in Portugal by neonatologists [24]. This classical table is considered highly biased (5635 births dated by LMP, within a middle class white population, over 50 years ago, at an altitude of 1600 m). Subsequent USA data found it inadequate to characterize contemporaneous and multiple region births [16]. We also find this table unfit to classify our newborn population (Fig. 6 and Table 4). Considering a 39 weeks GA, it would have a sensitivity of 20% for SGA babies. Alexander's large USA dataset performed very poorly in the validation, and has some obvious bias, such as the mentioned "hump" on P90 in premature babies.

The present series aim to be a standard for the country, rather than a population reference. That is why we chose to eliminate both mother and child identifiable pathology from analysis. While centiles would not greatly change, we agree that this leads to comparisons more meaningful [25]. While our data was sourced from live births, the methods minimized the risk of biased BW in premature babies, and compares well with ultrasound standards in rate of growth [20], as well as with the EFW validation sample. Furthermore, EFW based approaches also have caveats, namely the errors involved in the estimation [26], the difficulty in gathering large samples, especially in late term, and the bias associated with the often used [3] method of longitudinal serial scanning of individual gestations at multiple instances [23].

Conclusion

A birthweight standard is presented and validated. We describe the differences to other charts in use in our country and elsewhere, which can be very large, and, thus, lead to misclassification of the intended cutoffs. Specifically, the Lubchenco [1] table is unfit for use, and, while Yudkin's provides a comparable 50th centile, its larger range compromises the correct identification of extreme centiles.

The validation data provided a "real world" application scenario, with a novel approach to validation in charts based in BW, by the use of a large database of EFW for low GA comparison.

Although there should always be caution when classifying a fetus or newborn to a centile, and even greater caution acting on this information, this process is widely used. This reinforces the need for appropriately validated charts, and sound methodological practices when doing so. This will ensure the plotted centiles are indeed correct for the intended population, hence providing real, unbiased data to be interpreted.

Acknowledgments

The authors would like to thank Federação Portuguesa das Sociedades de Obstetrícia e Ginecologia (Portuguese Federation of Obstetrics and Gynecology Societies) for its financial support. We also thank the collaboration of all involved maternities. The following people were especially helpful in the data retrieval process: Tiago Costa, Eliana Sousa, António Lourenço, Rui Ribeiro, Antónia Santos, Rosa Monteiro, Paulo Brás, Abel Amaro, Paulo Moura, Natália Macedo, Jorge Branco, Ricardo Fontes, Gonçalo Inocêncio, Miriam Vieira, José Mugeiro, Teresa Branquinho, Ana Filipa Sousa, Belisa Vides, Carlos Fernandes, João Miranda, Hugo Bastos, Teresa Matias, Carolina Ferreira, Nuno Pereira, Henrique Ferreira, Lucinda Vasconcelos, Francisco Valente, Frederico Sotto-mayor, Daniel Santos, Luís Branco, Lucas Ribeiro, Jorge Santos, Paulo Rodrigues, Nuno Lucas, César Costa, Luís Mendes Graça, Hugo Pinhão, Ana Carina Alves, Paula Brás.

References

[1] Lubchenco L, Hansman C, Dressler M, Boyd E. Intrauterine growth as estimated from liveborn birth-weight data at 24 to 42 weeks of gestation. Pediatrics 1963;32:793.
[2] Zhang J, Merialdi M, Platt LD, Kramer MS. Defining normal and abnormal fetal growth: promises and challenges. Am J Obstet Gynecol 2010;202:522–8.
[3] Mayer C, Joseph KS. Fetal growth: a review of terms, concepts and issues relevant to obstetrics. Ultrasound Obstet Gynecol 2013;41:136–45.
[4] Sousa-Santos RF, Mendes-Castro A, Ferreira D, Miguelote RF, Cruz-Correia RJ, Bernardes JF. Gestational age and fetal growth assessment among obstetricians. J Matern Fetal Neonatal Med 2014;7:1–6.
[5] Chauhan SP, Gupta LM, Hendrix NW, Berghella V. Intrauterine growth restriction: comparison of American College of Obstetricians and Gynecologists practice bulletin with other national guidelines. Am J Obstet Gynecol 2009;200:409 e1–e6.
[6] Santos RF, Miguelote RF, Coelho DM, et al. Validity of classical fetal weight charts in the Portuguese population. Rev Bras Ginecol Obstet 2011;33:164–9.
[7] Ray JG, Sgro M, Mamdani MM, et al. Birth weight curves tailored to maternal world region. J Obstet Gynaecol Can 2012;34:159–71.
[8] Hemming K, Hutton JL, Glinianaia SV, Jarvis SN, Platt MJ. Differences between European birthweight standards: impact on classification of 'small for gestational age'. Dev Med Child Neurol 2006;48:906–12.
[9] Hadlock F, Shah Y, Kanon D, Lindsey J. Fetal crown-rump length: reevaluation of relation to menstrual age (5–18 weeks) with high-resolution real-time US. Radiology 1992;182:501–5.
[10] Hadlock FP, Harrist RB, Carpenter RJ, Deter RL, Park SK. Sonographic estimation of fetal weight. The value of femur length in addition to head and abdomen measurements. Radiology 1984;150:535–40.
[11] Tukey J. Exploratory data analysis. Addison-Wesley; 1977. p. 43–4.
[12] Yudkin PL, Aboualfa M, Eyre JA, Redman CWG, Wilkinson AR. New birthweight and head circumference centiles for gestational ages 24 to 42 weeks. Early Hum Dev 1987;15:45–52.
[13] Cole TJ, Green PJ. Smoothing reference centile curves: the LMS method and penalized likelihood. Stat Med 1992;11:1305–19.
[14] Rigby RA, Stasinopoulos DM. Generalized additive models for location, scale and shape. J R Stat Soc: Ser C Appl Stat 2005;54:507–54.
[15] Stasinopoulos DM, Rigby RA. Generalized additive models for location scale and shape (GAMLSS) in R. J Stat Softw 2007;23:1–46.
[16] Alexander G, Himes J, Kaufman R, Mor J, Kogan M. A United States national reference for fetal growth. Obstet Gynecol 1996;87:163–8.
[17] Villar J, Ismail LC, Victora CG, et al. International standards for newborn weight, length, and head circumference by gestational age and sex: the

Newborn Cross-Sectional Study of the INTERGROWTH-21st Project. Lancet 2014;384:857–68.
[18] Platt RW. The effect of gestational age errors and their correction in interpreting population trends in fetal growth and gestational age-specific mortality. Semin Perinatol 2002;26:306–11.
[19] Doubilet PM, Benson CB, Nadel AS, Ringer SA. Improved birth weight table for neonates developed from gestations dated by early ultrasonography. J Ultrasound Med 1997;16:241–9.
[20] Hadlock FP, Harrist RB, Martinez-Poyer J. In utero analysis of fetal growth: a sonographic weight standard. Radiology 1991;181:129–33.
[21] Gardosi J, Chang A, Kalyan B, Sahota D, Symonds EM. Customised antenatal growth charts. Lancet 1992;339:283–7.
[22] Direção-Geral da Saúde. Norma no. 023/2011: Exames Ecográficos na Gravidez. In: Divisão de Saúde Sexual R, Infantil e Juvenil, ed.; 2011.
[23] Altman DG, Chitty LS. Charts of fetal size: 1. Methodology. BJOG 1994;101:29–34.
[24] Sousa-Santos RF, Dinis-Ribeiro M, Santos CM, Cruz-Correia R, Bernardes J. Different standards in growth centiles for fetuses and newborns: a survey in Europe. Acta Obstet Ginecol Port 2015;9:214–20.
[25] Ferdynus C, Quantin C, Abrahamowicz M, et al. Can birth weight standards based on healthy populations improve the identification of small-for-gestational-age newborns at risk of adverse neonatal outcomes? Pediatrics 2009;123:723–30.
[26] Dudley NJ. A systematic review of the ultrasound estimation of fetal weight. Ultrasound Obstet Gynecol 2005;25:80–9.