DOI: 10.1111/j.1471-0528.2012.03451.x
www.bjog.org

Systematic review

Systematic review of methodology used in ultrasound studies aimed at creating charts of fetal size

C Ioannou,[a] K Talbot,[a] E Ohuma,[a] I Sarris,[a] J Villar,[a,b] A Conde-Agudelo,[c] AT Papageorghiou[a,b]

[a] Nuffield Department of Obstetrics and Gynaecology, University of Oxford, UK [b] Oxford Maternal and Perinatal Health Institute, Green Templeton College, University of Oxford, UK [c] Perinatology Research Branch, Eunice Kennedy Shriver National Institute of Child Health and Human Development/National Institutes of Health/Department of Health and Human Services, Bethesda, MD and Detroit, MI, USA
Correspondence: Dr AT Papageorghiou, Nuffield Department of Obstetrics and Gynaecology, John Radcliffe Hospital, Oxford, OX3 9DU, UK. Email aris.papageorghiou@obs-gyn.ox.ac.uk

Accepted 13 June 2012. Published Online 13 August 2012.

Background Reliable ultrasound charts are necessary for the prenatal assessment of fetal size, yet there is a wide variation of methodologies for the creation of such charts.

Objective To evaluate the methodological quality of studies of fetal biometry using a set of predefined quality criteria of study design, statistical analysis and reporting methods.

Search strategy Electronic searches in MEDLINE, EMBASE and CINAHL, and references of retrieved articles.

Selection criteria Observational studies whose primary aim was to create ultrasound size charts for bi-parietal diameter, head circumference, abdominal circumference and femur length in fetuses from singleton pregnancies.

Data collection and analysis Studies were scored against a predefined set of independently agreed methodological criteria and an overall quality score was given to each study. Multiple regression analysis between quality scores and study characteristics was performed.

Main results Eighty-three studies met the inclusion criteria. The highest potential for bias was noted in the following fields: 'Inclusion/exclusion criteria', as none of the studies defined a rigorous set of antenatal or fetal conditions which should be excluded from analysis; 'Ultrasound quality control measures', as no study demonstrated a comprehensive quality assurance strategy; and 'Sample size calculation', which was apparent in six studies only. On multiple regression analysis, there was a positive correlation between quality scores and year of publication: quality has improved with time, yet considerable heterogeneity in study methodology is still observed today.

Conclusions There is considerable methodological heterogeneity in studies of fetal biometry. Standardisation of methodologies is necessary in order to make correct interpretations and comparisons between different charts. A checklist of recommended methodologies is proposed.

Keywords Biometry, fetal growth, guideline, methodology, quality, review, ultrasound.

Please cite this paper as: Ioannou C, Talbot K, Ohuma E, Sarris I, Villar J, Conde-Agudelo A, Papageorghiou A. Systematic review of methodology used in ultrasound studies aimed at creating charts of fetal size. BJOG 2012;119:1425–1439.

Introduction

Prenatal evaluation of fetal size was made possible in the 1960s by the introduction of combined A- and B-mode ultrasonography in obstetrics.[1] Second- and third-trimester fetal biometry is now common practice and represents one of the most common medical investigations undertaken. However, a Cochrane review of trials evaluating routine ultrasonography beyond 24 weeks of gestation in low-risk pregnancies has not demonstrated any benefit in perinatal outcomes.[2] Furthermore, intrauterine growth restriction remains a leading cause of perinatal loss, accounting for at least one-fifth of stillbirths in the UK;[3] failure to diagnose and lack of effective treatment are the likely explanations.

The recognition of pathological growth is dependent on the existence of reliable standards. However, the establishment of normal charts for key biometric variables, such as bi-parietal diameter (BPD), head circumference (HC), abdominal circumference (AC) and femur length (FL), is not straightforward. Discrepancies in median values and percentile curves between studies have often been attributed to different distributions in racial,[4] gender[5] or other

biological and demographic determinants.[6] However, it is likely that differences in study design, data analysis and presentation also contribute to the observed discrepancies. Ultrasound measurement is often subject to observer error, and it is possible that systematic variations in measurement accuracy may exist between different studies. Suboptimal methodology when producing a fetal size chart is likely to affect the ability to discriminate the healthy from the compromised fetus.

Recommended methods in various aspects of study design have been published in recent years,[7–11] including appropriate statistical methods for modelling cross-sectional[10] (one scan per fetus) or longitudinal[11] (serial scans) ultrasound data. Conversely, other aspects, such as sample size, population selection and exclusion criteria, remain debated to this day.

The aim of this systematic review was to evaluate the methodological quality of studies of fetal biometry using a set of predefined quality criteria of study design, statistical analysis and reporting methods.

Table 1. Search strategy

Fetal Development/
*Gestational Age/
((fetal or foetal or fetus or foetus) adj2 growth).tw.
((fetal or foetal or fetus or foetus) adj biometr*).tw.
1 or 2 or 3 or 4
(growth adj (curve* or chart* or standard* or index or indices)).tw.
(reference adj (curve* or chart* or index or indices or equation* or value* or range* or equation* or centile* or percentile*)).tw.
(biometr* adj (curve* or chart* or index or indices or equation* or value* or range* or equation* or centile* or percentile*)).tw.
(size adj (chart* or curve*)).tw.
(dating adj (curve* or chart*)).tw.
*Reference Values/
6 or 7 or 8 or 9 or 10 or 11
Ultrasonography, Prenatal/
(ultrasound* or ultrasonogra* or sonogra*).tw.
13 or 14
5 and 12 and 15
exp animal/not human/
16 not 17
limit 18 to English language

* indicates keyword truncation.

Methods

This systematic review of observational studies was conducted and reported following the checklist proposed by the MOOSE group.[12] Three major electronic databases (MEDLINE, EMBASE and CINAHL) were systematically searched for the period 1968 to September 2011 to identify studies of two-dimensional ultrasound biometry. Reference lists of retrieved full-text articles were examined for additional, relevant citations. Studies were included if the primary objective was to create size charts for BPD, HC, AC and FL on B-mode ultrasound in normal singleton pregnancies using either a cross-sectional or longitudinal design. The search was not restricted by study design or methodology, but only articles written in English were considered. Articles were excluded if: (1) only A-mode ultrasound was used; (2) the primary aim was other than the construction of size charts, for instance, the prediction of gestational age or comparisons between different population groups; and (3) the entire gestation was not covered; for instance, size chart from 20 to 30 weeks of gestation.

The keyword search strategy, which was constructed by a professional information specialist, is presented in Table 1. Two reviewers (KT and CI) screened the titles and abstracts of all identified citations, and selected potentially eligible studies. The full-text versions of eligible studies were independently assessed by the same reviewers and any disagreements were resolved by consensus or consultation with a third reviewer (ATP). Authors' institutions were contacted in order to obtain a copy of the published article where this was not available from library sources. The flow chart of the literature search is presented in Figure 1. Studies excluded from this review and the reasons for exclusion are listed in Appendix S1.

A list of methodological quality criteria (shown in Table 2) was initially developed by one of us (AC-A) in advance of the review, and agreed between three of the authors (JV, AC-A and ATP) independent of those who performed the data abstraction. These quality criteria are based on available published research,[7–11] and are divided into three domains: study design, statistical methods and reporting methods; in total, 23 quality criteria were used for cross-sectional studies and 24 criteria for longitudinal studies.

The included studies were reviewed by two obstetricians (CI and KT) and a medical statistician (EO), and study details were abstracted onto an Excel spread sheet. Studies were assessed against each criterion within the checklist and were scored as either 'high' or 'low' risk of bias. Disagreements were resolved either by consensus or consultation with a third reviewer (ATP). The overall quality score was defined as the percentage of 'low risk of bias' marks over the total number of quality criteria for each study.

Multiple regression analysis was performed between quality scores and study characteristics which were not part of the scoring algorithm: year of publication, sample size of participating women, sample size of included ultrasound examinations, study duration, type of participating hospitals (teaching versus nonteaching), number of participating sites (single versus multi-site), number of sonographers (single versus multiple) and type of country (low-, middle- or high-income country, using the 2010 World Bank

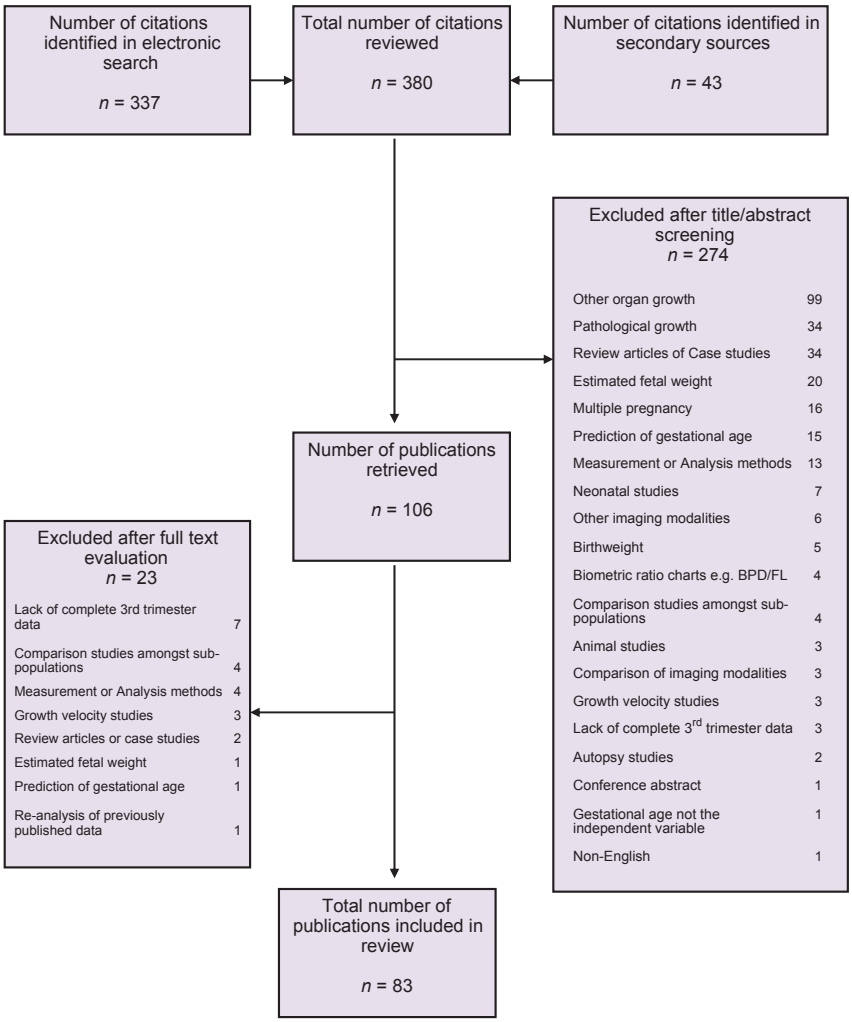

Figure 1. Consort flow diagram of literature assessment.

Classification of economies by gross national income). Statistical analyses were performed using Microsoft Excel 2010 and IBM SPSS Statistics version 19.

Results

Eighty-three publications from 32 countries were identified;[14–96] the earliest was published in 1971 and the latest in 2008. In two publications,[49,76] multiple study designs were used, for instance AC using a longitudinal and a cross-sectional design; the best described methodology with the highest quality score was used in the final analysis. The median sample size of participating women was 558 (minimum, 19; maximum, 17 660; interquartile range, 1120), whereas the median number of ultrasound examinations was 800 (minimum, 167; maximum, 50 131; interquartile range, 1770). Forty studies reported one biometric parameter only, whereas the remaining 43 studies included combinations of BPD, HC, AC or FL, but only 22 studies reported a complete set of all four parameters. In total, there were 60 studies for BPD, 34 for HC, 41 for AC and 43 for FL.

The study characteristics and overall quality score for each study are presented in Table 3. The breakdown of scores per study and per quality criterion is presented in Appendix S2. Additional characteristics not included in the scoring algorithm are presented in Appendices S3 (Additional maternal and pregnancy characteristics) and S4 (Additional study characteristics). Sixty-one studies had a cross-sectional design and 22 studies had a longitudinal design. Amongst the 61 cross-sectional studies, only 36 (59%) clearly specified that one examination per fetus was performed during the study period. Amongst the 22 longitudinal studies, only in 12 (55%) were ultrasound data analysed using a method that took into account their serial nature.

Table 2. Methodological quality criteria

Domain	Low risk of bias	High risk of bias
1. Study design		
1.1 Design	Clearly described as either cross-sectional or longitudinal	Not reported Mixture of cross-sectional and longitudinal data
1.2 Sample selection	Population-based study where there are attempts to identify and clearly define populations from a specific geographical area; from this underlying population, women are selected either consecutively or at random	Not population based; convenience sampling; arbitrary recruitment; or not reported
1.3 Number of occasions each fetus was measured (only for cross-sectional studies)	Each fetus was measured and included only once	Some fetuses were measured and included more than once
1.4 Method of selecting the gestational ages at which the fetuses were measured (only for longitudinal studies)	Interval of measures prospectively prespecified and justified	Interval of measures not prospectively prespecified and justified or not reported
1.5 Reason(s) for choosing a particular number of serial measurements (only for longitudinal studies)	Clear documentation of the intended number of serial measurements	No clear documentation of the intended number of serial measurements
1.6 Inclusion/exclusion criteria	The study made it clear that women at high risk of pregnancy complications were not included, and that women with abnormal outcome were excluded, i.e. an effort was made to include 'normal' outcome as best possible As a minimum, the study population should exclude: – multiple pregnancy – fetuses with congenital structural or chromosomal anomalies – fetal death – women with disorders that may affect fetal growth (at least should specify exclusion of women with pre-existing hypertension, diabetes mellitus, renal disease and smoking) – pregnancy complications (at least pre-eclampsia) – pregnancies conceived by assisted reproductive technology	The study population included both low-risk and high-risk pregnancies, or women with abnormal outcome were not excluded Study population that did not exclude fetuses or women with the characteristics previously described Exclusions which would have a direct effect on the estimated percentiles, such as fetuses found at birth to be large or small for dates
1.7 Sample size	*A priori* determination/calculation of sample size and justification	Lack of *a priori* sample size determination/ calculation and justification
1.8 Data collection	Prospective study and ultrasound data collected specifically for the purpose of constructing charts of fetal size or fetal growth	Retrospective study, or data not collected specifically for the purpose of constructing charts of fetal size or fetal growth, or unclear (e.g. use of routinely collected data)
1.9 Method of dating pregnancy	Clearly described Known last menstrual period (LMP) and regular menstrual cycles prior to pregnancy AND a sonogram before 14 weeks demonstrating a crown–rump length (CRL) that *corroborates LMP dates* (within how many days unspecified)	Not described clearly Gestational age assessment at >14 weeks, or gestational age assessment not including ultrasonographic verification
1.10 Collection of data on gestational age at inclusion	The gestational age was calculated precisely to the day	Truncation of gestational age to the number of 'completed weeks'
2. Statistical methods		
2.1 Number of measurements taken for each biometric variable	More than one measure per fetus per scan	Single measure or not specified
2.2 Statistical methods	Clearly described and identified	Not clearly described and identified
2.3 Assessment of increasing variability of the data with gestation	Performed	Not performed

Table 2. (Continued)

Domain	Low risk of bias	High risk of bias
2.4 Assessment of goodness of fit of the models	A test of goodness of fit of the models was reported	Goodness of fit of models was not reported
2.5 Scatter diagram of the data with the fitted percentiles superimposed	Study included scatter diagrams of the data with the percentiles superimposed	Study did not include scatter diagrams of the data with the percentiles superimposed
2.6 Change in reference percentiles across gestational age	Smooth change	Not smooth change
2.7 Methods used to estimate age-specific reference intervals for fetal size measurements	'Mean and standard deviation (SD) model', smoothed crude percentiles, or 'LMS method'[13]	Inadequate
3. Reporting methods		
3.1 Characteristics of study population	Presented in a table or clearly described, and includes minimum dataset of age, weight, height or body mass index and parity	Not presented in a table or not clearly described, or does not contain minimum dataset
3.2 Description of number approached/ enrolled	Described	Not described
3.3 Ultrasound machine(s) used	Clearly specified	Not clearly specified
3.4 Number of sonographers that took the measurements	Reported	Unreported
3.5 Description of measurement techniques	The study described sufficient and unambiguous details of the measurement techniques used for fetal size parameters, including imaging plane and calliper application method	The study did not describe sufficient and unambiguous details of the measurement techniques used for fetal size parameters
3.6 Contains quality control measures	Should include the following: – assessment of intraobserver variability – assessment of interobserver variability – image review – image scoring – image storage	Does not contain quality control measures
3.7 Report of mean and SD of each measurement and the sample size for each week of gestation	Presented in a table or clearly described	Not presented in a table or not clearly described
3.8 Report of regression equations for the mean (and SD if relevant) for each measurement	Reported	Not reported

In only 20 of the 83 studies (24%) were the ultrasound data collected prospectively and explicitly for research purposes, whereas, in 13 studies (16%), a retrospective analysis of an existing database was performed. In all the remainder, it was unclear whether a prospective or retrospective design was used, or whether the ultrasound examinations were performed for research purposes or as part of routine clinical care.

The frequencies of 'low risk of bias' in each of the three groups of methodological criteria are presented in Figures 2–4. Highest risk of bias was noted in the following fields: 'Inclusion/exclusion criteria', where none of the studies defined a rigorous set of antenatal or fetal conditions that should be excluded from analysis in order to ensure a normal pregnancy outcome (Figure 2, item 1.6); 'Ultrasound quality control measures', where no study demonstrated a comprehensive quality assurance strategy (Figure 4, item 3.6); and 'Sample size calculation', which was apparent in only six studies (Figure 2, item 1.7).

Although some individual criteria of participant selection were used in different studies, such as shown in Figure 5, there was no study which systematically used all of these. Conversely, in 22 studies (27%), inappropriate exclusions were applied, such as removing from the final analysis cases on the outer percentiles of the ultrasound measurement or birthweight.

None of the studies in this review used a comprehensive ultrasound quality control strategy incorporating the items in Figure 6. In approximately one-half of the studies (40 studies), ultrasound examinations were performed by multiple sonographers, yet an exercise to standardise participating sonographers was reported in only four studies. No study reported the use of an image scoring method for the purpose of ultrasound quality assurance; in only four

Table 3. Characteristics of the included studies

Reference	Year	Country	Number of women	Number of scans	Weeks	Measurement	Design	Data collection	Scans for research purposes only	Quality score (%)
Aickin et al.[14]	1976	New Zealand	446	1018	19–41	BPD	CM	Unclear	Unclear	17
Al-Meshari & Raber[15]	1987	Saudi Arabia	1570	1570	12–41	BPD	C1	Unclear	Unclear	30
Amoa et al.[16]	1993	Papua New Guinea	425	1229–1514	12–42	BPD, AC, FL	LC	Unclear	Unclear	21
Ashrafunnessa et al.[17]	2003	Bangladesh	710	710	12–42	BPD, AC, FL	C1	Prospective	Yes	70
Ayangade & Okonofua[18]	1986	Nigeria	558	712	12–40	BPD	CM	Unclear	Unclear	26
Beigi & ZarrinKoub[19]	2000	Iran	1324	15 594–15 693	12–40	BPD, FL	LC	Unclear	Unclear	13
Bergsjo et al.[20]	1976	Norway	131	421	24–44	BPD	CM	Unclear	No	13
Brons et al.[21]	1990	Netherlands	63	520	12–40	BPD, HC, AC, FL	LL	Unclear	Unclear	42
Browne et al.[22]	1992	USA	8285	7249–8108	10–44	BPD, HC, AC, FL	C1	Retrospective	No	35
Campbell & Newman[23]	1971	UK	574	1029	13–40	BPD	CM	Unclear	Unclear	17
Chan & Yeo[24]	1991	Singapore	1442	1442	17–40	BPD	C1	Retrospective	No	22
Chang et al.[25]	1996	Taiwan	2077	2077	16–41	BPD	C1	Unclear	Unclear	52
Chitty et al.[26]	1994	UK	610	425–610	12–42	AC	C1	Prospective	Yes	78
Chitty et al.[27]	1994	UK	649	649	12–42	FL	C1	Prospective	Yes	78
Chitty et al.[28]	1994	UK	594	594	12–42	BPD, HC	C1	Prospective	Yes	78
de la Vega et al.[29]	2008	Puerto Rico	548	548	14–38	BPD, HC, AC, FL	C1	Unclear	No	13
Deter et al.[30]	1982	USA	252	252	13–41	HC, AC	C1	Unclear	Unclear	35
Deter et al.[31]	1982	USA	20	Unclear	12–40	BPD, HC, AC	LL	Unclear	Unclear	33
Deter et al.[32]	1987	USA	20	Unclear	15–40	FL	LL	Unclear	Unclear	33
Di Battista et al.[33]	2000	Italy	238	1237–1539	12–40	BPD, HC, AC, FL	LL	Unclear	Unclear	42
Dubiel et al.[34]	2008	Poland	959	959	20–42	BPD, HC, AC, FL	CS	Prospective	Unclear	43
Elejalde & de Elejalde[35]	1986	USA	1068	2409	10–40	BPD, FL	CM	Unclear	No	22
Eriksen et al.[36]	1985	Denmark	41	493	13–40	BPD	LL	Prospective	Unclear	42
Exacoustos et al.[37]	1991	Italy	2317	2317	13–40	FL	C1	Unclear	No	48
Fescina et al.[38]	1982	–	30	692–722	13–40	BPD, HC, AC	LC	Unclear	Unclear	25
Flamme[39]	1972	Belgium	1394	4170	20–42	BPD	CM	Unclear	Unclear	0
Gallivan et al.[40]	1993	UK	67	434	20–40	AC	LL	Prospective	Yes	58
Guihard-Costa et al.[41]	1995	France	3433	3974–4248	8–41	BPD, FL	CM	Retrospective	Unclear	30
Hadlock et al.[42]	1982	USA	400	400	15–41	HC	C1	Unclear	Unclear	43
Hadlock et al.[45]	1982	USA	338	338	12–40	FL	CS	Unclear	No	35
Hadlock et al.[43]	1982	USA	400	400	15–41	AC	C1	Unclear	Unclear	39
Hadlock et al.[44]	1982	USA	533	533	12–40	BPD	C1	Unclear	Unclear	43
Hoffbauer et al.[46]	1979	Germany	430	800	12–40	BPD, HC, AC	CM	Unclear	Unclear	9
Issel et al.[48]	1975	Germany	Unclear	5400	12–43	BPD	CS	Unclear	Unclear	9
Issel[47]	1985	Germany	Unclear	1000	17–41	FL	CS	Retrospective	Unclear	9
Jeanty et al.[52]	1981	Belgium	450	Unclear	15–40	FL	CS	Unclear	Unclear	17
Jeanty et al.[49]	1984	Belgium	45	695	12–40	AC	LL	Prospective	Yes	67
Jeanty et al.[51]	1984	Belgium	45	696	10–40	BPD, HC	LL	Prospective	Yes	63

Table 3. (Continued)

Reference	Year	Country	Number of women	Number of scans	Weeks	Measurement	Design	Data collection	Scans for research purposes only	Quality score (%)
Jeanty et al.[50]	1984	Belgium	45	646	12–40	FL	LL	Prospective	Yes	63
Johnsen et al.[53]	2006	Norway	650	2489–2589	10–42	BPD, HC, AC, FL	LL	Prospective	Yes	67
Jung et al.[54]	2007	South Korea	10 455	10 455	12–40	BPD, HC, AC, FL	C1	Retrospective	No	61
Kurmanavicius et al.[55]	1999	Switzerland	6557	5462–6217	12–42	BPD, HC	C1	Retrospective	No	65
Kurmanavicius et al.[56]	1999	Switzerland	6557	5807–5860	12–42	AC, FL	C1	Retrospective	No	65
Lai & Yeo[57]	1995	Singapore	6374	6017–6131	14–41	BPD, HC, AC, FL	C1	Retrospective	No	61
Larsen et al.[58]	1990	Denmark	35	Unclear	14–40	BPD, AC, FL	LC	Unclear	Unclear	42
Lei & Wen[59]	1998	China	5496	5496	16–40	BPD, HC, AC, FL	C1	Prospective	No	22
Lessoway et al.[60]	1998	Canada	1396	747–790	11–42	BPD, HC, AC, FL	C1	Prospective	No	43
Leung et al.[61]	2008	China	709	679–708	12–40	BPD, HC, AC, FL	C1	Prospective	Yes	70
Levi et al.[63]	1973	Belgium	1011	3032	15–43	BPD	CM	Unclear	Unclear	17
Levi et al.[62]	1975	Belgium	Unclear	767	19–43	BPD, HC	CS	Unclear	Unclear	9
Lu et al.[64]	2008	Taiwan	Unclear	50 131	14–41	AC	CS	Retrospective	No	35
Mathai et al.[65]	1995	India	120	477–498	20–40	BPD, AC, FL	LC	Prospective	Yes	17
Merz et al.[66]	1987	Germany	530	530	13–42	FL	C1	Prospective	Unclear	26
Munjanja et al.[67]	1988	Zimbabwe	190	857–1233	12–40	BPD, HC	LC	Unclear	Unclear	13
Munoz et al.[68]	1986	South Africa	1842	1842	15–40	BPD	C1	Unclear	Unclear	52
Nasrat & Bondagii[69]	2005	Saudi Arabia	Unclear	Unclear	14–40	BPD, HC, AC, FL	C1	Retrospective	No	22
Neufeld et al.[70]	2004	Guatemala	319	684	14–38	BPD, HC, AC, FL	LC	Prospective	Yes	33
O'Brien & Queenan[71]	1981	USA	411	1016	14–40	FL	CM	Unclear	Unclear	22
Okonofua et al.[72]	1988	Nigeria	200	219	20–40	AC	CM	Unclear	No	17
Okupe et al.[73]	1984	Nigeria	552	1104	12–40	BPD	CM	Unclear	Unclear	17
Paladini et al.[74]	2005	Italy	626	623–625	16–40	BPD, HC, AC, FL	C1	Prospective	Yes	65
Pang et al.[75]	2003	China	500	2349	24–40	BPD, HC, AC, FL	LL	Prospective	Yes	50
Persson et al.[76]	1978	Sweden	93	707	15–42	BPD	LC	Unclear	Yes	46
Persson & Weldner[77]	1986	Sweden	19	167	11–39	BPD, FL	LC	Prospective	Yes	42
Pineau et al.[78]	2003	France	1336	1336	12–38	BPD, HC	C1	Retrospective	No	39
Queenan et al.[79]	1976	USA	468	738	17–43	BPD	CM	Unclear	Unclear	13
Saksiriwuttho et al.[80]	2007	Thailand	628	628	14–41	BPD, HC, AC, FL	C1	Prospective	Yes	35
Salomon et al.[81]	2006	France	Unclear	19 647	15–40	BPD, HC, AC, FL	CS	Unclear	No	48
Schluter et al.[82]	2004	Australia	17 660	15 871–20 520	11–41	BPD, HC, AC, FL	C1	Unclear	No	61
Shohat & Romano-Zelekha[83]	2001	Israel	1143	1143	13–41	BPD, FL	C1	Prospective	No	22
Siwadune et al.[84]	2000	Thailand	613	613	12–41	BPD	C1	Prospective	Yes	65
Smulian et al.[85]	2001	USA	10 070	10 070	11–42	AC	CS	Retrospective	No	43
Snijders & Nicolaides[86]	1994	UK	1040	1040	14–40	BPD, HC, AC, FL	C1	Retrospective	No	61
Sunsaneevithayakul et al.[87]	2000	Thailand	615	615	12–41	AC	C1	Prospective	Yes	70
Tamura & Sabbagha[88]	1980	USA	200	536	18–41	AC	CM	Unclear	Unclear	9
Thame et al.[89]	2003	Jamaica	499	2574	13–37	BPD, HC, AC, FL	LC	Prospective	Unclear	42

Table 3. (Continued)

Reference	Year	Country	Number of women	Number of scans	Weeks	Measurement	Design	Data collection	Scans for research purposes only	Quality score (%)
Titapant et al.[90]	2000	Thailand	608	608	12–41	FL	C1	Prospective	Yes	70
Todros et al.[91]	1987	Italy	1426	603–1283	12–42	BPD, HC, AC	C1	Unclear	Unclear	30
Verburg et al.[92]	2008	Netherlands	8313	20 277–22 271	10–40	BPD, HC, AC, FL	LL	Prospective	Yes	75
Warda et al.[93]	1985	USA	254	254	13–39	FL	C1	Unclear	No	57
Watson[94]	1986	Australia	202	657	12–40	BPD	CM	Unclear	Yes	26
Westerway et al.[95]	2000	Australia	3800	Unclear	11–40	BPD, HC, AC, FL	CM	Unclear	No	22
Wladimiroff et al.[96]	1978	Netherlands	303	303	24–41	BPD	C1	Unclear	No	17

AC, abdominal circumference; BPD, bi-parietal diameter; FL, femur length; HC, head circumference; CM, cross-sectional data, several scans per fetus; C1, cross-sectional data, one scan per fetus; CS, cross-sectional data, number of scans per fetus not stated; LC, longitudinal data with cross-sectional analysis; LL, longitudinal data with longitudinal analysis.

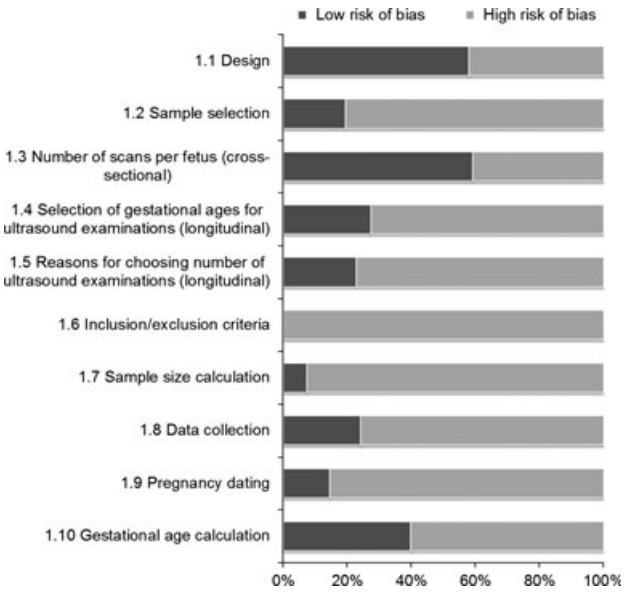

Figure 2. Methodological quality of included studies: study design.

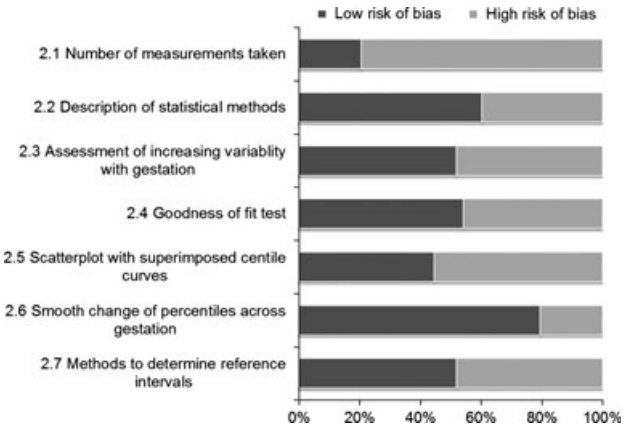

Figure 3. Methodological quality of included studies: statistical methods.

studies was it ensured that sonographers were blind to the actual measurement recorded during the examination.

Table 4 shows the different methods used for gestational age estimation. Last menstrual period (LMP) used in isolation remains the most popular method. Only 12 studies (14%) used a dating method considered to be at low risk of bias, namely either CRL alone or LMP confirmed by CRL.

Results from individual studies were reported in the form of tables, equations or charts as demonstrated in Figure 7. Although tables of median values (68 studies) and tables of percentile ranges (65 studies) were a common method of presentation, half of the time these contained the raw unmodelled data; fitted median values following

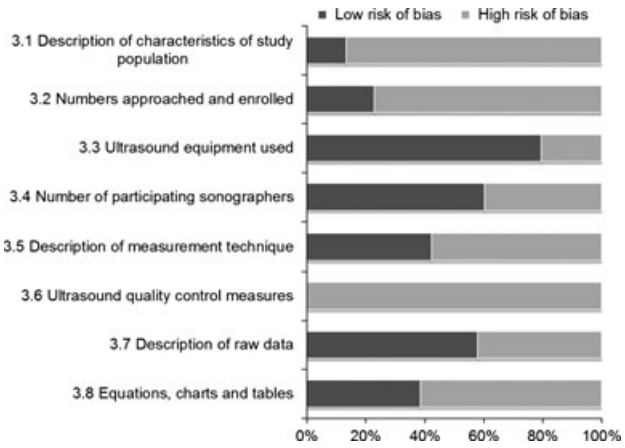

Figure 4. Methodological quality of included studies: reporting methods.

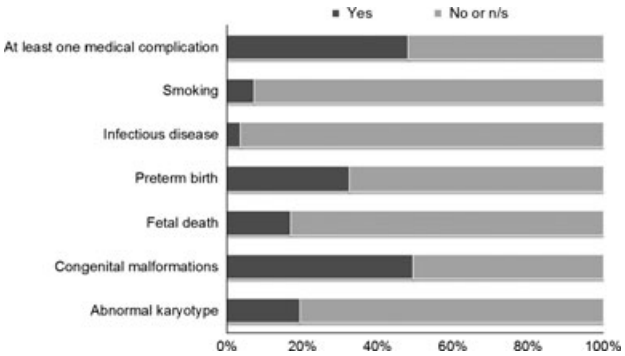

Figure 5. Exclusion criteria used in the included studies.

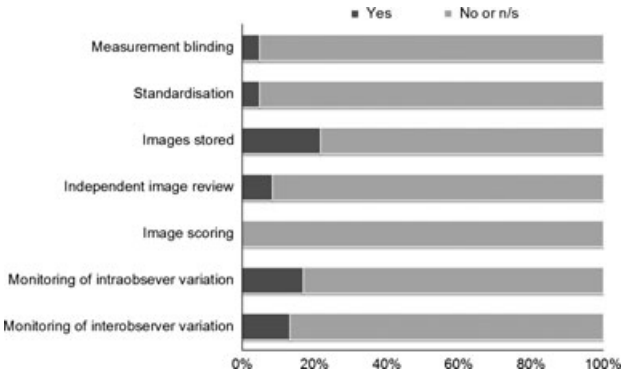

Figure 6. Ultrasound quality assurance measures in the included studies.

analytical modelling were presented in 35 studies and fitted percentiles in 33 studies. An equation for the median was reported in 50 of 83 studies, whereas the standard deviation was mathematically expressed in 32 of 83 studies, either as a fixed number or as a function of gestation.

Table 4. Dating methods used in the included studies

Type of dating	Number of studies (%)
LMP only	37 (45)
LMP confirmed by US parameter (non-CRL)	21 (25)
LMP confirmed by CRL	9 (11)
Other	6 (7)
Not stated	5 (6)
CRL only	3 (4)
US parameter only (non-CRL)	2 (2)

CRL, crown–rump length; LMP, last menstrual period; US, ultrasound.

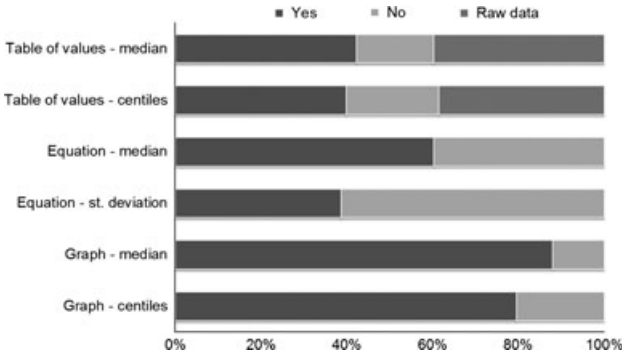

Figure 7. Use of presentation methods in the included studies.

Printed charts of the median and percentile curves were seen in the vast majority of the publications.

On univariate regression analysis, positive predictors of quality score were the year of publication ($P < 0.001$), sample size of participating women ($P = 0.04$) and teaching (as opposed to nonteaching) hospital status ($P = 0.003$). On multiple regression analysis, however, the effect of sample size and hospital type became nonsignificant, whereas only the year of publication persisted as a significant predictor of quality score with a coefficient of determination $R^2 = 0.24$. A scatterplot of quality scores versus year of publication is shown in Appendix S5.

Discussion

This review has revealed substantial heterogeneity of methodology used in ultrasound studies of fetal biometry. A predefined quality scoring sheet was used in the assessment of the included studies. This checklist is not intended to commend or discard studies, but rather to be used as a consensus guideline in order to improve consistency in fetal growth research.

This review has several strengths. In the literature search, there were no restrictions by year of publication, as some

of the older ultrasound charts may still be used in current clinical practice. The quality criteria were based on the best available evidence and were agreed independent of the reviewers who performed the data abstraction. The use of a quality score in percentage form allowed an objective rather than empirical assessment of quality and also enabled regression analyses in order to identify temporal or other trends. However, there are also some limitations. English language restriction was imposed and it is possible that studies of normative fetal biometry from non-English-speaking countries may have been missed. This was imposed for practical reasons; unlike systematic reviews of treatment effect, where it is imperative that all evidence is found, our aim was to assess the methodological quality of studies on fetal size. Another limitation was that the reviewers who performed the data abstraction were not blind to the origin and authors of the included studies. Finally, the older studies in this review were tested against some quality criteria which have only been established in recent years. Although it may seem unfair to judge previously published work by today's standards, it is important for clinicians who choose amongst any of the published ultrasound charts to have an up-to-date assessment of their methodological quality.

Multiple regression analysis demonstrated that quality scores have significantly improved in recent years. This is logical as the present criteria of quality have been developed through gradual improvement in both the ultrasound technology and the statistical methods of data analysis. However, considerable heterogeneity still persists today: even in the last decade, quality scores of published studies have varied considerably around the average trend. It is possible that such differences in study quality may explain discrepancies in the reported fetal size curves between different populations. Clearly defined and consistent methodology is therefore necessary in order to critically appraise such population differences.

For instance, none of the included studies scored a 'low risk of bias' for their inclusion and exclusion criteria. This scoring field considered that a study should exclude women with conditions and outcomes associated with pathological fetal growth. The aim of a fetal size chart should be to depict how infants *should* grow under optimal conditions (a 'prescriptive' standard) rather than *how they often* grow (a 'descriptive' reference).[97] To achieve this, it is necessary to consider factors that influence growth. A number of such factors are well established: maternal smoking;[98] maternal disease, e.g. chronic hypertension or diabetes;[99] pregnancy-induced hypertension;[100] pre-eclampsia;[100] abnormal karyotype;[99] congenital anomalies;[99] pre-term delivery;[101] stillbirth.[102] Some of these conditions were excluded in one or more studies in this review, but no study excluded all of them.

Authors have previously argued against this, claiming that an unselected population ensures a better representation of the underlying population,[9] or that some exclusion criteria, such as maternal smoking for instance, are not reasonable.[9,103] However, if strong evidence exists for each of them, then all should be excluded. It is unfortunate that the term 'supernormalisation'[103] has acquired a negative connotation in the past.

Conversely, it is arbitrary and illogical to exclude from analysis the outer percentiles of the ultrasound measurement values. For instance, measurements greater than 1.66 or two standard deviations from the mean[31,64] may well represent physiological healthy variability. Similarly, exclusion of cases with birthweight in the outer percentiles has been used by some studies[80,91] in order to define normal growth. The fundamental concern with this approach is the risk of excluding fetuses that are healthy, but constitutionally small or large for gestational age.

Only one-quarter of the studies in this review had a prospective design in which ultrasound examinations were performed for research purposes only. This is an important point. Most clinical ultrasound services now routinely collect information in computerised databases. Retrospective analysis of such databases is a practical solution to the generation of a large sample size; however, the ability of the researchers to address potential confounders is curtailed. For instance, reference curves can be skewed by the fact that a proportion of the examinations are clinically indicated as a result of suspected pathological growth. It is also difficult to retrospectively ascertain maternal or fetal complications, unless a reliable coding system is in place. The alternative is a prospective study. Strictly speaking, this indicates that the study design, participant recruitment and collection of clinical, demographic and ultrasound data are carried out with the objective of creating size charts. As a result of inconsistent terminology and ambiguous description, it was often difficult to identify prospective studies from retrospective database analyses in this review.

Accurate estimation of (gestational) age is a fundamental prerequisite for creating any size chart. It is already recognised that suboptimal dating in older ultrasound studies may have contributed to the observed flattening or reversal of growth at term.[104] In addition, in studies up to the mid-1980s, the research objectives often overlapped between the assessment of fetal size and pregnancy dating; several studies used a biometric parameter, such as BPD, both as a means of dating and as a predictor of growth, in an almost circular fashion.[31,42,45]

Several different dating strategies were encountered in this review. On certain occasions, dating was either vaguely stated or not described at all. There is now robust evidence that early ultrasonographic determination of gestational age

is more reliable than LMP alone.[105] It has also been argued that, for clinical management, ultrasonography alone may be marginally superior to LMP confirmed by ultrasound,[106] although such differences are small. However, when the objective is chart creation, the use of ultrasound for dating and then again for the assessment of fetal size creates a circular argument; for instance, it is possible that a smaller than expected CRL measurement may be indicative of early growth restriction or adverse pregnancy outcome,[107] but such an association will be missed if the gestational age is recalculated on the basis of CRL. Therefore, in the case of studies aiming to create fetal size charts, we believe that it is conceptually more appropriate to estimate the date of confinement using an independent method, such as the LMP, provided that this is corroborated by CRL measurement. In the methodological assessment, we classified as 'low risk of bias' studies that used dating either by CRL alone, or by LMP corroborated by CRL. In any case, the parameter used for sonographic dating and the gestational window during which this is applied should always be clearly specified, in order to demonstrate that methods are adequately standardised; this also highlights the recommended dating practice for institutions who wish to adopt these fetal size charts.

Table 5. Comparison of measurement values amongst highest scoring studies

Reference	28 weeks			32 weeks			36 weeks		
	10th centile	50th centile	SD	10th centile	50th centile	SD	10th centile	50th centile	SD
(A) Bi-parietal diameter (BPD) studies									
Chitty et al.[28]	69.29	73.35	3.17	78.95	83.32	3.41	86.78	91.45	3.64
Verburg et al.[92]	70.00	73.68	2.87	80.22	84.38	3.24	88.90	93.55	3.63
Leung et al.[61]	68.09	72.20	3.21	78.04	82.33	3.35	85.55	90.02	3.49
Johnsen et al.[53]	69.00	73.00	3.12	78.00	83.00	3.90	86.00	91.00	3.90
Paladini et al.[74]	66.20	70.90	3.67	75.70	80.30	3.59	84.10	88.10	3.12
Kurmanavicius et al.[55]	70.49	74.94	3.47	79.94	84.74	3.74	87.44	92.56	3.99
Siwadune et al.[84]	65.71	68.32	2.04	74.93	77.91	2.32	81.75	85.15	2.65
(B) Head circumference (HC) studies									
Chitty et al.[28]	249.10	262.50	10.45	282.70	297.30	11.39	309.00	324.80	12.32
Verburg et al.[92]	251.90	262.70	8.42	285.40	297.80	9.67	309.50	323.50	10.92
Leung et al.[61]	246.90	258.76	9.25	281.22	293.54	9.61	305.60	318.38	9.97
Johnsen et al.[53]	246.00	259.00	10.14	279.00	293.00	10.92	304.00	320.00	12.48
Paladini et al.[74]	250.00	263.50	10.53	280.90	296.10	11.86	303.30	320.00	13.03
Kurmanavicius et al.[55]	248.43	262.73	11.15	279.87	295.32	12.05	303.26	319.87	12.96
(C) Abdominal circumference (AC) studies									
Chitty et al.[26]	212.92	230.57	13.77	248.97	269.71	16.18	282.59	306.41	18.58
Verburg et al.[92]	224.30	239.40	11.78	261.10	278.60	13.65	292.10	312.20	15.68
Leung et al.[61]	218.95	233.90	11.66	256.61	273.56	13.22	290.68	309.64	14.79
Sunsaneevithayakul et al.[87]	217.04	232.58	12.12	254.89	273.26	14.33	289.72	309.12	15.13
Johnsen et al.[53]	223.00	240.00	13.26	262.00	282.00	15.60	299.00	321.00	17.16
Jeanty et al.[49]	208.01	225.24	13.44	245.13	262.36	13.44	276.07	293.30	13.44
Paladini et al.[74]	218.20	239.30	16.46	252.30	275.00	17.71	284.30	307.00	17.71
Kurmanavicius et al.[56]	213.21	231.78	14.49	249.60	270.61	16.39	283.33	306.78	18.29
(D) Femur length (FL) studies									
Chitty et al.[27]	49.25	52.70	2.69	57.43	61.18	2.93	64.14	68.19	3.16
Verburg et al.[92]	49.68	52.47	2.18	57.43	60.46	2.36	63.24	66.52	2.56
Leung et al.[61]	47.22	50.02	2.18	55.20	58.17	2.32	62.33	65.46	2.44
Titapant et al.[90]	46.45	48.95	1.95	54.39	57.10	2.11	60.95	63.89	2.29
Johnsen et al.[53]	48.00	52.00	3.12	56.00	60.00	3.12	63.00	67.00	3.12
Paladini et al.[74]	49.50	52.40	2.26	56.90	60.50	2.81	62.90	67.20	3.35
Kurmanavicius et al.[56]	48.55	52.30	2.93	56.96	60.85	3.03	64.28	68.28	3.12

For each parameter, the five highest scoring studies were identified; additional studies were included if more than one study occupied the fifth place; one study by Ashrafunnessa et al.[17] is not included despite having the third highest quality score (70%), because data for the fitted centiles were not presented, meaning that we could not derive values for this table. All BPD measurements were outer to outer, except in Siwadune et al.,[84] where the outer to inner convention was used; all measurements quoted in millimetres; all centiles are fitted centiles following modelling; SD, fitted standard deviation.

Considerable variation exists in the collection of ultrasound data, the number of sonographers and scan machines used, measurement method definition, standardisation and the monitoring of ultrasound data quality. We have proposed a comprehensive quality control strategy which includes saving and independently reviewing scan images, the use of an image scoring method and the assessment of intra- and interobserver variability of measurement. Studies with multiple sonographers are preferable as they reflect real clinical practice, provided that such strategies for quality assurance are in place. Explicit description of measurement planes and calliper application conventions are necessary. A formal standardisation exercise prior to the start of a multi-sonographer study has been shown to increase data consistency.[108] Blinding of the sonographers to their own measurements is a reasonable effort to remove observer bias.

A consensus has been reached in recent years regarding the appropriate methods for statistical modelling and data presentation.[9–11] Presentation of the raw measurement data only is not clinically useful or informative. Both the median and variance should be modelled as a function of gestational age in a manner that accounts for the increasing variability with gestation and provides smooth percentile curves; goodness of fit testing should demonstrate that these curves describe accurately the structure of the raw data.[10] Finally, the results should be presented in the form of tables of fitted percentile values, gestational curve charts and regression equations for both the mean and standard deviation.

When we assessed measurement differences from those studies that had the highest scores (Table 5), it was noted that, for the majority of biometric parameters, significant heterogeneity remained: for example, a BPD measurement of 88 mm at 36 weeks is around the 50th centile of one chart,[74] whereas the same fetus would be below the 10th centile by another,[92] and the same is seen at various other gestational ages. Differences in AC were similarly wide, with overlap between 10th and 50th centiles from different studies evident.[49,53] However, equivalent centiles of HC seem to be very similar between studies, with most differences well within 0.5 SD. This may be a reflection of the lower variability in HC measurements across different countries because of the properties of HC as a marker of fat-free mass, or may simply be caused by the fact that HC is a parameter that is more consistently measured in different locations. In other words, whether these differences in biometric parameters are a result of different measurement methodologies or differences in population characteristics is impossible to establish using this approach. It is only when studies of fetal size of the highest methodological quality are performed uniformly in different populations that differences in fetal size can be properly appraised or attributed to biological determinants.

Conclusion

This systematic review has demonstrated considerable heterogeneity of design in ultrasound studies of fetal biometry. The use of uniform methodology of the highest quality is essential in order to establish whether population differences in fetal measurements are biological or caused by differences in measurement. A checklist of the recommended design is proposed in order to aid such uniformity.

Disclosure of interests

CI, IS, EO, JV and ATP are part of the INTERGROWTH-21st project, an international study of fetal growth (www.intergrowth21.org.uk/).

Contribution to authorship

CI, KT and ATP designed the study, analysed the data, interpreted the results, drafted the manuscript and made the decision to submit. AC-A, JV and ATP defined the quality criteria *a priori*. CI, KT and EO extracted the data. All authors had full access to the data, interpreted the results, edited and approved the final manuscript.

Details of ethical approval

No ethics approval was required.

Funding

This project was supported by INTERGROWTH-21st (Grant ID# 49038) from the Bill & Melinda Gates Foundation to the University of Oxford, for which we are very grateful. CI and ATP are supported by the Oxford Partnership Comprehensive Biomedical Research Centre with funding from the Department of Health NIHR Biomedical Research Centres funding scheme.

Acknowledgements

We would like to thank Nia Roberts, outreach librarian at the Bodleian Health Care Libraries, for her assistance in the literature search.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Excluded studies.

Appendix S2. Risk of bias amongst included studies.

Appendix S3. Additional maternal and pregnancy characteristics.

Appendix S4. Additional study characteristics.

Appendix S5. Scatterplot of quality scores.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information

supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author. ■

References

1 Campbell S. An improved method of fetal cephalometry by ultrasound. *J Obstet Gynaecol Br Commonw* 1968;75:568–76.
2 Bricker L, Neilson JP, Dowswell T. Routine ultrasound in late pregnancy (after 24 weeks' gestation). *Cochrane Database Syst Rev* 2008 Oct 8;(4):CD001451.
3 Centre for Maternal and Child Enquiries (CMACE). *Perinatal Mortality 2009*. London: CMACE, 2011.
4 Raman S, Teoh T, Nagaraj S. Growth patterns of the humeral and femur length in a multiethnic population. *Int J Gynaecol Obstet* 1996;54:143–7.
5 Schwarzler P, Bland JM, Holden D, Campbell S, Ville Y. Sex-specific antenatal reference growth charts for uncomplicated singleton pregnancies at 15–40 weeks of gestation. *Ultrasound Obstet Gynecol* 2004;23:23–9.
6 Krampl E, Lees C, Bland JM, Espinoza Dorado J, Moscoso G, Campbell S. Fetal biometry at 4300 m compared to sea level in Peru. *Ultrasound Obstet Gynecol* 2000;16:9–18.
7 Silverwood RJ, Cole TJ. Statistical methods for constructing gestational age-related reference intervals and centile charts for fetal size. *Ultrasound Obstet Gynecol* 2007;29:6–13.
8 Royston P, Wright EM. How to construct 'normal ranges' for fetal variables. *Ultrasound Obstet Gynecol* 1998;11:30–8.
9 Altman DG, Chitty LS. Design and analysis of studies to derive charts of fetal size. *Ultrasound Obstet Gynecol* 1993;3:378–84.
10 Altman DG, Chitty LS. Charts of fetal size: 1. Methodology. *Br J Obstet Gynaecol* 1994;101:29–34.
11 Royston P, Altman DG. Design and analysis of longitudinal studies of fetal size. *Ultrasound Obstet Gynecol* 1995;6:307–12.
12 Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of observational studies in epidemiology (MOOSE) group. *J Am Med Assoc* 2000;283:2008–12.
13 Cole TJ. The LMS method for constructing normalized growth standards. *Eur J Clin Nutr* 1990;44:45–60.
14 Aickin DR, Chatfield WR, Rippon PE. Fetal cephalometry during normal pregnancy in a New Zealand population. *Aust N Z J Obstet Gynaecol* 1976;16:210–2.
15 Al-Meshari AA, Raber H. Fetal biparietal diameter in Saudi Arabia. *Ann Saudi Med* 1987;7:227–33.
16 Amoa AB, Wapi J, Klufio CA. Longitudinal fetal biometry of normal pregnant Melanesian Papua New Guineans to construct standards of reference for Papua New Guinea. *P N G Med J* 1993;36:219–27.
17 Ashrafunnessa, Jehan AH, Chowdhury SB, Sultana F, Haque JA, Khatun S, et al. Construction of fetal charts for biparietal diameter, fetal abdominal circumference and femur length in Bangladeshi population. *Bangladesh Med Res Counc Bull* 2003;29:67–77.
18 Ayangade SO, Okonofua FE. Normal growth of the fetal biparietal diameter in an African population. *Int J Gynaecol Obstet* 1986;24:35–42.
19 Beigi A, ZarrinKoub F. Ultrasound assessment of fetal biparietal diameter and femur length during normal pregnancy in Iranian women. *Int J Gynaecol Obstet* 2000;69:237–42.
20 Bergsjo P, Bakke T, Bjerkedal T. Growth of the fetal skull, with special reference to weight-for-dates of the newborn child. *Acta Obstet Gynecol Scand* 1976;55:53–7.
21 Brons JT, van Geijn HP, Bezemer PD, Nauta JP, Arts NF. The fetal skeleton; ultrasonographic evaluation of the normal growth. *Eur J Obstet Gynecol Reprod Biol* 1990;34:21–36.
22 Browne PC, Hamner LH III, Clark WS. Sonographic fetal growth curves from an indigent population in Atlanta, Georgia: I. Singleton pregnancies. *Am J Perinatol* 1992;9:467–76.
23 Campbell S, Newman GB. Growth of the fetal biparietal diameter during normal pregnancy. *J Obstet Gynaecol Br Commonw* 1971;78:513–9.
24 Chan WB, Yeo GS. A comparison of fetal biparietal diameter measurements between local Chinese and Caucasian populations. *Singapore Med J* 1991;32:214–7.
25 Chang CH, Chang FM, Yu CH, Yao BL, Liang RI, Ko HC. Re-analysis of fetal biparietal diameter during gestation in Taiwanese by Altman's method. *J Med Ultrasound* 1996;4:162–8.
26 Chitty LS, Altman DG, Henderson A, Campbell S. Charts of fetal size: 3. Abdominal measurements. *Br J Obstet Gynaecol* 1994;101:125–31.
27 Chitty LS, Altman DG, Henderson A, Campbell S. Charts of fetal size: 4. Femur length. *Br J Obstet Gynaecol* 1994;101:132–5.
28 Chitty LS, Altman DG, Henderson A, Campbell S. Charts of fetal size: 2. Head measurements. *Br J Obstet Gynaecol* 1994;101:35–43.
29 de la Vega A, Ruiz-Febo N, Roberts ZC. Fetal ultrasound biometry: normative charts for a Puerto Rican population. *P R Health Sci J* 2008;27:81–4.
30 Deter RL, Harrist RB, Hadlock FP, Carpenter RJ. Fetal head and abdominal circumferences: II. A critical re-evaluation of the relationship to menstrual age. *J Clin Ultrasound* 1982;10:365–72.
31 Deter RL, Harrist RB, Hadlock FP, Poindexter AN. Longitudinal studies of fetal growth with the use of dynamic image ultrasonography. *Am J Obstet Gynecol* 1982;143:545–54.
32 Deter RL, Rossavik IK, Hill RM, Cortissoz C, Hadlock FP. Longitudinal studies of femur growth in normal fetuses. *J Clin Ultrasound* 1987;15:299–305.
33 Di Battista E, Bertino E, Benso L, Fabris C, Aicardi G, Pagliano M, et al. Longitudinal distance standards of fetal growth. Intrauterine and Infant Longitudinal Growth Study: IILGS. *Acta Obstet Gynecol Scand* 2000;79:165–73.
34 Dubiel M, Krajewski M, Pietryga M, Tretyn A, Breborowicz G, Lindquist P, et al. Fetal biometry between 20–42 weeks of gestation for Polish population. *Ginekol Pol* 2008;79:746–53.
35 Elejalde BR, de Elejalde MM. The prenatal growth of the human body determined by the measurement of bones and organs by ultrasonography. *Am J Med Genet* 1986;24:575–98.
36 Eriksen PS, Secher NJ, Weis-Bentzon M. Normal growth of the fetal biparietal diameter and the abdominal diameter in a longitudinal study. An evaluation of the two parameters in predicting fetal weight. *Acta Obstet Gynecol Scand* 1985;64:65–70.
37 Exacoustos C, Rosati P, Rizzo G, Arduini D. Ultrasound measurements of fetal limb bones. *Ultrasound Obstet Gynecol* 1991;1:325–30.
38 Fescina RH, Ucieda FJ, Cordano MC, Nieto F, Tenzer SM, López R. Ultrasonic patterns of intrauterine fetal growth in a Latin American country. *Early Hum Dev* 1982;6:239–48.
39 Flamme P. Ultrasonic fetal cephalometry: percentiles curve. *Br Med J* 1972;3:384–5.
40 Gallivan S, Robson SC, Chang TC, Vaughan J, Spencer JA. An investigation of fetal growth using serial ultrasound data. *Ultrasound Obstet Gynecol* 1993;3:109–14.
41 Guihard-Costa AM, Larroche JC, Droulle P, Narcy F. Fetal biometry. Growth charts for practical use in fetopathology and antenatal ultrasonography. Introduction. *Fetal Diagn Ther* 1995;10:211–78.

42 Hadlock FP, Deter RL, Harrist RB, Park SK. Fetal head circumference: relation to menstrual age. *Am J Roentgenol* 1982;138:649–53.

43 Hadlock FP, Deter RL, Harrist RB, Park SK. Fetal abdominal circumference as a predictor of menstrual age. *Am J Roentgenol* 1982;139:367–70.

44 Hadlock FP, Deter RL, Harrist RB, Park SK. Fetal biparietal diameter: a critical re-evaluation of the relation to menstrual age by means of real-time ultrasound. *J Ultrasound Med* 1982;1:97–104.

45 Hadlock FP, Harrist RB, Deter RL, Park SK. Fetal femur length as a predictor of menstrual age: sonographically measured. *Am J Roentgenol* 1982;138:875–8.

46 Hoffbauer H, Pachaly J, Arabin B, Baumann ML. Control of fetal development with multiple ultrasonic body measures. *Contrib Gynecol Obstet* 1979;6:147–56.

47 Issel EP. Ultrasonic measurement of the growth of fetal limb bones in normal pregnancy. *J Perinat Med* 1985;13:305–13.

48 Issel EP, Prenzlau P, Bayer H, Lüder R, Schulte R, Wohlfahrth G, et al. The measurement of fetal growth during pregnancy by ultrasound (B-scan). *J Perinat Med* 1975;3:269–75.

49 Jeanty P, Cousaert E, Cantraine F. Normal growth of the abdominal perimeter. *Am J Perinatol* 1984;1:129–35.

50 Jeanty P, Cousaert E, Cantraine F, Hobbins JC, Tack B, Struyven J. A longitudinal study of fetal limb growth. *Am J Perinatol* 1984;1:136–44.

51 Jeanty P, Cousaert E, Hobbins JC, Tack B, Bracken M, Cantraine F. A longitudinal study of fetal head biometry. *Am J Perinatol* 1984;1:118–28.

52 Jeanty P, Kirkpatrick C, Dramaix-Wilmet M, Struyven J. Ultrasonic evaluation of fetal limb growth. *Radiology* 1981;140:165–8.

53 Johnsen SL, Wilsgaard T, Rasmussen S, Sollien R, Kiserud T. Longitudinal reference charts for growth of the fetal head, abdomen and femur. *Eur J Obstet Gynecol Reprod Biol* 2006;127:172–85.

54 Jung SI, Lee YH, Moon MH, Song MJ, Min JY, Kim JA, et al. Reference charts and equations of Korean fetal biometry. *Prenat Diagn* 2007;27:545–51.

55 Kurmanavicius J, Wright EM, Royston P, Wisser J, Huch R, Huch A, et al. Fetal ultrasound biometry: 1. Head reference values. *Br J Obstet Gynaecol* 1999;106:126–35.

56 Kurmanavicius J, Wright EM, Royston P, Zimmermann R, Huch R, Huch A, et al. Fetal ultrasound biometry: 2. Abdomen and femur length reference values. *Br J Obstet Gynaecol* 1999;106:136–43.

57 Lai FM, Yeo GS. Reference charts of foetal biometry in Asians. *Singapore Med J* 1995;36:628–36.

58 Larsen T, Petersen S, Greisen G, Larsen JF. Normal fetal growth evaluated by longitudinal ultrasound examinations. *Early Hum Dev* 1990;24:37–45.

59 Lei H, Wen SW. Ultrasonographic examination of intrauterine growth for multiple fetal dimensions in a Chinese population. Central-South China Fetal Growth Study Group. *Am J Obstet Gynecol* 1998;178:916–21.

60 Lessoway VA, Schulzer M, Wittmann BK, Gagnon FA, Wilson RD. Ultrasound fetal biometry charts for a North American Caucasian population. *J Clin Ultrasound* 1998;26:433–53.

61 Leung TN, Pang MW, Daljit SS, Leung TY, Poon CF, Wong SM, et al. Fetal biometry in ethnic Chinese: biparietal diameter, head circumference, abdominal circumference and femur length. *Ultrasound Obstet Gynecol* 2008;31:321–7.

62 Levi S, Erbsman F. Antenatal fetal growth from the nineteenth week. Ultrasonic study of 12 head and chest dimensions. *Am J Obstet Gynecol* 1975;121:262–8.

63 Levi S, Smets P. Intra-uterine fetal growth studied by ultrasonic biparietal measurements. The percentiles of biparietal distribution. *Acta Obstet Gynecol Scand* 1973;52:193–8.

64 Lu SC, Chang CH, Yu CH, Kang L, Tsai PY, Chang FM. Reappraisal of fetal abdominal circumference in an Asian population: analysis of 50,131 records. *Taiwan J Obstet Gynecol* 2008;47:49–56.

65 Mathai M, Thomas S, Peedicayil A, Regi A, Jasper P, Joseph R. Growth pattern of the Indian fetus. *Int J Gynaecol Obstet* 1995; 48:21–4.

66 Merz E, Kim-Kern MS, Pehl S. Ultrasonic mensuration of fetal limb bones in the second and third trimesters. [Erratum appears in J Clin Ultrasound 1987;15(9):704]. *J Clin Ultrasound* 1987;15:175–83.

67 Munjanja SP, Masona D, Masvikeni S. Fetal biparietal diameter and head circumference measurements: results of a longitudinal study in Zimbabwe. *Int J Gynaecol Obstet* 1988;26:223–8.

68 Munoz WP, Moore PJ, MacKinnon A, Haines LM. Biparietal diameter and menstrual age in the black population attending Edendale Hospital. *J Clin Ultrasound* 1986;14:681–8.

69 Nasrat H, Bondagji NS. Ultrasound biometry of Arabian fetuses. *Int J Gynaecol Obstet* 2005;88:173–8.

70 Neufeld LM, Haas JD, Grajeda R, Martorell R. Ultrasound measurement of fetal size in rural Guatemala. *Int J Gynaecol Obstet* 2004;84:220–8.

71 O'Brien GD, Queenan JT. Growth of the ultrasound fetal femur length during normal pregnancy. Part I. *Am J Obstet Gynecol* 1981;141:833–7.

72 Okonofua FE, Ayangade SO, Ajibulu OA. Ultrasound measurement of fetal abdominal circumference and the ratio of biparietal diameter to transverse abdominal diameter in a mixed Nigerian population. *Int J Gynaecol Obstet* 1988;27:1–6.

73 Okupe RF, Coker OO, Gbajumo SA. Assessment of fetal biparietal diameter during normal pregnancy by ultrasound in Nigerian women. *Br J Obstet Gynaecol* 1984;91:629–32.

74 Paladini D, Rustico M, Viora E, Giani U, Bruzzese D, Campogrande M, et al. Fetal size charts for the Italian population. Normative curves of head, abdomen and long bones. *Prenat Diagn* 2005;25:456–64.

75 Pang MW, Leung TN, Sahota DS, Lau TK, Chang AM. Customizing fetal biometric charts. *Ultrasound Obstet Gynecol* 2003;22:271–6.

76 Persson PH, Grennert L, Gennser G, Gullberg B. Normal range curves for the intrauterine growth of the biparietal diameter. *Acta Obstet Gynecol Scand Suppl* 1978;78:15–20.

77 Persson PH, Weldner BM. Normal range growth curves for fetal biparietal diameter, occipito frontal diameter, mean abdominal diameters and femur length. *Acta Obstet Gynecol Scand* 1986; 65:759–61.

78 Pineau JC, Guihard-Costa AM, Droulle P. Only two-phase models, computed independently for males and females, are appropriate to describe fetal head growth. *Fetal Diagn Ther* 2003;18:207–16.

79 Queenan JT, Kubarych SF, Cook LN, Anderson GD, Griffin LP. Diagnostic ultrasound for detection of intrauterine growth retardation. *Am J Obstet Gynecol* 1976;124:865–73.

80 Saksiriwuttho P, Ratanasiri T, Komwilaisak R. Fetal biometry charts for normal pregnant women in northeastern Thailand. *J Med Assoc Thai* 2007;90:1963–9.

81 Salomon LJ, Duyme M, Crequat J, Brodaty G, Talmant C, Fries N, et al. French fetal biometry: reference equations and comparison with other charts. *Ultrasound Obstet Gynecol* 2006;28:193–8.

82 Schluter PJ, Pritchard G, Gill MA. Ultrasonic fetal size measurements in Brisbane, Australia. *Australas Radiol* 2004;48:480–6.

83 Shohat T, Romano-Zelekha O. Ultrasonographic measurements of fetal femur length and biparietal diameter in an Israeli population. *Isr Med Assoc J* 2001;3:166–8.

84 Siwadune T, Sunsaneevithayakul P, Titapant V, Boriboonhirunsarn D, Kanokpongsakdi S. Charts of Thai fetal biometries: 2. Biparietal diameter. *J Med Assoc Thai* 2000;83:292–8.

85 Smulian JC, Ananth CV, Vintzileos AM, Guzman ER. Revisiting sonographic abdominal circumference measurements: a comparison of outer centiles with established nomograms. *Ultrasound Obstet Gynecol* 2001;18:237–43.

86 Snijders RJ, Nicolaides KH. Fetal biometry at 14–40 weeks' gestation. *Ultrasound Obstet Gynecol* 1994;4:34–48.

87 Sunsaneevithayakul P, Boriboonhirunsarn D, Siwadune T, Titapant V, Vantanasiri C. Charts of Thai fetal biometries: 4. Abdominal circumference. *J Med Assoc Thai* 2000;83:307–14.

88 Tamura RK, Sabbagha RE. Percentile ranks of sonar fetal abdominal circumference measurements. *Am J Obstet Gynecol* 1980;138:475–9.

89 Thame M, Osmond C, Fletcher H, Forrester TE. Ultrasound derived fetal growth curves for a Jamaican population. *West Indian Med J* 2003;52:99–110.

90 Titapant V, Siwadune T, Boriboonhirunsarn D, Sunsaneevithayakul P, Swasdimongkol S, Sutanthavibul A. Charts of Thai fetal biometries: 3. Femur length. *J Med Assoc Thai* 2000;83:299–306.

91 Todros T, Ferrazzi E, Groli C, Nicolini U, Parodi L, Pavoni M, et al. Fitting growth curves to head and abdomen measurements of the fetus: a multicentric study. *J Clin Ultrasound* 1987;15:95–105.

92 Verburg BO, Steegers EA, De Ridder M, Snijders RJ, Smith E, Hofman A, et al. New charts for ultrasound dating of pregnancy and assessment of fetal growth: longitudinal data from a population-based cohort study. *Ultrasound Obstet Gynecol* 2008;31:388–96.

93 Warda AH, Deter RL, Rossavik IK, Carpenter RJ, Hadlock FP. Fetal femur length: a critical reevaluation of the relationship to menstrual age. *Obstet Gynecol* 1985;66:69–75.

94 Watson DS. Biparietal diameter in the Australian aboriginal fetus. *Br J Obstet Gynaecol* 1986;93:339–42.

95 Westerway SC, Davison A, Cowell S. Ultrasonic fetal measurements: new Australian standards for the new millennium. *Aust N Z J Obstet Gynaecol* 2000;40:297–302.

96 Wladimiroff JW, Bloemsma CA, Wallenburg HC. Ultrasonic assessment of fetal head and body sizes in relation to normal and retarded fetal growth. *Am J Obstet Gynecol* 1978;131:857–60.

97 de Onis M, Garza C, Victora CG, Onyango AW, Frongillo EA, Martines J. The WHO Multicentre Growth Reference Study: planning, study design, and methodology. *Food Nutr Bull* 2004; 25:S15–26.

98 DiFranza JR, Lew RA. Effect of maternal cigarette smoking on pregnancy complications and sudden infant death syndrome. *J Fam Pract* 1995;40:385–94.

99 Hendrix N, Berghella V. Non-placental causes of intrauterine growth restriction. *Semin Perinatol* 2008;32:161–5.

100 Rasmussen S, Irgens LM. The effects of smoking and hypertensive disorders on fetal growth. *BMC Pregnancy Childbirth* 2006;6:16.

101 Zeitlin J, Ancel PY, Saurel-Cubizolles MJ, Papiernik E. The relationship between intrauterine growth restriction and preterm delivery: an empirical approach using data from a European case–control study. *BJOG* 2000;107:750–8.

102 Stacey T, Thompson JM, Mitchell EA, Ekeroma AJ, Zuccollo JM, McCowan LM. The Auckland Stillbirth study, a case–control study exploring modifiable risk factors for third trimester stillbirth: methods and rationale. *Aust N Z J Obstet Gynaecol* 2011;51: 3–8.

103 Rasmussen S. Charts to assess fetal wellbeing. *Ultrasound Obstet Gynecol* 2011;37:2–5.

104 Gardosi J. Customized growth curves. *Clin Obstet Gynecol* 1997;40:715–22.

105 Gardosi J. Dating of pregnancy: time to forget the last menstrual period. *Ultrasound Obstet Gynecol* 1997;9:367–8.

106 Mongelli M, Wilcox M, Gardosi J. Estimating the date of confinement: ultrasonographic biometry versus certain menstrual dates. *Am J Obstet Gynecol* 1996;174:278–81.

107 Salomon LJ. Early fetal growth: concepts and pitfalls. *Ultrasound Obstet Gynecol* 2010;35:385–9.

108 Sarris I, Ioannou C, Dighe M, Mitidieri A, Oberto M, Qingqing W, et al. Standardization of fetal ultrasound biometry measurements: improving the quality and consistency of measurements. *Ultrasound Obstet Gynecol* 2011;38:681–7.

