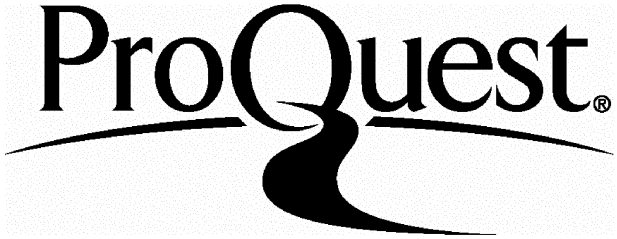
DIAGNOSIS OF GROWTH RETARDATION IN SMALL FETUSES: SERIAL ULTRASOUND ASSESSMENT OF ABDOMINAL CIRCUMFERENCE AND FETAL WEIGHT

A thesis presented for the Degree of MD in the
Faculty of Medicine of the University of London by

Tou Choong CHANG

DEPARTMENT OF OBSTETRICS AND GYNAECOLOGY
UNIVERSITY COLLEGE AND MIDDLESEX SCHOOL OF MEDICINE
LONDON WC1E 6HX

1993



ABSTRACT

Aim

The aim of this work was to evaluate the ability of serial ultrasound measurements of abdominal circumference (AC) and estimates of fetal weight to diagnose intrauterine growth retardation (IUGR) in a group of small fetuses delivered at term.

Factors influencing study design

1. The interpretation of serial ultrasound values of AC and estimated fetal weight (EFW) required appropriately derived reference standards for these measurements.
2. To determine the minimum interval between ultrasound assessments, AC and EFW were subjected to tests of reproducibility.
3. As there was no accepted method of quantifying serial values of AC and EFW, different statistical methods of describing serial data were tested against neonatal morphometric indices of "wasting", the "gold standard" used to define IUGR.

Methods

1. To construct reference standards for AC and EFW, four different classes of formulae were fitted to the longitudinal data of 67 normal fetuses by least squares fitting.
2. The intra-observer and inter-observer reproducibility of AC and EFW were evaluated using one-way analysis of variance and limits of agreement respectively.
3. Three different methods of quantifying serial ultrasound data were evaluated in their ability to predict a reduced neonatal ponderal index, mid-arm circumference / head circumference ratio and subscapular and triceps skinfold thickness in 104 small fetuses (defined as an AC < 10th centile in the third trimester of pregnancy). These different statistical measures were compared using receiver operating characteristic (ROC) curves.
4. The best method of quantifying serial values of AC and EFW was compared with umbilical artery pulsatility index (PI) and single estimates of fetal size in the prediction of abnormal neonatal morphometry using ROC curves.
5. The study group was divided by their antenatal growth profile into those who were normally grown and those with IUGR. Perinatal morbidity and biochemical indices of IUGR were then compared in these two groups.

Results

1.Reference standards for AC and EFW were constructed using a log quadratic formula; $\log_{10}(AC) = a + b\,(GA) + c\,(GA)^2$. The fitting of this formula to the longitudinal AC (and EFW) data for each fetus resulted in the smallest residual fitting errors and did not systematically over- or under-estimate the final values of $\log_{10}(AC)$ [or $\log_{10}(EFW)$].
2. The tests of reproducibility suggested a minimum interval between ultrasound

assessments of two weeks.

3. The best method of quantifying serial values of AC and EFW was found to be a change in standard deviation scores (Δ SDS) between the first and last ultrasound assessment. The SDS for AC (or EFW) at each ultrasound assessment was calculated as follows: SDS = (Measured AC - Mean AC) / Standard deviation of AC.

4. This method of quantifying serial AC and EFW data was superior to umbilical artery PI and the last estimate of fetal size prior to delivery in the prediction of abnormal neonatal morphometry and perinatal morbidity associated with IUGR.

5. Separation of small fetuses on the basis of their Δ SDS values resulted in two groups with some differences in perinatal morbidity and biochemical indices of IUGR.

Conclusion

In a group of small fetuses, serial ultrasound assessment of AC and EFW predicted subsequent neonatal morphometry indicative of IUGR. The method of quantifying fetal growth described in this study was useful in separating small fetuses into those with IUGR and those with normal growth.

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AC	Abdominal circumference
Δ AC.SDS	Change in the standard deviation score of abdominal circumference
AGA	Appropriate for gestational age
BE	Base Excess
BMUS	British Medical Ultrasound Society
BPD	Biparietal diameter
CI	Confidence intervals
cm	Centimetres
CTG	Cardiotocography
CW	Continuous wave
dB	Decibels
EFW	Estimated fetal weight
Δ EFW.SDS	Change in the standard deviation score of estimated fetal weight
FDA	Food and Drug Administration
FGV	Fetal growth velocity
FL	Femur length
FL / AC	Femur length / Abdominal circumference
FPR	False positive rate
g	Grams
GA	Gestational age
HC	Head circumference
HC / AC	Head circumference / Abdominal circumference
IGF	Insulin-like growth factor
I - m	Maximum intensity
I - SPPA	Spatial Peak Pulse Average Intensity
I - SPTA	Spatial Peak Time Averaged Intensity
IUGR	Intrauterine growth retardation
kg	Kilograms
KHz	KiloHertz
l	Litre
MAC / HC	Mid-arm circumference / Head circumference
MHz	MegaHertz
m	Metres
ml	Millilitres
mm	Millimetres
mmol	Millimoles
mW	MilliWatts

n	Number
NICU	Neonatal intensive care unit
NS	Not significant
OAPR	Odds of being affected given a positive result
OR	Odds ratio
PI	Pulsatility index
pmol	Picomoles
PPV	Positive predictive value
PW	Pulsed wave
ROC	Receiver operating characteristic
SD	Standard deviation
S / D	Systolic / Diastolic ratio
SDS	Standard deviation score
Δ SDS	Change in standard deviation score
SE	Standard error
sec	Second
SGA	Small for gestational age
U	Units
W	Watts

CHAPTER 1

INTRODUCTION

The antenatal identification of smallness for gestational age (SGA) at birth (birthweight < 10th centile for gestational age) is an important aspect of obstetric care in view of the associated increased risk of perinatal morbidity and mortality (Dobson et al. 1982, Teberg et al. 1982, Steer 1989), as well as adverse neurodevelopmental outcome (Dijxhoorn et al. 1987, Stewart 1989). Many previous studies have used SGA to define intrauterine growth retardation (IUGR) (Villar and Belizan 1986, Benson et al. 1986). However, this is inappropriate as many SGA neonates are constitutionally small but not growth retarded (Roord and Raemaker 1979, Walther and Raemaker 1982, Patterson and Pouliot 1987a, Fay et al. 1991a). Conversely, many neonates with a birthweight > 10th centile for gestational age have morphometric features of "wasting", indicative of IUGR (Hill et al. 1984, Georgieff et al. 1988, Sumners et al. 1990). As will be evident from the review in Section 2.2, neonatal morphometric indices of malnutrition such as the ponderal index, mid-arm circumference / head circumference ratio (MAC / HC) and skinfold thickness measurements are superior to birthweight for the purposes of defining IUGR (Georgieff et al. 1986, Haas et al. 1987) . Studies have shown that SGA neonates with morphometric evidence of IUGR are at greater risk of perinatal morbidity and mortality than SGA neonates with no evidence of "wasting" (Patterson and Pouliot 1987a, Villar et al. 1990, Fay et al. 1991a). Therefore, the antenatal identification of small fetuses with growth retardation is important in order to concentrate increased surveillance in this sub-group of fetuses.

Whilst clinical risk factors (Galbraith et al. 1979, Wennergren et al. 1982) and clinical assessment of uterine size (Quaranta et al. 1981, Secher et al. 1990) have been used to identify IUGR antenatally, ultrasound has generally been regarded as the best method of assessing intrauterine growth. However, it will be evident from the review of literature in Section 2.3 that many ultrasound parameters purportedly reported to be useful in the diagnosis of IUGR have in fact been evaluated in their ability to predict SGA. There are comparatively few studies evaluating the ability of ultrasound to predict neonatal morphometry indicative of IUGR. As abdominal circumference (AC) and estimated fetal weight (EFW) have been shown to be the best ultrasound parameters for predicting SGA (Chang et al. 1992), serial ultrasound assessment of AC and EFW may be useful in the identification of small fetuses with IUGR. Fetal growth is a dynamic process of progressive increase in fetal anthropometry; serial ultrasound assessment of fetal size allows this growth to be quantified (Altman and Hytten 1989, Deter et al. 1990, Deter and Harrist 1992). However, despite their widespread use in clinical practice, serial values of AC and EFW have not been critically assessed in their ability to separate small fetuses into those with growth failure and those with normal growth.

A study was therefore constructed to address this problem. Serial ultrasound values of AC and EFW were evaluated in their ability to predict neonatal morphometry

indicative of IUGR and perinatal morbidity associated with IUGR in a group of small fetuses delivered after 36 weeks gestation. Only pregnancies which ended after 36 weeks gestation were considered to exclude the confounding effects of prematurity on perinatal morbidity. In order to address this problem, satisfactory reference ranges for AC and EFW were necessary to provide standards for evaluating normal growth. Accepted methods of quantifying serial measurements of fetal size were also required to describe fetal growth. However, as will be evident from Section 3.1.2, all previously reported reference standards for AC and EFW were inappropriate for the evaluation of fetal growth. This has also been noted by other workers (Evans et al. 1990). Furthermore, although the quantification of serial values of AC and EFW has been described by Deter and co-workers (Deter et al. 1984, Deter and Rossavik 1987, Deter et al. 1988, Deter et al. 1989b, Deter et al. 1990, Deter and Harrist 1992), these have never been validated against standard neonatal morphometry indicative of IUGR (Section 3.1.3).

It was therefore necessary to construct reference standards for AC and EFW for the evaluation of fetal growth (Chapter 5). These reference standards were derived from longitudinally collected data which are more likely to represent the normal growth trajectory of fetuses than cross-sectionally collected data (Deter et al. 1982, Evans et al. 1990, Sparks and Cetin 1991). These ultrasound parameters were also subjected to tests of reproducibility so that inferences about fetal growth could be made in the light of the error of these measurements (Chapter 6). Different statistical methods of quantifying serial ultrasound measurements of AC and estimates of fetal weight were validated against ponderal index, MAC / HC ratio and skinfold thickness at birth (Section 7.4.1).

The usefulness of serial ultrasound values of AC and EFW in the diagnosis of IUGR was further evaluated by comparison with other ultrasound parameters currently used to evaluate fetal growth, such as umbilical artery Doppler waveform indices and single estimates of fetal size (Section 7.4.2). The clinical implications of separating small fetuses into those with ultrasound evidence of IUGR or normal growth were determined by comparing measures of perinatal morbidity and biochemical indices of IUGR in the resultant two groups (Section 7.4.3).

This study addresses the problem of identifying a sub-group of small fetuses with growth retardation by serial ultrasound. It evaluates critically the deficiencies of previous methods of quantifying serial ultrasound measurements, suggests a new method of quantifying serial measurements irrespective of gestational age or interval between scans, and verifies its usefulness by validating it against neonatal morphometry, the "gold standard" for defining IUGR. The clinical implications of this sub-division of small fetuses, as well as future research emanating from this work, are discussed.

CHAPTER 2

REVIEW OF LITERATURE

2.1 Introduction

The usefulness of any ultrasound parameter in the diagnosis of IUGR is evaluated by testing its ability to predict a "gold standard". However, there is little agreement in the obstetric literature as to which gold standard to use to define IUGR. Smallness for gestational age, neonatal morphometric indices which reflect "wasting" in the newborn, perinatal morbidity associated with IUGR and biochemical indices of malnutrition have all been used to define IUGR. The advantages and disadvantages of using each of these outcome measures for the purposes of defining IUGR are discussed in Section 2.2.

Numerous studies have reported the use of different ultrasound parameters in the diagnosis of IUGR. These are reviewed in Section 2.3. However, these ultrasound studies varied tremendously in the choice of outcome criteria used to define IUGR. In view of the heterogeneity of these studies, ultrasound studies were grouped together according to the outcome criteria used.

For any given outcome criteria used to define IUGR, the usefulness of any ultrasound parameter has usually been assessed by constructing contingency (two-by-two) tables from the data and calculating sensitivities, specificities, positive and negative predictive values. The limitations of this method of analysis will be discussed. The results of studies which report data using the same ultrasound parameters and outcome criteria will be grouped and summary statistics presented where appropriate.

2.2 Definition of Intrauterine Growth Retardation (IUGR)

2.2.1 Smallness-for-gestational age

2.2.1.1 Use of birthweight standards

The National Institute of Child Health and Human Development in the United States of America defined IUGR as follows: "Both for clinical and research purposes, IUGR birth should be defined as resulting in a birthweight less than the 10th percentile for gestational age, using criteria appropriate to the population under study" (Miller and Merritt 1979). As a result of this, many previous publications have used SGA (birthweight < 10th centile for gestational age) to define IUGR (Villar and Belizan 1986, Benson et al. 1986). The inherent assumption here is that a small fetus / SGA neonate has had a decrease in growth rate in-utero which ultimately results in low birthweight.

Different birthweight standards have been reported for the classification of neonates into those who are SGA or appropriate for gestational age (AGA) (birthweight > 10th centile for gestational age). In the United Kingdom, the most commonly used reference standards for the classification of birthweight are those reported by Thomson et al. (1968), Neligan (1974), Gairdner and Pearson (1985) and Yudkin et al. (1987). The obvious advantage with the use of birthweight to define IUGR is that it is an easily definable end-point measured routinely at all deliveries.

2.2.1.2 Limitations with the use of birthweight

1. Although a proportion of SGA infants are growth-retarded, the majority exhibit no symptoms or signs associated with IUGR (Hill et al. 1984). These infants have no subcutaneous fat or muscle wasting, are well-proportioned, and may be constitutionally small (Walther and Raemaker 1982). Conversely, there are newborns who, despite having a birthweight greater than the 10th centile, are growth-retarded, with wasting and reduced subcutaneous fat deposition (Patterson and Pouliot 1987a, Villar et al. 1990). Therefore, the use of SGA to define IUGR clearly has marked limitations (Altman and Hytten 1989). Its use would result in unacceptably large numbers of false positives and false negatives when used to define IUGR.

2. Three studies have shown that the use of a birthweight < 10th centile to define IUGR identified only a small proportion of all babies with perinatal morbidity associated with IUGR. In a study of 355 patients, birthweight was compared with other neonatal morphometric indices in the prediction of significant perinatal morbidity, defined as

operative delivery for fetal distress, 5-minute Apgar score < 7, meconium aspiration, polycythaemia or hypoglycaemia (Patterson and Pouliot 1987a). Of the 33 neonates with significant perinatal morbidity, only 5 (15.2%) had a birthweight < 10th centile for gestational age. In contrast, 18 (54.5%) had either a low ponderal index or MAC / HC ratio.

In another study, a retrospective analysis of 44,830 patients over a gestational age range from 28 to 41 weeks showed that birthweight <10th centile for gestational age was a poor predictor of abnormal perinatal outcome at all gestational ages (Patterson et al. 1986). A low birthweight predicted abnormal perinatal outcome in 29, 39 and 36% of cases at 30, 36 and 38 weeks respectively. Whilst the poor predictiveness of the low birthweight before 34 weeks gestation may be explained by the confounding effects of prematurity, a low prevalence (3.5%) of poor perinatal outcome after 34 weeks may partly account for the low predictiveness of birthweight after this gestational age. Nevertheless, the study highlighted the limitations of birthweight centiles in the prediction of poor perinatal outcome at all gestations.

A third study analysed the ability of low birthweight to predict poor perinatal outcome in 2314 consecutive births greater than 37 weeks gestation (Fay et al. 1991a). Of the 77 births complicated by poor perinatal outcome (defined as operative delivery for fetal distress or 5-minute Apgar score < 7), only 23 (29.9%) had a birthweight less than the 10th centile. In contrast, 35 (45.5%) of these infants had a low ponderal index. The above studies show that birthweight is an imprecise end-point for the definition of IUGR.

3. A variety of birthweight standards have been reported for the purposes of defining IUGR in previous studies. Many of the birthweight charts used were derived from another totally different reference population. For example,the Denver charts in the U.S.A. (Lubchenco et al. 1963) were derived from a population about 5000 feet above sea-level and comprised many underprivileged people of mixed racial origin - yet these charts have been used indiscriminately in many studies for the definition of IUGR. This is inappropriate as it leads to misclassification of SGA at birth.

4. Published reference ranges for birthweight have reported differing tenth centile values. This was highlighted by Goldenberg et al. (1989) who reviewed thirteen studies which reported reference ranges for birthweight at different gestational ages. For example, at 40 weeks gestation, the 10th centile values reported by Hardy et al. (1979) and Miller and Merritt (1979) were 2604 g and 3050 g respectively, a difference of 446 g. An infant classified SGA by one birthweight standard would have been classified AGA according to another birthweight standard.

5. Studies which reported reference standards for birthweight differed in the assessment of gestational age, inclusion criteria and whether the studies were hospital or population based (Goldenberg et al. 1989). Although congenital anomalies should be excluded from data for constructing normal reference standards, only 3 studies allowed for this (Lubchenco et al. 1963, Thomson et al. 1968, Miller and Merritt 1979). Kiely et al. (1992) also showed that the tenth centile of a reference United States population increased by 90 to 180 grams compared with reference standards derived 18 years previously. These differences, both in the methodology of the studies and population differences, suggest that birthweight standards have to be critically assessed prior to their use to define SGA.

2.2.2 Neonatal morphometric indices

2.2.2.1 Introduction

The disadvantages of using birthweight to define IUGR have been enumerated in Section 2.2.1. In view of this, neonatal morphometric indices of wasting or malnutrition in the newborn have been suggested as alternatives to birthweight for the definition of IUGR.

The following neonatal morphometric indices have all been reported to be useful for the purposes of defining IUGR and each will be critically reviewed: ponderal index (Miller and Hassanein 1971), MAC / HC ratio (Sasanow et al. 1986), skinfold thickness measurements (Oakley et al. 1977b), body mass index (Wolfe et al. 1990), weight / length ratio (Wolfe et al. 1990) and estimations of percentage body fat (Dauncey et al. 1977).

2.2.2.2 Ponderal index

The ponderal index, an assessment of the amount of body mass, is calculated as follows: Ponderal index = weight / length^3 (g / cm^3) x 100. Infants who are proportionately small, either because of genetic or constitutional reasons, will have a normal ponderal index whilst those with reduced fat deposition will have a reduced ponderal index (Walther and Raemaker 1982). This index is therefore a method of assessing the relative fatness or thinness of an infant in relation to its length. Reference ranges for the ponderal index at various gestational ages have been published (Miller and Hassanein 1971, Meadows et al. 1986, Sarmandal and Grant 1990, Fay et al. 1991a); there appears to be a gradual increase in the ponderal index from 30 to 37 weeks, with no significant increase after 37 weeks. A tenth centile value of 2.32 after 37 weeks gestation has commonly been used to define IUGR in many studies evaluating the use of ultrasound in the diagnosis of IUGR (Ott 1985, Vintzileos et al. 1986, Patterson and Pouliot 1987a, Beattie and Dornan 1989, Sijmons et al. 1989, Weiner and Robinson 1989). Differences in the tenth centile values between different reference ranges are minimal (Fay et al. 1991a). The additional advantage of the ponderal index is that it is independent of race, parity and fetal sex (Miller and Hassanein 1971).

At least six studies have reported the usefulness of the neonatal ponderal index in predicting perinatal complications associated with IUGR. The two studies reviewed in Section 2.2.1 (Patterson and Pouliot 1987a and Fay et al. 1991a) showed that neonates with low ponderal indices were more likely to have had fetal distress in labour and a low Apgar score. Jarai et al. (1977) evaluated the relationship between ponderal index and

hypoglycaemia in 233 neonates admitted to the neonatal unit. Of the 42 babies who were hypoglycaemic, none had a ponderal index > 2.4 and 66.7% had a ponderal index < 2.0. This contrasts with 15% of normoglycaemic infants with a ponderal index > 2.4, and 38.2% of normoglycaemic infants with a ponderal index < 2.0. The reliability of ponderal index in the prediction of subsequent hypoglycaemia is consistent with findings that babies with a low ponderal index have insufficient hepatic glycogen stores and subcutaneous fat reserves, thereby reducing the gluconeogenic response to hypoglycaemia (Lubchenco and Bard 1971). In a study of 500 neonates, 80 neonates with a ponderal index < 10th centile for gestational age were more likely to be asphyxiated, acidotic, hypoglycaemic, hypothermic and polycythaemic compared with infants with a normal ponderal index (Walther and Raemaker 1982). In contrast, division of the same group of infants into SGA and AGA did not reveal any significant differences in the incidence of asphyxia or acidosis between the two groups.

Two large population studies (Haas et al. 1987, Villar et al. 1990) provided further evidence for the close association between ponderal index and morbidity and mortality associated with IUGR. These studies were particularly important as they investigated the ability of the ponderal index to separate SGA neonates into those with and without perinatal morbidity. In an epidemiological study of more than 12,000 infants in Mexico, Haas et al. (1987) reported a strong association between ponderal index and perinatal mortality. In this study, SGA and AGA neonates were divided into those with a low or a normal ponderal index. Those SGA neonates with a normal ponderal index had twice the mortality of AGA infants with a normal ponderal index. Other SGA infants with a low ponderal index had a 5.7 fold increased risk of perinatal mortality compared with SGA neonates with no morphometric evidence of IUGR. In a recent prospective follow-up study of 16,850 Guatemalan infants, Villar et al. (1990) demonstrated the independent effect of a low ponderal index on neonatal morbidity among SGA infants. Among the 4422 SGA infants, those with a low ponderal index were more likely to have low Apgar scores, aspiration syndrome, hypoglycaemia, fetal distress, risk of infection and longer hospital stays than SGA infants with a normal ponderal index. Among the 12,428 AGA infants, similar findings were noted when comparing the relative neonatal morbidities in infants with normal or low ponderal indices.

The simplicity of calculation, availability of reference ranges, and the many studies which have demonstrated the close correlation between ponderal index and perinatal morbidity have led to a recent increase in use of the ponderal index for the purposes of defining IUGR.

2.2.2.3 Mid-arm circumference / head circumference ratio

Another morphometric measure that has been used to define IUGR at birth is the MAC / HC ratio. This ratio relies on the principle that weight (muscle and fat mass) is lost in preference to other morphometric measurements during periods of starvation and malnutrition (Georgieff et al. 1986). It also relies on the principle that head-sparing occurs in IUGR; the MAC / HC ratio therefore accentuates a decreased muscle and fat mass in the mid-arm when compared with the head circumference. Kanawati et al. (1970) had previously shown that such a ratio was useful in the diagnosis of malnutrition in infants greater than 3 months of age. Reference ranges for the MAC / HC ratio in the second half of pregnancy have been reported (Sasanow et al. 1986). The ratio increases steadily with gestational age, reflecting the increased deposition in fat stores in the upper arm with gestational age.

The use of the MAC / HC ratio to define IUGR has been assessed in five studies. Georgieff et al. (1984) first reported a correlation between a low MAC / HC ratio and perinatal complications of IUGR in both AGA or SGA infants. In a subsequent study involving 73 neonates, this ratio was compared with birthweight in its ability to identify perinatal complications associated with IUGR (Georgieff et al. 1986). Infants were classified into those who were SGA or AGA, and perinatal morbidity was defined as the presence of hypoglycaemia, polycythaemia or hypocalcaemia. Fourteen of the 17 SGA infants with perinatal morbidity had a MAC / HC ratio more than 2 standard deviations (SD) below the mean. None of the asymptomatic infants had an abnormal MAC / HC ratio. The incidences of hypoglycaemia and polycythaemia in AGA neonates with a low MAC / HC ratio were similar to those in SGA neonates, further confirming the MAC / HC ratio to be superior to birthweight in the prediction of perinatal morbidity associated with IUGR. In two other studies which involved 96 and 64 neonates respectively, the MAC / HC ratio was found to be more discriminatory than birthweight in identifying neonates who would subsequently develop symptoms associated with IUGR (Excler et al. 1985, Meadows et al. 1986). The latter study also demonstrated no inter-racial differences and therefore advocated the use of a single reference standard for neonates of all racial origins.

Whilst Patterson and Pouliot (1987a) reported the MAC / HC ratio to be as good as the ponderal index, Georgieff et al. (1988) reported the MAC / HC ratio to be superior to the ponderal index in the prediction of perinatal morbidity associated with IUGR. In the latter study involving 60 infants, the MAC / HC ratio identified a higher percentage of neonates with symptoms associated with IUGR compared with the ponderal index (77% vs. 57%). The above studies suggest that the MAC / HC ratio is an alternative to the ponderal index for the purposes of defining IUGR.

2.2.2.4 Skinfold thickness measurements

Estimation of the fat content of the fetus has been shown in studies on body composition to be important in the recognition of impaired fetal growth (Usher et al. 1970, Widdowson 1971, Brans et al. 1975). This is especially pertinent in the third trimester as the rate of fat deposition in the last 3 months of pregnancy is much greater than the first 6 months of pregnancy. Widdowson et al. (1979) showed by dissection studies that total body fat increased from 49 g (4% of body weight) at 28 weeks to 476 g (14% of body weight) at 40 weeks.

Estimation of skinfold thickness has been advocated as the most appropriate non-invasive method of quantifying neonatal fat stores for the purposes of defining IUGR (Whitelaw et al. 1979). This is based on some classic studies describing the distribution of fat in the neonate (Forbes et al. 1962, Usher et al. 1970, Brans et al. 1974). Two recent studies have demonstrated the usefulness of skinfold thickness measurements in accurately assessing the amount of body fat in the neonate. Petersen et al. (1988) demonstrated the close inverse relationship between lean body mass, calculated using dual photon absorptiometry, and skinfold thickness as reflecting fat stores in the body. In the well-nourished infant, the mean value for lean body mass was 87% and the total skinfold thickness was 21.9 mm. This contrasted with a lean body mass of 98% and total skinfold thickness of 16.5 mm in the poorly nourished group. The difference in body fat content in the malnourished infant, a tenth of the value in the well-nourished infant, was therefore accurately reflected in the diminution in skinfold thickness. In another study involving 16 infants, the percentage body fat of the neonate was estimated by total body electrical conductivity (Cochran et al. 1986). The sum of subscapular and triceps skinfold thickness measurements were found to have a strong linear correlation (correlation coefficient = 0.82) with percentage body fat as estimated by this method.

Subscapular and triceps skinfold thicknesses are commonly used to quantify neonatal fat. Reference ranges for triceps and subscapular skinfold thickness have been constructed from the data of 1293 Caucasian infants born after 37 weeks of gestation (Oakley et al. 1977b). These standards were used in a prospective study in the prediction of hypoglycaemia (Oakley et al. 1977a). In a study of 100 fetuses, the sum of the triceps and subscapular skinfold thicknesses showed a close inverse relationship with the degree of hypoglycaemia. The usefulness of skinfold thickness in the prediction of perinatal morbidity associated with IUGR was further assessed in a recent study of 53 neonates (Sumners et al. 1990). Triceps, subscapular and quadriceps skinfold thicknesses correlated well with the risk of hypothermia, a known complication of IUGR. Six out of 10 babies with a triceps skinfold thickness less than the 3rd centile, and 5 out of 6 babies with a subscapular thickness less than the 3rd centile suffered from hypothermia; in

contrast, only one baby with a skinfold thickness more than the 3rd centile required attention for hypothermia.

The advantages with the use of skinfold thickness measurements to define IUGR in the newborn are its simplicity and ease of use, and the existence of reference standards derived from a large cohort of neonates.

2.2.2.5 Other morphometric measurements

Other morphometric measures of weight / length ratio (Hill et al. 1984), weight / length^2 ratio (Wolfe et al. 1990) and estimation of percentage body fat (Dauncey et al. 1977) have also been used to define IUGR at birth.

To overcome the disadvantage of compounding the error of length measurements by the cube power in the calculation of ponderal index, other simpler weight to length ratios such as weight / length^2 and weight / length have been advocated (Wolfe et al 1990). In a study of 119 term infants, all three weight to length ratios were correlated with skinfold thickness measurements (Wolfe et al. 1990). The correlation of these ratios with skinfold thickness measurements decreased as the length was squared or cubed. The weight / length ratio accounted for 52% of the variance in skinfold measurements and cubing the crown-heel length complicated the calculation with no apparent improvement in the prediction of neonatal body fat. Similar findings were reported in another study where 46 infants were separated at birth into those who were clinically well-nourished (n = 13) and those who were malnourished (n = 33), using a clinical subjective assessment of the amount of subcutaneous fat (Hill et al. 1984). The clinically well-nourished infant was identified as having an abundant amount of subcutaneous fat especially in the cheeks, lateral abdominal wall and in the arms, legs, thighs and buttocks. The use of the ponderal index, weight / length^2 and the weight / length ratio resulted in the misclassification of malnourished infants in 30.4, 18.0 and 8.7% of cases respectively. Another potential advantage of the weight / length ratio is that it is independent of gestational age (Miller and Hassanein 1971). Further evidence for this was provided in a recent study of 12,238 births more than 36 weeks gestation where weight / length ratio, birthweight and weight / length ratio adjusted for gestational age were compared in their ability to predict adverse neonatal outcome (Wolfe et al. 1992). The weight / length ratio was the best predictor of hypoglycaemia, low 5-minute Apgar score and polycythaemia. Adjustment for gestational age added little to the predictive ability of the weight / length ratio.

Estimation of fat content using measures other than skinfold thickness have also

been used to assess the degree of growth retardation in the neonate. Although estimations of body fat can be accurately determined using techniques such as hydrostatic weighing, measurements of total body water by deuterium oxide, measurement of total body potassium with ^{40}K, xenon or dual photon absorption, these methods are of limited value because they cannot be used routinely on newborns. Dauncey et al. (1977) reported a formula for estimating fat content using a combinations of skinfold thickness and nine other body dimensions. This theoretical model is based on the three assumptions: firstly, that 80% of total adipose tissue is in the subcutaneous tissue. Secondly, a constant ratio exists between subcutaneous fat and internal body fat, and thirdly the fat distribution throughout the body is constant. It is also highly reliant on the skinfold measurements. This method of calculating percentage body fat has been used in clinical studies to estimate percentage body fat (Clapp et al. 1990, Catalano et al. 1992).

2.2.2.6 Limitations of neonatal morphometric indices

1. There are reservations about the use of ponderal index to define IUGR in routine clinical practice because of concerns about errors when making measurements. The very formula for its derivation requires the length to be cubed. If length measurements in newborns are not made accurately, any measurement errors will be compounded by the need to cube it for use in the formula. A specially designed infantometer is essential for measuring the crown-heel length to reduce inaccuracies (Miller and Hassanein 1971, Colley et al. 1991) and the tonic neck reflex manoeuvre (Miller and Hassanein 1971) employed to ensure that the knees and hips are extended during measurement. As only a few studies have reported data on the reproducibility of the ponderal index, this would need to be investigated further.

2. A disadvantage with the use of the MAC / HC ratio is the assumption that head growth continues at a normal rate in IUGR due to the "brain-sparing" effect. This latter phenomenon is related to the observation that blood supply to the brain is maintained during hypoxia related to IUGR in preference to other visceral organs (Peeters et al. 1979). The assumption that head growth is totally unaffected by the process of IUGR may not be appropriate as some anthropometric studies have shown that growth of head circumference is still affected, albeit to a lesser extent than other parts of the body, in IUGR (Crane and Kopta 1980, Kramer et al. 1989, Colley et al. 1991).

3. Not all studies agree that skinfold thickness measurements accurately reflect fat content in the neonate. In a study of 48 infants at 5 weeks of age, total body water and fat-free mass were determined by ascertaining the heavy water ($H_2{}^{18}O$) content (Davies and Lucas 1990). The percentage body fat calculated by this method correlated poorly with skinfold thickness measurements. This poor correlation was not related to poor reproducibility of skinfold measurements. The authors attributed the poor correlation to the possibility that skinfold thickness, a measure of external body fat, may not accurately reflect internal body fat content. These results contrast with those reported by Cochran et al. (1986) and Petersen et al. (1988). The different methods of quantifying internal body fat content used in these different studies may possibly have accounted for some of these discrepancies.

4. Another potential limitation with the use of skinfold thickness to define IUGR is the different methodology of measurement described in the literature. Different reference standards have been reported using different methodology. In the derivation of the British reference standards (Oakley et al. 1977b, Whitelaw 1979), measurements were made once the caliper reading were stable. Other reference standards were derived using measurements made 60 seconds after application of calipers (Brans et al. 1974). This was based on the rationale that oedema in the neonate would otherwise have led to a falsely

high estimate of subcutaneous fat content. These different reference standards are therefore not comparable and strict adherence to the methodology is necessary in the use of such standards.

5. Different sites have also been used to measure skinfold thickness. The choice of the most appropriate site presents a problem: in adults, skinfold measurements in the thigh area correlate most strongly with total body fat (Lohman et al. 1981). No comparable study has been undertaken in neonates to determine the site which most accurately reflects total body fat and which therefore should be used to define IUGR. Moreover, only reference standards for triceps and subscapular skinfolds have been reported in the neonate (Oakley et al. 1977b, Whitehead et al. 1989). The paucity of data on the reproducibility of skinfold thickness measurements in the neonate further suggest that more studies need to be performed to determine the accuracy of this method of defining IUGR.

6. Although the weight / length ratio has been shown by some studies to be superior to ponderal index in the prediction of perinatal morbidity associated with IUGR, other studies have shown that this ratio may less discriminatory than the ponderal index in predicting neonatal complications. In the study by Jarai et al. (1977) (see Section 2.2.2.1), the weight / length ratio was of no use in differentiating between the hypoglycaemic and normoglycaemic infants as all the weight / length ratios were very low. This suggested that when the incidence of neonatal complications was high and deficits in weight markedly exceeded those in length, the discriminatory power of the ponderal index was better than that of the weight / length ratio. By far the greatest disadvantage of the weight / length ratio, however, is the absence of any reference standards reported in the literature for use in the definition of IUGR.

8. Although the non-invasive method described by Dauncey et al. (1977) for calculating percentage body fat and hence defining IUGR is potentially useful in clinical practice, it has not been validated by other standard methods of calculating fat content (Section 2.2.2.5). The model espoused by Dauncey et al. (1977) also assumes that the fat distribution throughout the body is constant, in contrast to the findings of Davies and Lucas (1990). The heavy reliance of this model on skinfold thickness measurements suggests that the latter is as useful as the former when used to define IUGR, without the need of a complicated formula to calculate percentage body fat.

2.2.3 Measures of adverse perinatal outcome

2.2.3.1 Introduction

An alternative to neonatal morphometric indices for the definition of IUGR are measures of adverse perinatal outcome associated with IUGR. It can be argued that whilst it may be important to define IUGR at birth using neonatal morphometric indices, it may be of more clinical relevance to define IUGR using measures of perinatal morbidity which reflect the clinical consequences of growth failure.

Although there is much in the literature concerning the perinatal outcome of SGA infants, there is comparatively little on the outcome of neonates with morphometric evidence of IUGR. Although SGA per se already confers an increased risk of significant perinatal morbidity and mortality, it is the SGA neonate with abnormal morphometry who is at further increased risk of adverse perinatal outcome (Patterson and Pouliot. 1987a, Villar et al. 1990, Fay et al. 1991). Studies which have reported data on adverse perinatal outcome associated with SGA and abnormal neonatal morphometry will therefore be discussed separately.

2.2.3.2 Adverse perinatal outcome and SGA

The risks associated with being SGA are well-documented. Perinatal mortality is significantly increased in infants born SGA at any gestational age (Starfield et al. 1982, Teberg et al. 1982). Numerous studies have also reported that, compared with AGA infants, SGA infants are at increased risk of fetal distress in labour and acidaemia (Dijxhoorn et al. 1987), hypoglycaemia (Lubchenco and Bard 1971), meconium aspiration (Steer 1989), pulmonary haemorrhage (Sly and Drew 1981) and polycythaemia (Makanson and Oh 1980).

Whilst the prediction of the SGA neonate remains an important objective of antenatal care, others question the true morbidity associated with SGA. In a study of 164 infants born with a birthweight < 5th centile for gestational age (of which 60 were below the 2.3rd centile), only one was hypoglycaemic enough to warrant intravenous glucose (Jones and Robertson 1986). Only five infants were admitted to neonatal intensive care unit (NICU) for complications associated with SGA (2 had hypoglycaemia, 1 had hypothermia and 1 had meconium aspiration). The incidence of significant morbidity associated with SGA appeared low. This is not surprising as SGA infants are a heterogeneous group and the vast majority of such infants are healthy and constitutionally small. Further evidence was provided by a study in which the incidence of IUGR within

the SGA population was ascertained (Fay et al. 1991b). A total of 418 infants were classified at birth using locally derived birthweight standards. Of the 42 diagnosed to be SGA, a thorough morphological and behavioural assessment by the paediatricians revealed that evidence of IUGR was present in 17 (40%) of these SGA infants. No differences were noted in the perinatal morbidity between SGA and AGA infants, but marked morbidity was noted in the sub-population of SGA infants with IUGR.

Therefore, classification of infants into those with and without morphometric evidence of IUGR is of greater value than classification by birthweight in the prediction of perinatal morbidity.

2.2.3.3 Adverse perinatal outcome and abnormal neonatal morphometry

The majority of perinatal outcome measures which have been used to define IUGR have been reported in studies based on abnormal neonatal morphometric indices of ponderal index (Jarai et al. 1977, Walther and Raemaker 1982, Patterson and Pouliot. 1987a, Villar et al. 1990, Fay et al. 1991a), MAC / HC ratio (Meadows et al. 1986, Georgieff et al. 1988) and skinfold thickness (Sumners et al. 1990).

In the largest study to date, Villar et al. (1990) evaluated the perinatal outcome in a cohort of 3450 SGA infants. Infants with a low ponderal index (n = 432) had a statistically higher risk (between 1.5 and 12 times) of a low 1- and 5-minute Apgar score, meconium aspiration, hypoglycaemia and "perinatal distress" than infants with an adequate ponderal index (n = 3018). In particular, growth retarded infants had a 6 fold increase in risk of hypoglycaemia, defined as blood glucose $\leq$ 35 mg / dl in the first 72 hours of life. Growth retarded infants also had a 5-fold increased risk of "perinatal distress", defined as the presence of an abnormal fetal heart rate pattern, low Apgar score, meconium-stained amniotic fluid or acidosis. Meconium aspiration syndrome (defined as the presence of meconium in the trachea whilst suctioning the airway after birth, clinical manifestations of aspiration or radiological evidence of meconium aspiration) was 10 times more likely in growth retarded SGA infants. Likewise, when compared with normally grown SGA infants, growth retarded infants were 3.5 times more likely to remain in hospital beyond 7 days. A logistic regression analysis revealed that this effect of ponderal index on neonatal morbidity was independent of birthweight and gestational age.

Other studies have likewise used perinatal outcome measures to define IUGR. However, the choice of adverse perinatal outcome measures used varied from study to study. The specific cut-off criteria used to define adverse perinatal outcome were presented in some but not all studies. Jarai et al. (1977) demonstrated a relationship between a low

ponderal index and hypoglycaemia (≤ 20 mg / dl in preterm infants, ≤ 30 mg / dl in term infants) as well as neonatal mortality. Walther and Raemaker (1982) evaluated the relationship between a low ponderal index and abnormal perinatal outcome in term infants. Abnormal perinatal outcome parameters evaluated were asphyxia (1- and / or 5-minute Apgar score ≤ 3 or need for at least one minute of positive pressure ventilation prior to respiration), acidosis (arterial pH < 7.09), hypothermia (rectal temperature < 35.5 degrees), hyperviscosity (venous haematocrit ≥ 65%), hypoglycaemia (blood glucose ≤ 1.6 mmol / l) and hyperbilirubinaemia (total bilirubin concentration > 200 µmol / l). Georgieff et al. (1988) in a study based on the MAC / HC ratio defined symptomatic IUGR in the postnatal period as the presence of hypoglycaemia (blood glucose < 30 mg / dl) , hypocalcaemia (serum calcium < 7.0 mg / dl) or polycythaemia (haematocrit > 65%). In another study also based on the MAC / HC ratio, hypoglycaemia (blood glucose < 2 mmol / l), poor temperature control (< 36 degrees) and necrotising enterocolitis were the criteria for defining adverse outcome (Meadows et al. 1986). Patterson and Pouliot (1987a) defined adverse outcome as operative delivery for fetal distress, 5-minute Apgar score ≤ 7, meconium aspiration, polycythaemia or hypoglycaemia. Sumners et al. (1990) defined an abnormal perinatal outcome as the presence of hypoglycaemia or temperature instability. Fay et al. (1991a) used a 5-minute Apgar score < 7 and emergency Caesarean section for fetal distress to define an abnormal outcome.

The detailed relationships between the neonatal morphometric indices and adverse perinatal outcome measures have been already been described in detail in Sections 2.2.2.2 to 2.2.2.4. The above summary demonstrates the absence of standardised measures of adverse perinatal outcome for the purposes of defining IUGR. Nevertheless, the most commonly used measures reported in the literature were operative delivery for fetal distress, acidaemia at birth, a low Apgar score at 5 minutes and hypoglycaemia.

2.2.3.4 Limitations of measures of adverse perinatal outcome

1. It will be evident from the above review that acidaemia and a low Apgar score were used in almost all the studies to define the perinatal complications of IUGR. The fact that IUGR predisposes to hypoxia is well established (Peeters et al. 1979, Rankin and McLoughlin 1979, Soothill et al. 1986). However, many have raised doubts as to whether the pathophysiology of asphyxia can be appropriately assessed by the Apgar score or acidaemia at birth (Nelson and Karin 1988, Ruth and Raivio 1988, Marlow 1992). Furthermore, these measures are non-specific and can be affected by many other obstetric conditions unrelated to IUGR.

2. Both acidaemia at birth and a low Apgar score are poor predictors of long-term morbidity. Numerous studies have shown a poor correlation between acidaemia at birth and subsequent neurological outcome (Jurgens-van der Zee et al. 1979, Dijxhoorn et al. (1986, Dijxhoorn et al. 1987, Ruth and Raivio 1988). The poor discrimination observed with arterial pH may be due to the failure to differentiate between respiratory and metabolic acidosis. Better discrimination may be achieved using base deficits to quantify the degree of perinatal asphyxia (Rosen and Murphy 1991). Numerous studies have also shown the 5-minute Apgar score to be a poor predictor of serious long-term morbidity (Nelson and Ellenberg 1981, Dijxhoorn et al. 1986, Dijxhoorn et al. 1987, Ruth and Raivio 1988). Furthermore, the 5-minute Apgar score also correlates poorly with acidaemia at birth (Sykes et al. 1983, Ruth and Raivio 1988, Steer et al. 1989 and Hoffman et al. 1991). Portman et al. (1990) suggested the incorporation of the Apgar score into a scoring system which includes other markers of perinatal asphyxia would improve the prediction of perinatal morbidity. It is evident that these measures cannot provide sole evidence that the infant has suffered sufficiently prolonged and severe asphyxia seondary to IUGR to produce permanent neurological deficit.

3. Although operative delivery for fetal distress has been used to define perinatal complications associated with IUGR, it is nevertheless another non-specific measure of asphyxia. Some studies have reported a significant association between abnormal cardiotocographic (CTG) tracings in labour and subsequent abnormal neurological outcome (Dijxhoorn et al. 1987) and acidaemia at birth (Gilstrap et al. 1987. However, numerous other studies have failed to demonstrate this relationship (Sykes et al. 1983, MacDonald et al. 1985, Steer et al. 1989). Another limiting factor with the use of this measure to define IUGR remains the correct interpretation and identification of abnormal CTG traces in labour.

4. Although meconium-stained liquor is noted to be more common in SGA compared with AGA newborns (Dijxhoorn et al. 1987, Steer 1989), it is a non-specific outcome measure

of morbidity associated with IUGR. Expulsion of meconium during labour is common, occurring in 18% of all cases (Nelson and Ellenberg 1984). In the same study, the presence of meconium did not predict subsequent outcome; 99.6% of infants who had meconium-stained amniotic fluid did not develop cerebral palsy. Gilstrap et al. (1987) also showed no relationship between meconium- stained amniotic fluid and subsequent acidaemia at birth.

5. Although admission to NICU remains one of the commonest measures of perinatal morbidity, it has to be used with caution for the definition of IUGR. The incidence of admission to NICU is greatly influenced by the maturity of the newborns. Therefore, the population of infants to be studied has to be clearly defined in relation to the gestational ages at delivery.

2.2.4 Biochemical indices of IUGR

2.2.4.1 Introduction

An alternative method of defining IUGR is the measurement of biochemical indices which either reflect the metabolic consequences of or the endocrine changes in IUGR. Biochemical indices can be ascertained from umbilical arterial or venous samples obtained at delivery, or from the neonate during the first few days after birth. Recently, cordocentesis has also enabled the biochemical status of the growth retarded fetus to be evaluated in-utero. The following is a review of metabolic (glucose, amino acid profile and triglycerides) and endocrine [insulin, insulin-like growth factors (IGF)] factors which may be useful for the purposes of defining IUGR .

2.2.4.2 Metabolic responses to IUGR

Studies based on biochemical data obtained at delivery (Bozzeti et al. 1988), in the neonatal period (Jarai et al. 1977, Lubchenco and Bard 1971, Hawdon et al. 1992a) and during cordocentesis (Soothill et al. 1987a, Economides et al. 1989b) all suggest that the the process of IUGR leads to hypoglycaemia. In IUGR, glucose uptake by the fetus is reduced together with an increased maternal-fetal glucose gradient, suggesting a reduced supply of substrates due to placental hypoperfusion (Economides et al. 1989b). This leads to reduced glycogen stores and increased tissue catabolism. Therefore, the growth retarded fetus / neonate has less glycogen stores to mobilise to produce glucose, and gluconeogenesis from amino acids and lipolysis of fatty acids is also reduced (Hawdon et al. 1992a).

Amino acids are also major substrates for fetal energy production and growth. The total uptake of amino acids by the fetus exceeds the net accretion rate, consistent with the role of amino acids as substrates for energy production (Battaglia et al. 1969). Studies which have evaluated the amino acid profile of SGA and AGA neonates have shown a significantly lower total α-aminonitrogen concentration in SGA infants than the AGA infants (Cetin et al. 1988). The neutral branched amino acids, valine, leucine and isoleucine, all essential to the growing fetus, account for most of the difference in total α-aminonitrogen levels between AGA and SGA neonate. The fetus may also lose the branched-chain amino acid, isoleucine, to the placenta (Hayashi et al. 1978, Cetin et al. 1988). Such a net loss of protein could explain the phenomenon of "wasting" evident in the newborn. Cordocentesis data from small fetuses showed the response of the non-essential amino acids to be variable (Economides et al. 1989a, Bernadini et al. 1991). Alanine was increased, whereas serine and tyrosine were decreased in such small fetuses.

The plasma non-essential / essential acid ratio was increased, suggesting a degree of intrauterine malnutrition. This increase in gluconeogenic substrates has also been shown in studies based on SGA newborns (Haymond et al. 1974, Mestyan et al. 1975).

Lipid metabolism is important for cellular and organ growth. Studies based on cord blood data obtained at delivery (Gustafson et al. 1972, Dhanireddy et al. 1981) and during cordocentesis (Economides et al. 1988) have suggested that triglyceride levels are raised in small fetuses. There are two possible mechanisms for this: the first is that IUGR resulting in hypoxia causes the lipolysis and mobilization of fats in order to provide additional substrate for oxidation. The second is that triglyceride levels are raised because of reduced uptake into adipose tissue. Data from the cordocentesis study of Economides et al. (1988) suggested that in small fetuses, plasma triglyceride concentrations were increased but the other products of lipolysis (non-esterified fatty acids and glycerol) were not. Therefore, it is unlikely that triglyceride levels are raised secondary to increased lipolysis, but rather due to reduced utilisation for fat deposition.

2.2.4.3 Endocrine control of fetal growth

Many fetal endocrine systems have been examined for a possible role in controlling fetal growth. In a review of the endocrine control of fetal growth, Chard (1989) summarised the role of primary endocrine factors like the fetal pituitary-adrenal axis, placental growth hormone, thyroid hormones, growth hormone and placental lactogen. Although all these have previously been implicated in determining fetal growth, there is much contradictory evidence to suggest that they play no more than a permissive role. Only two, insulin and the insulin-like growth factors (IGF), are useful for the purposes of defining IUGR.

Fetal hyperinsulinism is a known response of the the fetal pancreas to maternal hyperglycaemia in diabetes mellitus (Hill and Milner 1985). Conversely, fetal hypoinsulinaemia is associated with IUGR. Although insulin can exert a direct somatotrophic effect on cells and organs, it is difficult to ascertain whether this is the primary mode of action, or whether it exerts its effects secondarily through altered carbohydrate and lipid metabolism (Hill 1989). Data from cordocentesis studies suggest that hypoinsulinaemia is not secondary to hypoglycaemia; the small fetus has a lower insulin-glucose ratio than a fetus of normal size (Economides et al. 1989b). It is more likely to be the result of pancreatic β-cell dysfunction, a finding confirmed by van Assche et al. (1977) who found reduced endocrine pancreatic tissue at autopsy in SGA infants. Hypoinsulinaemia therefore reduces the availability of nutrients in skeletal muscles, liver and adipose tissue. The result will be decreased glycogen and fat stores and impaired fetal

growth.

Numerous workers have reported a relationship between cord insulin levels at delivery and birthweight (Lin et al. 1981, Weiss et al. 1984, Spellacy et al. 1987, Stanley et al. 1992). Only one has reported reference ranges for insulin, based on birthweight distribution (Weiss et al. 1984). In this study, a normal range was reported for cord insulin levels based on 180 infants with a birthweight < 90th centile from which obese mothers and diabetics were excluded. In the largest study to date, cord insulin from 209 unselected singleton births were measured and its distribution relative to birthweight centiles was examined (Stanley et al. 1992).There was only a weak correlation between decreasing birthweight and insulin levels. There was considerable overlap in the 95% range of insulin levels when infants were classified according to birthweights less than the 10th centile, between 10th and 25th centiles and between the 25th and 50th centiles. Low insulin levels also occurred in each of the birth weight centile groups, highlighting the marked limitations of establishing centile values for cord insulin based on birthweight distribution.

Insulin-like growth factors are mitogenic peptides which are structurally homologous to pro-insulin. They are synthesized in connective tissue and cells of mesenchymal origin in a wide variety of adult and fetal organs. A large amount of indirect evidence suggests that IGF-1 is involved in the regulation of fetal growth (Gluckman 1989, Hill 1989). Cord blood data at delivery have shown that IGF-1 levels correlate with birthweight (Foley et al. 1980, Bennett et al. 1983, Gluckman et al. 1983). A recent study using cordocentesis data has allowed further assessment of IGF-1 levels in the fetus (Lassarre et al. 1991). In this study, IGF-1 and IGF-2 were measured in cord blood obtained at cordocentesis in 103 subjects between 20 and 37 weeks of gestation. IGF-1 and IGF-2 levels were constant until 33 weeks of pregnancy but thereafter increased with gestation until term. In 16 small fetuses, IGF-1 levels were significantly lower than normal size fetuses of the same gestation. These findings suggest that in the last few months of life, IGF-1 may be involved in the control of fetal size.

The above review suggests that the nutritional status of a fetus / neonate may be quantified by measuring glucose, essential and non-essential amino acids, triglycerides, insulin and IGF-1 levels.

2.2.4.4 Limitations of biochemical indices of IUGR

1. The measurement of biochemical indices from cord blood at delivery may not reflect the true hormonal milieu of the fetus prior to labour. Maternal fasting, anaesthesia, surgery and the stress of labour have a variable effect on fetal metabolism. This may limit the usefulness of cord data obtained at delivery for the purposes of defining IUGR.

2. The metabolic status of the neonate in the first few days after birth may be altered by differing feeding practices (Hawdon et al. 1992a). The response of the growth retarded infant to substrates may differ according to the method of feeding and the maturity of the infant (Lucas et al. 1988). Therefore, if biochemical indices in the neonate are to be evaluated, reference ranges different to those established from cord blood data will have to be constructed which take these factors into account.

3. Biochemical data obtained at cordocentesis would circumvent the problems mentioned above. Although reference ranges for glucose (Economides et al. 1989b), triglycerides (Economides et al. 1988), insulin (Economides et al. 1989b) and IGF-1 (Lassarre et al. 1990) have been reported, such standards were constructed from the data of high-risk fetuses where the procedure was undertaken for exclusion of a chromosomal anomaly, assessment of acid-base status and interventional procedures like fetal blood transfusion for rhesus iso-immunisation. Although only the results of karyotypically normal fetuses were used, it remains questionable as to whether such reference ranges can be used to evaluate cord biochemical data obtained at delivery or in the neonatal period.

4. Reference ranges constructed from cordocentesis data were derived by classifying fetuses into whether they were small or normal sized based on ultrasound measurements. However, not all small fetuses are growth retarded. A proportion of small fetuses with normal growth would therefore have been included in the derivation of these standards.

5. Reference ranges derived from cord blood obtained at delivery have been reported for only two of the biochemical indices reviewed, insulin and glucose. However, the reference standards for insulin were derived from infants divided by birthweight criteria (Weiss et al. 1984). Such reference ranges are therefore of dubious value for the purposes of differentiating IUGR from non-IUGR.

6. Although reference standards are available for the assessment of umbilical venous glucose levels at delivery, the cut-off criteria for defining hypoglycaemia varied from study to study (Pildes et al. 1967, Lubchenco and Bard 1971, Koh et al. 1988, Hawdon et al. 1992a). Like reference standards for insulin, such reference standards were also based on infants divided by birthweight criteria with all the disadvantages already mentioned.

2.2.5 Choice of optimal outcome measure for the definition of IUGR

It is evident from the review of literature in Sections 2.2.1. to 2.2.4 that there is little agreement as to which outcome measure to use to define IUGR. This choice is important as any ultrasound parameter which purports to be of diagnostic value in the identification of IUGR must be tested against a "gold standard". Neonatal morphometric indices, measures of adverse perinatal outcome and cord biochemical data all provide, to a certain extent, evidence of IUGR in the newborn.

Neonatal morphometric indices, such as the ponderal index, MAC / HC ratio and subscapular and triceps skinfold thickness, have three important advantages. First, these indices provide quantitative measures of wasting and lack of fat, evidence that the neonate had suffered some degree of impaired growth in-utero (Hill et al. 1984, Patterson and Pouliot 1987a, Villar et al. 1990, Fay et al. 1991a). Second, these morphometric indices correlate well with measures of perinatal morbidity associated with IUGR (Section 2.2.2.3). Third, reference standards for ponderal index, MAC / HC ratio and skinfold thickness have been published for the confirmation of IUGR in the newborn (Miller and Hassanein 1971, Sasanow et al. 1986, Oakley et al. 1977b).

Many limitations with measures of adverse perinatal outcome preclude their use as the sole outcome measure for the purposes of IUGR (Section 2.2.3.4). These measures of morbidity are not specific to IUGR as they can also be affected by many other unrelated obstetric conditions (Freeman and Nelson 1988). Nevertheless, adverse perinatal outcome can be used as a supplementary outcome measure of IUGR. Amongst all the biochemical markers of IUGR, reference standards for biochemical cord data at delivery have only been reported for glucose and insulin (Section 2.2.4). However, the definition of hypoglycaemia varied according to which reference standard was used (Hawdon et al. 1992a). Although reference standards derived from cordocentesis data have been reported for numerous biochemical indices, such standards are inappropriate for the assessment of indices measured in cord blood at delivery. Furthermore, no one biochemical marker can be used in isolation to confirm IUGR. Growth retarded neonates tend to be hypoglycaemic, hypoinsulinaemic, hypertriglyceridaemic and have low IGF-1 levels. The use of cordocentesis to ascertain the biochemical status of such fetuses would be inappropriate and unnecessary (Nicolini et al. 1989). The use of cord biochemistry at delivery would therefore serve only as an ancillary marker of growth retardation.

The evidence strongly suggests neonatal morphometric indices to be the best outcome measure for the purposes of defining IUGR. Neonatal morphometry should therefore be used as the "gold standard" against which ultrasound parameters are to be evaluated.

2.3 Ultrasound diagnosis of IUGR

2.3.1 Ultrasound and prediction of SGA

2.3.1.1 Review of literature

Numerous ultrasound parameters, including Doppler waveform indices, have been used to predict SGA. At least 117 studies have reported the use of ultrasound in the diagnosis of IUGR, as defined by a low birthweight centile. This consisted of 86 which used non-Doppler ultrasound measurements and 37 Doppler studies. Table 2.1 lists the various ultrasonic measurements that have been reported for this purpose. Reviews of the use of ultrasound (Benson et al. 1986, Deter et al. 1986, Villar and Belizan 1986, Secher et al. 1987b, Low 1991, Divon and Hsu 1992) to predict SGA at birth have generally reported results for individual studies with little summary statistics on ultrasound parameters. It was therefore unclear from the literature as to which ultrasound measure was the best predictor of SGA at birth.

As summary statistics for each ultrasound parameter would allow different ultrasound parameters to be compared, all previous studies were reviewed and grouped so that common sensitivities, common false positive rates (FPR) and common odds ratios (OR) could be calculated for each ultrasound parameter when appropriate. Studies were included in the calculation of summary statistics only if the following criteria were met: a) The criteria for the antenatal diagnosis were clearly defined [eg. AC < 10th centile] b) The postnatal criterion for the diagnosis of SGA was birthweight < 10th centile for gestational age c) Data for normal and SGA fetuses were reported, enabling the construction of a two-by-two table. The antenatal and postnatal reference standards used in the individual studies were accepted. Studies with the same diagnostic and postnatal criteria were then grouped according to whether the population was high risk (ie. subjects with a poor obstetric history, including a previous SGA infant, or a clinical suspicion of SGA) or low risk (ie. unselected subjects). Because a large number of non-Doppler studies used diagnostic criteria which were unique, only those which fulfilled the same antenatal and postnatal criteria were included in our study. In view of the smaller number of studies involving Doppler ultrasound, all those with appropriate postnatal criteria were included.

The sensitivity, FPR and OR [with 95% confidence interval (CI)] (Kahn and Sempos 1989) were calculated for individual studies. Where more than one study used the same antenatal and postnatal criteria, an analysis was then performed to determine if there was a statistically significant difference in the sensitivities, FPRs and ORs within each diagnostic group (Zelen 1971, Breslow and Day 1980). This was done using a test of heterogeneity (Fisher's exact test) for sensitivities and FPRs and either an asymptotic test

or an exact test (Gart 1970) for ORs, depending on the numbers available for analysis. For large numbers, an exact test was not computationally feasible (Mantel et al. 1959). Where the p was > 0.01 (ie the individual sensitivities, FPRs or ORs were not significantly different), a common sensitivity, common FPR or common OR was calculated.

Thirty six non-Doppler and 24 Doppler studies satisfied these criteria. The remaining studies were excluded either because the antenatal criteria were unique or because different birth weight criteria were used, ranging from < 1 SD to < 2.5th centile. A common sensitivity, OR and FPR could not be calculated in a large proportion of groups due to significant differences between the individual studies. The results of individual studies in high and low risk populations are shown in Tables 2.2 and 2.3. Abdominal circumference and EFW were the best predictors of a SGA infant at birth. In high risk patients AC < 10th centile predicted 84% of SGA fetuses with a common OR of 18.4, compared with a common OR of 39.1 using EFW < 10th centile. Both ultrasonic measurements had comparable false positive rates. Comparable values were lower in low risk patients but the same trend was seen. The results suggest that the use of AC < 10th centile would detect the highest percentage of SGA neonates but the odds of being small in any individual fetus are greatest if the EFW is < 10th centile.

The results of Doppler studies in high and low risk subjects are shown in Tables 2.4 and 2.5 respectively. In a high risk population the common OR for umbilical artery systolic-diastolic (S / D) ratio ≥ 3 was significantly lower than the common OR for EFW < 10th centile. The common OR for umbilical artery S / D > 95th centile was also significantly lower than the common OR for AC < 10th centile. Abnormal uteroplacental waveforms had a lower common OR than umbilical artery S / D ratio. Individual Doppler studies that have incorporated internal carotid waveforms in high risk subjects have reported higher sensitivities, ORs and lower FPRs. In the low-risk population, the common sensitivity and OR for uteroplacental S / D ratio > 95th centile were very low with high FPRs. Other individual studies generally reported equally poor sensitivities, ORs and FPRs with the exception of one study (Maulik et al. 1990) in which umbilical artery S / D ratio > 3 detected 74% of SGA fetuses with OR of 10.2; however, the false positive rate was high (68%).

Six studies (Berkowitz et al. 1988a, Divon et al. 1988, Gaziano et al. 1988, Chambers et al. 1989, Newnham et al. 1990, Miller and Gabert 1992) compared AC or EFW with Doppler measurements in the same group of subjects in the prediction of SGA (Table 2.6). In three of these studies, the Doppler ORs were significantly lower than those for AC or EFW (Chambers et al. 1989, Newnham et al. 1990, Miller and Gabert 1992). In the remaining three studies the same trend was apparent (Berkowitz et al. 1988a, Divon et al. 1988, Gaziano et al. 1988), although the differences were not statistically significant.

That AC and EFW were superior to umbilical artery and uteroplacental waveform indices in the prediction of SGA was evident when comparing sensitivities and ORs both within the same study populations and in different studies. More recent studies on the Doppler assessment of the of internal carotid (Arduini et al. 1987, Degani et al. 1990) and middle cerebral (Arabin et al. 1992, Gramellini et al. 1992) circulations have reported higher sensitivities and ORs. These results, if confirmed, would suggest that 'cerebral sparing', as evidenced by the low pulsatility indices in the the cerebral vessels of such fetuses, may be better than umbilical or uteroplacental waveforms in predicting SGA fetuses.

2.3.1.2. Limitations of previous studies

1. Most previous studies investigating the ultrasonic prediction of SGA have reported results as sensitivities and positive predictive values. However, the positive predictive value reported in any individual study is dependent on the prevalence of SGA in the study group (Villar and Belizan 1986, Wald and Cuckle 1989). To overcome this problem, some workers have suggested that the positive predictive value should be quoted for a standardised population prevalence of SGA of 10% using Bayes theorem (Simon et al. 1990b). While the sensitivity of an ultrasound test has generally been thought to be prevalence-independent, its use in comparing different ultrasound parameters in the prediction of SGA may be one exception to this (Stempel 1982). It was therefore important to compare different ultrasound parameters using another statistical method which is prevalence-independent. The OR was therefore used in this review to compare individual studies.

3. Divon and Hsu (1992) recently used ORs to compare individual studies of Doppler ultrasound in the prediction of SGA. However, no summary statistic was reported to enable comparison between different Doppler parameters in this or any other previous review article. The common OR was therefore used in this review to derive a clinically useful summary statistic for comparing different ultrasound parameters (Kahn and Sempos 1989).

4. It was apparent from the review of studies that numerous clinical differences existed between studies which evaluated the same ultrasound parameter. The methods used to determine a particular ultrasound measurement varied. This was particularly evident in studies evaluating the use of EFW where different formulae were used to calculate fetal weight (Divon et al. 1988, Gaziano et al. 1988, Simon et al. 1990a). The methods of quantifying total intrauterine volume also varied between the studies of Gohari et al.

(1977), Chinn et al. (1981) and Gierrson et al. (1985b). Reduced amniotic fluid volume was also defined using different criteria in different studies. Such heterogeneity in methods of measurement limit the direct comparison of these individual studies.

5. The gestational age (GA) at ultrasound assessment and the interval between assessment and delivery varied from study to study. Some studies included second trimester measurements (Ott 1985, Warsof et al. 1986, Gaziano et al. 1988) while most others were limited to the third trimester.

6. The antenatal reference standards used for individual ultrasound parameters also varied from study to study; for example, numerous cross-sectionally derived (Fescina et al. 1982, Hadlock et al. 1982a, Woo et al. 1984) and longitudinal (Deter et al. 1982a, Jeanty et al. 1984a) reference standards for AC were used in the studies reviewed. The antenatal cut-off criteria used to define abnormality also varied from study to study [eg. < 2.5th centile (Hadlock et al. 1982a), < 5th centile (Selbing et al. 1984) and <10th centile (Neilson et al. 1980, Woo et al. 1984)].

7. A variety of birthweight standards were used to define SGA in the individual studies. Many of the birthweight charts used were derived from another totally different reference population. The inappropriateness of this indiscriminate use of birthweight standards, together with the discrepancy in the 10th centile values between different birthweight standards, have been discussed in Section 2.2.1.2.

8. Even when similar antenatal and postnatal criteria and reference standards were used, there were often statistically significant differences in the sensitivities, ORs and FPRs between individual studies precluding the calculation of a common statistic. Thompson and Pocock (1991) recently reviewed these problems in relation to the meta-analysis of clinical trials, highlighting the problems of statistical and clinical heterogeneity between individual studies.

9. The most important limitation with all studies reviewed here remains the use of SGA at birth to define IUGR. All the studies reviewed in this section evaluated the ability of ultrasound to predict birthweight < 10th centile. The many disadvantages with this outcome criterion have already been enumerated in Section 2.2.1.2. Smallness for gestational age is not synonymous with IUGR and more appropriate outcome criteria need to be used to define IUGR.

Table 2.1 Summary of ultrasound parameters used in previous studies to predict SGA in high and low risk populations.

Ultrasound parameter	No of studies	High-risk	Low-risk
EFW	21	12	9
BPD	18	9	9
AC	16	8	8
FL / AC	9	7	2
Amniotic fluid volume	7	7	-
HC / AC	5	4	1
Total intrauterine volume	4	3	1
Placental grade	4	2	2
Fetal ponderal index	4	3	1
Trunk area x crown-rump length	3	2	1
Abdominal area	2	1	1
Chest area	2	1	1
Abdominal diameter	1	-	1
Head Area / abdominal area	1	-	1
Liver size	1	-	1
FL / HC	1	-	1
FL	1	1	-
Distal femoral epiphyseal ossification	1	-	1
Thoracic diameter	1	-	1
Thigh circumference	1	1	-
FL / thigh circumference	1	1	-
Trunk area	1	1	-
Doppler			
Umbilical artery	30	24	6
Uteroplacental	7	3	4
Descending aorta	3	3	-
Internal carotid artery	2	2	-

Abbreviations; BPD, biparietal diameter; FL, femur length; HC, head circumference.

Table 2.2 Ultrasonic measurements in the prediction of SGA in high risk subjects.

Ultrasound Criteria	Reference	GA	INDIVIDUAL Se	INDIVIDUAL OR	INDIVIDUAL FPR	COMMON† Se	COMMON† OR (95% CI)	COMMON† FPR
BPD < 10th centile	Geirsson et al. (1985b)	35-37	41.7	4.7	42.3	-	**5.8** (3.6, 9.4)	**47.7**
	Brown et al. (1987)	III	72.7	6.3	48.8			
AC <10th centile	Chambers et al. (1989)	III	72.9	10.8	16.2	**84.4**	**18.4** (9.8, 34.3)	-
	Brown et al. (1987)	III	95.4	31.2	49.4			
HC/AC >95th centile	Hill et al. (1989)	III	53.0	3.8	18.6	**47.6**	**3.3** (1.6, 7.2)	**24.5**
	Hassan et al. (1989)	III	27.8	2.3	50.0			
EFW <10th centile	Palo et al. (1989)	III	82.5	47.7	9.1	-	**39.1** (28.9, 52.8)	-
	Simon et al. (1990b)	III	78.8	54.0	23.0			
	Simon et al. (1988)	III	33.3	19.4	46.3			
	Divon et al. (1988)	III	86.7	42.0	22.0			
	Brown et al. (1987)	III	65.9	47.4	12.1			
	Gaziano et al. (1988)	15-44	43.5	45.4	23.1			
	Berkowitz et al. (1988a)	30-42	78.6	20.7	36.5			
	Ott and Doyle (1984)	NK	89.2	42.7	40.4			
EFW <5th centile	Simon et al. (1988)	III	51.6	23.1	36.0	-	-	**23.2**
	Simon et al. (1990b)	III	52.6	48.1	12.3			
	Hill et al. (1989)	III	80.3	3.1	27.4			
FL/AC >23	Palo et al. (1989)	31-42	44.7	7.5	12.5	**49.1**	**2.8** (1.7, 4.7)	-
	Ott (1985)	20-42	52.3	2.2	69.4			
FL/AC >23.5	Ott (1985)	20-42	36.9	1.7	71.1	-	-	-
	Divon et al. (1986)	III	55.5	10.7	18.5			
	Divon et al. (1988)	III	44.4	8.6	25.9			

Table 2.2.(continued) Ultrasonic measurements in the prediction of SGA in high risk subjects.

Ultrasound Criteria	Ref	GA	INDIVIDUAL Se	OR	FPR	COMMON† Se	OR (95% CI)	FPR
FL/AC >24	Ott (1985)	20-42	27.7	1.8	69.5	-	**2.7**	**58.2**
	Brown et al. (1987)	III	54.5	3.6	51.5		(1.8, 4.1)	
Placenta Grade 3	Kazzi et al. (1983a)	III	61.5	2.3	72.1	**61.7**	**3.1**	-
	Kazzi et al. (1983b)	III	61.9	4.4	40.9		(1.8, 5.3)	
Reduced Amniotic fluid	Manning et al. (1981)	27-40	83.9	133.1	10.3	-	-	-
	Chamberlain et al. (1984)	NK	5.5	11.4	61.4			
	Divon et al. (1988)	III	15.5	7.4	22.2			
	Hassan et al. (1987)	III	44.4	2.4	52.9			
	Fescina et al. (1987)	II / III	27.8	21.3	9.1			
	Philipson et al. (1983)	II / III	82.6	7.2	20.0			
	Ott and Doyle (1984)	III / III	53.3	31.9	60.4			
Reduced TIUV	Geirrson et al. (1985b)	III	69.2	13.1	32.5	**61.0**	-	-
	Gohari et al. (1977)	II / III	75	∞	0.0			
	Chinn et al. (1981)	22-41	48.2	2.1	69.0			

† Common sensitivity, OR and FPR only given if the test of equal sensitivities, ORs or FPRs is not rejected at the 1% level.

Abbreviations; Se, sensitivity; NK, not known; TIUV, total intrauterine volume; II, second trimester; III, third trimester.

Table 2.3 Ultrasonic measurements in the prediction of SGA in low risk subjects.

Ultrasound Criteria	Reference	GA	INDIVIDUAL Se	OR	FPR	COMMON† Se	OR (95% CI)	FPR
BPD < 10th centile	Rosendahl et al. (1988)	34	27.3	11.4	69.8	-	-	-
	Warsof et al. (1986)	>25	53.6	6.1	31.7			
	Geirsson et al. (1985a)	32-36	32.0	3.8	76.5			
	Gerhard et al. (1987)	35-40	29.7	12.2	54.2			
	Hughey et al. (1984)	III	40.0	6.7	93.3			
AC <25th centile	Warsof et al. (1986)	III	80.0	20.7	28.6	**80.7**	**21.1** (14.1, 31.8)	-
	Sarmandal et al. (1990)	34-36	85.7	23.6	69.2			
AC <10th centile	Warsof et al. (1986)	>25	48.2	12.9	38.9	-	**13.5** (11.5, 15.9)	**39.8**
	Simon et al. (1990a)	III	64.2	16.2	39.0			
	Duff et al. (1986)	33-35	56.4	18.8	55.1			
AC <5th centile	Simon et al. (1990a)	III	47.7	8.2	46.3	-	-	**45.6**
	Ferrazi et al. (1986)	III	88.3	39.9	42.6			
	Newnham et al. (1990)	34	47.2	13.6	59.5			
EFW < 10th centile	Secher et al. (1987a)	37	34.3	16.5	54.0	-	-	-
	Secher et al. (1987b)	32-37	31.1	10.7	54.9			
	Secher et al. (1986a)	32-37	38.2	20.6	40.3			
	Weiner et al. (1989)	28-43	73.3	18.8	54.1			
	Simon et al. (1990a)	III	68.2	28.8	29.4			
Placenta Grade 3	Vosmer et al. (1989)	26-43	38.9	0.9	87.7	**36.1**	**0.8** (0.4, 1.6)	**89.0**
	Miller et al. (1988)	>37	34.5	0.8	89.7			

† Common sensitivity, OR and FPR only given if the test of equal sensitivities, ORs or FPRs is not rejected at the 1% level.

Abbreviations; Se, sensitivity; III, third trimester.

Table 2.4 Doppler waveform indices in the prediction of SGA in high risk subjects.

Ref	GA	INDIVIDUAL			COMMON†		
		Se	OR	FPR	Se	OR (95% CI)	FPR
Umbilical artery S / D >3							
Lowery et al. (1990)	NK	63.6	3.4	61.3	**52.6**	**6.9**	-
Mulders et al. (1987)	NK	53.3	7.6	33.3		(4.8. 10.0)	
Fleischer et al. (1985)	18-42	78.3	17.4	51.3			
Berkowitz et al. (1988a)	30-42	38.9	6.3	44.7			
Divon et al. (1988)	III	39.3	14.7	18.5			
Gaziano et al. (1988)	15-44	79.2	7.2	78.9			
Ott (1990b)	III	48.3	4.8	46.1			
Umbilical artery S / D >95th centile							
Dempster et al. (1989)	>30	41.5	3.3	39.3	-	**5.8**	**43.1**
Trudinger et al. (1985)	III	72.1	7.4	43.6		(4.8, 7.0)	
Al-Ghazali et al. (1988)	16-42	40.0	5.3	33.3			
Trudinger et al. (1986)	12-44	50.2	5.9	44.1			
Uterine artery RI > 0.5							
Chambers et al. (1989)	III	29.4	1.7	32.4	-	**2.3**	-
Jacobson et al. (1990)	24	70.6	4.3	69.2		(1.2, 4.3)	
Umbilical artery PI ≥ 1SD							
Arduini et al. (1987)	26-28	60.9	4.1	50.0			
Umbilical artery PI ≥2 SD							
Degani et al. (1990)	29-40	76.3	3.2	27.5			
Aortic PI >1 SD							
Arduini et al. (1987)	26-28	65.2	3.8	53.1			
Aortic blood flow class							
Laurin et al. (1987)	III	50.7	27.7	7.3			
Internal carotid artery PI <1 SD							
Arduini et al. (1987)	26-28	69.6	25.5	20.0			
Internal carotid artery PI <2 SD							
Degani et al. (1990)	29-40	78.9	34.1	3.2			
Umbilical / Internal carotid PI ≥1 SD							
Arduini et al. (1987)	26-28	78.3	39.4	18.1			
Umbilical / Internal carotid PI ≥2 SD							
(Degani et al. (1990)	29-40	84.2	47.2	3.0			

† Common sensitivity , OR and FPR only given if the test of equal sensitivities, OR's or FPRs is not rejected at the 1% level.

Abbreviation; Se, sensitivity.

Table 2.5 Doppler waveform indices in the prediction of SGA in low risk subjects.

Reference	GA	INDIVIDUAL			COMMON†		
		Se	OR	FPR	Se	OR (95% CI)	FPR
Uterine S/D ≥95th centile							
Newnham et al. (1990)	34	8.6	2.2	84.2	**5.9**	**1.6**	**89.3**
Hanretty et al. (1989)	34	0	0	100.0		(0.47, 5.6)	
Uterine RI ≥95th centile							
Bewley et al. (1991)	16-24	15.3	4.0	65.3			
Uterine RI ≥0.55							
Steele et al. (1988)	24	33.3	5.1	60.0			
Umbilical artery PI ≥95th centile							
Sijmons et al. (1989)	34	21.7	4.7	47.1			
Umbilical artery S/D ≥95th centile							
Newnham et al. (1990)	34	16.7	3.9	76.9			
Umbilical artery S/D ≥3							
Maulik et al. (1990)	32-36	74.4	10.2	68.0			

† Common sensitivity, OR and FPR only given if the test of equal sensitivities, ORs or FPRs is not rejected at the 1% level.

Abbreviations; Se, sensitivity; RI, resistance index.

Table 2.6 Comparison of AC / EFW with Doppler ultrasound in studies with similar subjects.

Ref	U/S criteria	Sensitivity (%)	OR (95% CI)	Doppler criteria	Sensitivity (%)	OR (95% CI)
Berkowitz et al. (1988a)	EFW < 10th centile	78.6	20.1 (8.5, 50.0)	UA S / D > 3	38.9	6.3 (2.9, 13.8)
Chambers et al. (1989)	AC < 10th centile	72.9	10.8 (4.8, 23.8)	Uterine artery RI > 0.5	29.4	1.7 (0.8, 3.7)
Divon et al. (1988)	FL/AC > 23.5	44.4	8.6 (3.2, 22.7)	UA S / D > 3	39.3	14.7 (5.0, 43.3)
Gaziano et al. (1988)	EFW <10th centile	43.5	45.4 (11.7, 227.2)	UA S / D > 3	79.2	7.2 (2.7, 22.3)
Newnham et al. (1990)	AC < 5th	47.2	13.6 (6.2, 29.6)	Uterine artery S / D > 95th centile	8.6	2.2 (0.5, 6.1)
Miller and Gabert (1992)	Relative EFW < 0.8	82.6	28.1 (10.6, 74.4)	UA S / D > 3	67.3	6.4 (2.9, 14.1)
Miller and Gabert (1992)	Relative EFW < 0.784	80.4	42.1 (14.9, 119.0)	UA S / D > 2.085	82.6	12.3 (5.0, 30.4)

Abbreviations; RI, resistance index; UA, umbilical artery.

2.3.2 Ultrasound and prediction of neonatal morphometry

2.3.2.1 Review of literature

Despite the inappropriateness of using birthweight to define IUGR, ultrasound studies have continued to use this outcome criteria to define IUGR. However, recent ultrasound studies have begun to use neonatal morphometric indices such as ponderal index, MAC / HC ratio and subscapular and triceps skinfold thickness to define IUGR. The advantages of using each of these outcome criteria have been enumerated in Section 2.2.2.

In contrast to the 117 studies reported in the literature on ultrasound in the prediction of SGA, only eight studies have reported data on ultrasound in the prediction of abnormal neonatal morphometry. Neonatal ponderal index was used as the outcome criteria in five ultrasound studies evaluating non-Doppler ultrasound measurements (Ott 1985, Vintzileos et al. 1986, Patterson et al. 1987b, Weiner and Robinson 1989, Sarmandal and Grant 1990) and three Doppler ultrasound studies (Beattie and Dornan 1989, Sijmons et al. 1989, Trudinger et al. 1991). In addition, Sarmandal and Grant (1990) also used the MAC / HC ratio to define IUGR whilst Beattie and Dornan (1989) also used skinfold thickness and the MAC / HC ratio to define IUGR.

Four of these studies were based on high-risk obstetric populations. In a study of 326 high risk fetuses, Ott (1985) evaluated the FL / AC ratio in the prediction of an abnormal neonatal ponderal index. Three different antenatal values (23.0, 23.5 and 24.0) were used to define abnormality. In another study of high-risk fetuses, Patterson et al. (1987b) evaluated the use of amniotic fluid volume to predict a low ponderal index. An average of the vertical and two perpendicular horizontal diameters was used to calculate the average amniotic fluid diameter. An antenatal cut-off value of < 3.2 cm, determined using a receiver-operating characteristic (ROC) curve, was used to define abnormality. In another study based on 113 high-risk fetuses, the in-utero ponderal index was used to predict subsequent abnormal neonatal ponderal index (Vintzileos et al. 1986). The in-utero ponderal index was calculated by dividing the EFW (derived using a formula incorporating BPD and AC) by the cube of the femur length. All of the above three studies provided data such that contingency tables could be constructed. In the only other study which addressed the ultrasound prediction of neonatal morphometric indices of IUGR in a high-risk population, two-by-two tables could not be constructed from the data presented. Trudinger et al. (1991) evaluated the use of umbilical artery Doppler in 2178 high-risk fetuses. Neonatal ponderal index was significantly lower (ponderal index = 0.218) in those with absent end-diastolic flow in the umbilical artery waveform compared with those who had a normal S / D ratio (ponderal index = 0.260).

Sensitivities, FPRs and ORs (and 95% confidence intervals) were calculated from the data reported in the three studies of Ott (1985), Patterson et al. (1987b) and Vintzileos et al. (1986) [Table 2.7]. The sensitivities for FL / AC ratio were generally poor (< 52%). An equally poor sensitivity of 53.8% was noted when a measure of reduced amniotic fluid was used. A much higher sensitivity was achieved using the fetal ponderal index, detecting 76.9 % of all neonates with subsequent abnormal neonatal morphometry. The highest ORs were obtained using amniotic fluid diameter (6.31) and fetal ponderal index (14.67); corresponding values for FL / AC ratio were lower (all ORs < 4.33). Although the fetal ponderal index appeared to have an advantage over the other ultrasonic parameters, it had a high false positive rate (64.3%). Its use in screening a high-risk population would therefore result in an inappropriately large number of normally grown fetuses being incorrectly identified as being growth retarded. In the study by Vintzileos et al. (1986), the incidence of IUGR, as defined by an abnormal neonatal ponderal index, was 11.5%. If this test were to be applied to the obstetric population at large, in which the prevalence of IUGR would be lower, the false positive rate for fetal ponderal index would be considerably larger as prevalence decreases (Villar and Belizan 1986).

The remaining four studies assessed the use of ultrasound parameters in the prediction of abnormal neonatal morphometry in low-risk subjects (Beattie and Dornan 1989, Sarmandal and Grant 1990, Sijmons et al. 1989, Weiner and Robinson 1989). In a study of 310 low-risk women, Sarmandal and Grant (1990) evaluated the use of FL / AC ratio and AC in the prediction of IUGR as defined by either a ponderal index < 10th centile or a MAC / HC ratio < 10th centile. High antenatal cut-off values of AC < 25th centile and AC / FL < 25th centile were used to define abnormality. The rationale suggested by the authors for these high antenatal cut-offs was that for the purposes of screening, the largest number of fetuses should be detected. In another study of 121 patients, the ultrasonic parameters of AC < 2.5th centile, EFW > 10th centile, head circumference / abdominal circumference (HC / AC) ratio > 95th centile and the FL / AC ratio >24.0 were evaluated in their ability to IUGR, as defined by a ponderal index < 10th centile (Weiner and Robinson 1989). Different centiles were used to define abnormality for different ultrasound parameters, the reasons for which were not given by the authors. Sijmons et al. (1989) evaluated the use of umbilical artery PI to predict abnormal neonatal morphometry, as defined by either a ponderal index < 3rd centile or < 10th centile. An umbilical artery PI > 95th centile for gestational age was used to define abnormality. In another large prospective blinded study based on low-risk pregnancies, Beattie and Dornan (1989) assessed the ability of umbilical artery Doppler to predict IUGR as defined by either a low ponderal index, MAC / HC ratio or skinfold thickness. Umbilical artery Doppler waveform indices were measured at 28, 34 and 38 weeks gestation in 2097 pregnancies. However, contingency tables could not be constructed due to the lack of

relevant data reported in the paper. Statistical testing in the paper was performed using the χ^2 test, which failed to reach significance for any Doppler parameter in the prediction of abnormal neonatal morphometry.

The sensitivities, FPRs and odds ratios (and 95% confidence intervals) calculated from the data reported in the above three studies are shown in Table 2.8. The ultrasound parameters of AC < 2.5th centile (Weiner and Robinson 1989) and AC < 25th centile (Sarmandal and Grant 1990) resulted in the highest sensitivities of 81.8% and 62.0% and odds ratios of 5.6 and 7.6 respectively when used to predict a neonatal ponderal index < 10th centile. The use of AC < 25th centile (Sarmandal and Grant 1990) also had a sensitivity of 67% and odds ratio of 6.7 in the prediction of MAC / HC < 10th centile. All other ultrasound parameters had sensitivities and odds ratios which were disappointingly low. In particular, the ratios HC / AC and FL / AC in the study of Weiner and Robinson (1989) had low sensitivities and odds ratios which were not significantly different from zero. The sensitivity and odds ratio of FL / AC were also low in the study of Sarmandal and Grant (1990), suggesting that these ratios are of limited value in the diagnosis of IUGR. It was surprising that the FL / AC ratio, an indirect measure of in-utero ponderal index, fared no better than AC in the diagnosis of IUGR. Although umbilical artery PI > 95th centile (Simon et al. 1989) had reasonable odds ratios ($\geq$ 4.0), the sensitivities ($\leq$ 25%) were very low.

2.3.2.2 Limitations of previous studies

1. Comparatively few studies have reported data on ultrasound in the prediction of neonatal morphometry in high-risk and low-risk populations. Only seven ultrasound parameters, AC, EFW, HC / AC, FL / AC, fetal ponderal index and umbilical artery pulsatility index, have been evaluated in the prediction of abnormal neonatal morphometry.

2. The sensitivities and ORs of the ultrasound parameters evaluated were generally lower than when the same ultrasound parameters were used to predict birthweight. This was particularly so with AC and EFW (see Tables 2.3 and 2.8). This is not altogether surprising as a single estimate of fetal size is unlikely to confer much information on the dynamic changes occurring in growth failure (Altman and Hytten 1989).

3. The FPRs for all ultrasound parameters reviewed were very high, particularly in the low-risk population in which all ultrasound parameters had a FPR > 75%. Even though the single estimate of AC < 2.5th centile had a high sensitivity (81.8%) in the low risk population, the high FPR of 80.9% renders it a poor discriminator between normal and subnormal growth.

4. No study has reported data on the serial ultrasound assessment of AC or EFW in the prediction of abnormal neonatal morphometry. This is despite the widespread use of these measurements in clinical practice. As AC and EFW are the best ultrasound parameters for predicting SGA (Section 2.3.1), serial ultrasound values of AC and EFW may be an appropriate method of predicting SGA neonates with abnormal neonatal morphometry. Although recent studies have begun to shift the focus of antenatal ultrasound surveillance away from single estimates of fetal size towards a dynamic assessment of serial fetal measurements over time (Deter et al. 1990, Deter and Harrist 1992), the ability of serial values of AC and EFW to predict neonatal morphometry, especially in small fetuses, remains to be determined.

Table 2.7 Ultrasound measurements in the prediction of abnormal neonatal morphometry in high risk subjects.

Reference	Ultrasound Criteria	GA	Se	OR	(95% CI)
Ott (1985)	FL / AC > 23.0	20 - 42	51.5	2.8	(1.31, 6.17)
Ott (1985)	FL / AC > 23.5	20 - 42	39.4	4.3	(1.8, 10.2)
Ott (1985)	FL / AC > 24.0	20 - 42	33.3	3.3	(1.4, 7.9)
Patterson et al. (1987b)	Reduced amniotic fluid diameter	II / III	53.8	6.3	(1.6, 27.6)
Vintzileos et al. (1986)	Fetal ponderal index < 10th centile	26 - 40	76.9	14.7	(3.8, 72.1)

Abbreviation; Se, sensitivity.

Table 2.8 Ultrasound measurements in the prediction of abnormal neonatal morphometry in low risk subjects.

Reference	Ultrasound Criteria	Outcome Criteria	GA	Se	OR (95% CI)	FPR
Sarmandal and Grant (1990)	AC < 25th centile	Ponderal index	34-36	62.0	5.7 (2.5,12.7)	76.0
Sarmandal and Grant (1990)	AC < 25th centile	MAC / HC	34-36	67.0	6.7 (3.0,15.3)	75.0
Sarmandal and Grant (1990)	Fetal ponderal index < 25th	Ponderal index	34-36	51.7	3.7 (1.7,8.2)	80.0
Sarmandal and Grant (1990)	Fetal ponderal index < 25th	MAC / HC	34-36	46.7	3.1 (1.4,6.7)	80.8
Weiner and Robinson (1989)	AC < 2.5th centile	Ponderal index	28-43	81.8	7.7 (1.6,38.1)	80.8
Weiner and Robinson (1989)	EFW < 10th centile	Ponderal index	28-43	36.4	3.3 (0.9,13.0)	78.9
Weiner and Robinson (1989)	HC / AC > 95th centile	Ponderal index	28-43	16.7	2.2 (0.2,21.4)	88.9
Weiner and Robinson (1989)	FL / AC > 24	Ponderal index	28-43	55.6	3.3 (0.8,13.6)	84.8
Sijmons et al. (1989)	Umbilical PI > 1.46	Ponderal index *	28	20.0	3.9 (1.0,15.4)	86.9
Sijmons et al. (1989)	Umbilical PI > 1.46	Ponderal index	28	19.4	4.5 (1.7,12.0)	69.6
Sijmons et al. (1989)	Umbilical PI > 1.27	Ponderal index *	34	27.3	4.0 (1.0,16.4)	90.0
Sijmons et al. (1989)	Umbilical PI > 1.27	Ponderal index	34	24.1	3.8 (1.5,10.1)	76.7

All outcome < 10th centile except those with an asterisk (*) where ponderal index < 3rd centile.

Abbreviation; Se, sensitivity.

2.3.3 Ultrasound and prediction of adverse perinatal outcome.

2.3.3.1 Review of literature

Numerous studies have evaluated the use of ultrasound to predict adverse perinatal outcome. The limitations of using perinatal morbidity as an outcome criterion to define IUGR have been enumerated in Section 2.2.3.4. Nevertheless, many ultrasound studies have reported results using adverse perinatal outcome as the sole outcome measure.

To date, there has only been one randomised controlled study on ultrasound evaluation of fetal size and subsequent perinatal outcome (Secher et al. 1987a). In this study, 1570 women were screened for ultrasound evidence of a small fetus (EFW more than 85% below the expected mean) at 32 and 34 weeks gestation; the resultant 184 pregnancies were then randomised into those whose results of a subsequent scan at 37 weeks gestation were revealed to the clinician and those whose results were concealed. The results of this study are summarised in Table 2.9. There was a significantly increased incidence of induction of labour in women in the revealed group [OR 3.55 (95% CI 1.87, 6.75)] with no reduction in neonatal morbidity or mortality.

Four randomised controlled trials investigated the ultrasound assessment of fetal size in low-risk pregnancies in the third trimester and subsequent perinatal outcome (Bakkateig et al. 1984, Eik-Nes et al. 1984a, Neilson et al. 1984, Larsen et al. 1992). The data from these studies are also summarised in Table 2.9. In the study of Bakketeig et al. (1984), patients who had ultrasound measurements of biparietal diameter (BPD) at 32 weeks gestation were admitted to hospital more often than unscreened patients [OR 1.92 95% CI (1.37, 2.68)] with no reduction in the incidence of adverse perinatal outcome. In contrast, the trial conducted by Eik-Nes et al. (1984a) showed that significantly fewer patients who had ultrasound measurements of BPD and abdominal diameter (AD) at 32 weeks gestation were admitted to hospital compared with controls [OR 0.60 (95% CI 0.49, 0.75)]; this did not lead to any reduction in neonatal morbidity or mortality. No significant effects on admission rates, neonatal morbidity or mortality were noted in the study of Neilson et al. (1984) where ultrasound measurements of crown rump lengths and trunk areas were performed between 34 and 36 weeks gestation. In the recent study of Larsen et al. (1992), pregnancies in whom EFWs obtained after 28 weeks gestation were revealed to the clinicians had significantly more elective deliveries based on the diagnosis of smallness [OR 2.37 (95% CI 1.32, 4.29)] and more healthy preterm babies admitted to the NICU [OR 1.66 (95% CI 1.11, 2.49)] without any improvement in perinatal morbidity or mortality.

All the above trials suggest that isolated measurements of fetal size by ultrasound,

whether in low-risk or selected patients, did not result in any significant reduction in perinatal morbidity or mortality. Similar findings were noted in a study in which AC was used to predict subsequent compromise (defined as an ominous CTG prompting caesarean section) in 145 high-risk pregnancies with clinical suspicion of a small fetus (Chambers et al. 1989). The authors reported their results by assuming a fixed sensitivity of 100%; the specificity of AC were 12% and FPR 82%. These results suggest that the use of a cut-off value which detected all perinatal compromise would have led to an unacceptably high degree of intervention in normal pregnancies.

Danielian et al. (1992) evaluated the perinatal outcome of neonates who had a birthweight less than that predicted from an ultrasound scan performed in the third trimester. A total of 197 women who had a routine ultrasound scan between 28 and 34 weeks gestation were included in the analysis. For each fetus, EFW was calculated and the centile at that particular gestation determined using the reference ranges of Jeanty et al. (1984b). Projected birthweight was then calculated, using the nomogram of Altman and Coles (1980) assuming no change in centile between EFW and birthweight. Actual birthweight and projected birthweight were then compared and the percentage change calculated. A difference of more than 5% between projected and actual birthweight was associated with a significantly higher incidence of CTG abnormalities in labour [OR 2.54, (95% CI 1.36, 4.78)] and need for operative delivery [OR 1.94 (95% CI 1.15, 3.27)]. Whilst the five randomised controlled trials and the study of Chambers et al. (1989) addressed the predictive ability of isolated ultrasound measurements of fetal size, the study of Danielian et al. (1992) introduced the concept of change in size as a measure of fetal growth.

Numerous studies have also reported data on the ability of umbilical artery Doppler waveform indices to predict adverse perinatal outcome, both in high-risk and low-risk pregnancies. Five randomised controlled trials on the use of umbilical artery Doppler waveform indices have been reported, four based on high-risk pregnancies in the third trimester (Trudinger et al. 1987, Tyrell et al. 1990, Almstrom et al. 1992, Newnham et al. 1991) and one in low-risk pregnancies (Davies et al. 1992). The results of these studies are summarised in Table 2.10. In the study of Trudinger et al. (1987), there were no significant differences in any of the outcome measures evaluated between fetuses evaluated with umbilical artery Doppler ultrasound compared with controls. The only significant finding in the study of Newnham et al. (1991) was an increased incidence of low Apgar scores at 1-minute in the study group compared with the control group. Tyrell et al. (1990) evaluated the use of routine vs highly selective umbilical artery Doppler waveform indices as part of a modified biophysical profile; the only significant finding was a lower number of infants with a low Apgar score at 5 minutes in the group monitored routinely. The randomised study of Almstrom et al. (1992) specifically addressed the use of umbilical

artery Doppler waveform indices vs. cardiotocography in third trimester small fetuses (EFW > -2 SD below the mean). The only significant finding was a reduction in the number of fetuses requiring emergency Caesarean section for fetal distress in the group evaluated using Doppler ultrasound. The only randomized study which has evaluated the use of umbilical artery Doppler ultrasound in low-risk fetuses to date reported a significantly increased perinatal mortality rate in fetuses investigated using Doppler compared with controls (Davies et al. 1992).

These generally unfavourable results from randomized trials contrast with observational studies which have generally reported umbilical artery Doppler waveform indices to be useful in the prediction of fetal distress in labour (Lowery et al. 1990, van Vugt 1991), acidaemia at delivery (Yoon et al. 1992), necrotising enterocolitis (Hackett et al. 1987, Elwood et al. 1991), poor neonatal outcome (Berkowitz et al. 1988, Hanretty et al. 1989, Lowery et al. 1990, Maulik et al. 1990, Ferrazzi et al. 1991) and perinatal mortality (Trudinger et al. 1985). Nevertheless, a few observational studies have also reported umbilical artery Doppler to be a poor predictor of adverse perinatal outcome (Abramovich et al. 1991, Vintzileos et al. 1991).

Seven studies have specifically reported data on the ability of umbilical artery Doppler waveform indices to predict adverse perinatal outcome in small fetuses (Laurin et al. 1987, Reuwer et al. 1987, Rochelson et al. 1987, Berkowitz et al. 1988b, Chambers et al. 1989, Burke et al. 1990, Gudmundsson and Marsal 1991b). Apart from the studies of Reuwer et al. (1987) and Rochelson et al. (1987) which reported an association between abnormal umbilical Doppler indices and outcome, the other studies provided data such that sensitivities and odds ratios could be calculated. In a study of 159 small fetuses (EFW > -1.5 SD at 32 weeks gestation), umbilical artery PI was evaluated in its ability to predict operative delivery for fetal distress, acidaemia and low Apgar scores (Laurin et al. 1987). Berkowitz et al. (1988b) evaluated the ability of umbilical artery S / D ratio to predict adverse perinatal outcome (defined as operative delivery for fetal distress, respiratory distress syndrome, other neonatal morbidity related to IUGR or neonatal death) in 43 fetuses with an EFW < 10th centile. Chambers et al. (1989) evaluated 145 pregnancies with clinical suspicion of a small fetus where umbilical artery resistance index was used to predict subsequent compromise (defined as an ominous CTG prompting Caesarean section). Burke et al. (1990) evaluated the ability of the umbilical artery S / D ratio to predict Caesarean section for fetal distress and admission to NICU in a group of fetuses with an AC < 5th centile for gestational age. In a study of 142 small fetuses (EFW more than 15% below expected weight at 32 weeks gestation), the abilities of umbilical artery PI and blood flow class (a semi-quantitative measure of the amount of diastolic flow) to predict fetal distress were investigated using ROC curves (Gudmundsson and Marsal. 1991a). In a subsequent publication by the same workers using the same cohort of

fetuses, further information was provided about the ability of umbilical artery Doppler to predict operative delivery for fetal distress and low Apgar scores (Gudmundsson and Marsal. 1991b).

The results of these Doppler studies based on small fetuses are summarised in Table 2.11. Of the outcome measures evaluated by Laurin et al. (1987), the sensitivity (83.3%) and OR [43.6 95% CI (14.0, 134.6)] could only be calculated for the prediction of operative delivery for fetal distress. By contrast, Berkowitz et al. (1988b) reported a lower sensitivity (66.6%) and an OR which was not significantly different from zero. Chambers et al. (1989) fixed the sensitivity of the test at 100%, as a result of which an OR could not be calculated for this study. Burke et al. (1990) reported sensitivities between 55 and 60% and odds ratios between 3.7 and 4.1. The results reported by Gudmundsson and Marsal (1991b) suggest that whilst umbilical artery Doppler predicted subsequent operative delivery for fetal distress and low 1-minute Apgar scores, the odds ratios for the prediction of a low 5-minute Apgar score were not significantly different from zero.

Recent studies have also suggested that fetal Doppler waveform indices of the aortic, middle cerebral and renal arteries may be useful in the prediction of adverse perinatal outcome (Gudmundsson and Marsal 1991b, Arabin et al. 1992, Gramellini et al. 1992, Hecher et al. 1992). Some studies further suggest fetal Doppler ratios to be superior to umbilical artery Doppler in the prediction of adverse perinatal outcome (Arabin et al. 1992, Gramellini et al. 1992, Hecher et al. 1992). These results, if confirmed, would suggest that 'cerebral sparing', as evidenced by the low pulsatility indices in cerebral vessels compared with peripheral vessels of such fetuses, may improve the prediction of subsequent morbidity.

2.3.3.2 Limitations of previous studies

1. There is a paucity of data on the use of non-Doppler ultrasound in high-risk pregnancies to predict adverse perinatal outcome. The studies reviewed showed that single values of AC and EFW were poor predictors of adverse perinatal outcome. This is not surprising as single values of AC and EFW only convey information about fetal size and do not address the antenatal diagnosis of impaired fetal growth (Altman and Hytten 1989). Therefore, perinatal morbidity associated with IUGR is unlikely to be predicted accurately by single estimates of fetal size.

2. The study of Danielian et al. (1992) is the only one to date which has addressed the diagnosis of impaired fetal growth using change in size. There are, however, many limitations with this study. The comparison of EFW and birthweight meant that the respective centiles had to be derived from two different charts with all the problems of compatibility of the two reference standards. This could have been overcome if a second ultrasound was performed later on in pregnancy, thereby allowing serial EFW data to be evaluated against one reference standard. Furthermore, the diagnosis of IUGR could only be made retrospectively, thereby precluding the use of this method in the antenatal diagnosis of impaired growth.

3. Although recent studies have begun to shift the focus of antenatal ultrasound surveillance away from single to serial estimates of size over time (Deter et al. 1988, Deter et al. 1989a, Deter et al. 1989b), no study has evaluated the ability of serial ultrasound estimates of fetal size to predict perinatal morbidity. Furthermore, the ability of serial values of AC and EFW to predict adverse perinatal outcome in a group of small fetuses has yet to be evaluated.

4. The generally poor results from randomised controlled trials of umbilical artery Doppler in high-risk third trimester pregnancies do not confirm those from most observational studies. Nevertheless, studies confined to small fetuses generally show umbilical artery Doppler to be predictive of subsequent perinatal morbidity (Table 2.11). However, these studies reported results using different outcome criteria to define adverse perinatal outcome, making comparisons between studies difficult. Furthermore, all these studies included many neonates who delivered prematurely; perinatal morbidity in these studies was therefore not solely confined to that related to IUGR.

5. Whilst Chambers et al. (1989) compared ultrasound assessment of fetal size with umbilical artery Doppler waveform indices, no study to date has compared serial ultrasound values of AC or EFW with umbilical artery Doppler in the prediction of adverse perinatal outcome in small fetuses.

Table 2.9 Randomized controlled trials of ultrasound assessment of fetal size in third trimester and perinatal outcome.

Reference	Measurement	Outcome	Ultrasound	Control	OR	(95% CI)
Bakketeig et al. (1984)	BPD	**Low Apgar score**	34 / 510	23 / 499	1.47	(0.86, 2.51)
Eik-Nes et al. (1984a)	BPD, AD		41 / 809	35 / 819	1.20	(0.75, 1.89)
Neilson et al. (1984)	TA, CRL		37 / 433	40 / 444	0.94	(0.59, 1.51)
Secher et al. (1986a)*	EFW, AD		8 / 96	10 / 88	0.71	(0.27, 1.88)
Larsen et al. (1992)	EFW		22 / 484	22 / 481	0.99	(0.52, 1.89)
Bakketeig et al. (1984)	BPD	**Admission to NICU**	21 / 510	25 / 499	0.81	(0.45, 1.47)
Eik-Nes et al. (1984a)	BPD, AD		68 / 809	66 / 819	1.05	(0.74, 1.49)
Secher et al. (1986a)*	EFW, AD		8 / 96	8 / 88	0.91	(0.33, 2.53)
Larsen et al. (1992)	EFW		75 / 484	48 / 481	1.66	(1.11, 2.49)
Bakketeig et al. (1984)	BPD	**Perinatal deaths**	5 / 510	5 / 499	0.98	(0.28, 3.40)
Eik-Nes et al (1984a)	BPD, AD		3 / 809	7 / 819	0.45	(0.13, 1.57)
Neilson et al. (1984)	TA, CRL		0 / 433	0 / 444	1.00	(1.00, 1.00)
Secher et al. (1986a)*	EFW, AD		0 / 96	0 / 88	1.00	(1.00, 1.00)
Larsen et al. (1992)	EFW		5 / 484	3 / 481	1.67	(0.35, 8.83)

Abbreviations; AD, abdominal diameter; TA, trunk area, CRL, crown rump length.

* Study of Secher et al. (1986) was based on small fetuses (EFW < 85% of expected weight).

Table 2.10 Randomized controlled trials of umbilical artery Doppler waveform indices and perinatal outcome.

Reference	Population	Outcome	Doppler	Control	OR	(95% CI)
Trudinger et al. (1987)	High-risk	**CSFD**	3 / 127	11 / 162	0.38	(0.13, 1.13)
Tyrell et al. (1990)	High-risk*		50 / 225	39 / 215 *	1.29	(0.80, 2.07)
Almstrom et al. (1992)	SGA †		11 / 178	30 / 190 †	0.35	(0.17, 0.73)
Davies et al. (1992)	Low-risk		72 / 1238	70 / 1222	1.01	(0.72, 1.43)
Trudinger et al. (1987)	High-risk	**Apgar < 7 at 5 min.**	6 / 127	8 / 162	0.95	(0.32, 2.81)
Tyrell et al. (1990)	High-risk *		3 / 250	12 / 250 *	0.24	(0.06, .086)
Newnham et al. (1991)	High-risk		15 / 275	6 / 270	2.55	(0.97, 6.67)
Almstrom et al. (1992)	SGA†		4 / 214	5 / 212 †	0.79	(0.21, 3.02)
Davies et al. (1992)	Low-risk		9 / 1246	6 / 1229	1.48	(0.52, 4.22)
Trudinger et al. (1987)	High-risk	**Perinatal deaths §**	0 / 127	2 / 162	0.17	(0.01, 2.74)
Tyrell et al. (1990)	High-risk *		9 / 275	9 / 270 *	0.98	(0.38, 2.53)
Newnham et al. (1991)	High-risk		1 / 250	2 / 250	0.50	(0.04, 5.66)
Davies et al. (1992)	Low-risk		16 / 1246	4 / 1229	3.98	(1.31, 12.08)
Trudinger et al. (1987)	High-risk	**Admission to NICU**	27 127	38 / 162	0.88	(0.51, 1.54)
Tyrell et al. (1990)	High-risk *		18 / 250	19 / 250 *	0.94	(0.48, 1.86)
Newnham et al. (1991)	High-risk		103 / 275	106 / 275	0.93	(0.66, 1.31)
Almstrom et al. (1992)	SGA †		76 / 214	92 / 212 †	0.72	(0.48, 1.06)
Davies et al. (1992)	Low-risk		44 / 1246	43 / 1229	1.01	(0.65, 1.55)

Abbreviations; CSFD, Caesarean section for fetal distress.

* Randomized controlled trial of routine vs highly selective use of umbilical artery Doppler waveforms and biophysical profile.

† Randomized controlled trial of umbilical artery Doppler waveforms vs. cardiotocography in small fetuses (EFW more than 2 SD below mean).

§ Perinatal deaths excluding congenital malformations.

Table 2.11 Doppler ultrasound in the prediction of adverse perinatal outcome in small fetuses.

Reference	Doppler	Prevalence	Outcome	Se	OR	(95% CI)
Laurin et al. (1987)	Umbilical PI ≥ 2 SD	30 / 156	ODFD	83.3	43.6	(14.0,134.6)
Berkowitz et al. (1988b)	Umbilical S / D > 3.0	21 / 43	Adverse outcome	66.7	3.3	(0.9,12.2)
Chambers et al. (1989)	Umbilical RI > 2 SD	24 / 145	CSFD	100*	-	-
Burke et al. (1990)	Umbilical S / D > 2 SD	53 / 179	Admission to NICU	52.8	4.1	(2.1, 8.2)
Burke et al. (1990)	Umbilical S / D > 2 SD	10 / 179	CSFD	60.0	3.7	(1.0, 13.8)
Gudmundsson and Marsal (1991b)	Umbilical PI > 2 SD	39 / 139	ODFD	82.0	41.1	(14.3,118.5)
Gudmundsson and Marsal (1991b)	Umbilical PI > 2 SD	19 / 139	1-min Apgar ≤ 7	63.1	4.0	(1.4,11.1)
Gudmundsson and Marsal (1991b)	Umbilical PI > 2 SD	6 / 139	5-min Apgar ≤ 7	50.0	2.32	(0.4,12.2)

* Authors assumed a sensitivity of 100%.

Abbreviations; RI, resistance index; Se, sensitivity; CSFD, Caesarean section for fetal distress; ODFD, operative delivery for fetal distress.

2.3.4 Ultrasound and prediction of abnormal biochemical indices of IUGR

2.3.4.1 Review of literature

Biochemical indices used to define IUGR have been discussed in Section 2.2.5. Most of the literature on ultrasound and the prediction of abnormal biochemical indices has been based on blood biochemistry obtained at cordocentesis.

In a study of fetuses with an AC < 5th centile for gestational age, the oxygen tension in venous blood obtained at cordocentesis was below the mean for gestational age in 33 out of 38 fetuses (Soothill et al. 1987a). Fourteen of these fetuses had oxygen tensions levels more than two standard deviations below the mean. The severity of hypoxia also correlated with venous glucose levels. In a cordocentesis study of 208 normal sized fetuses and 196 small fetuses (AC > -2 SD for gestational age), the latter were significantly more likely to be hypoxic, acidaemic and hyperlacticaemic (Nicolaides et al. 1989).

Five further studies based on biochemical indices obtained at cordocentesis have shown that within a group of small fetuses, umbilical and fetal Doppler waveform assessment can identify those with hypoxia and low pH (Soothill et al. 1986, Nicolaides et al. 1989, Vyas et al. 1988, Vyas et al. 1989, Bilardo et al. 1990). In a study based on 29 fetuses with an AC < 5th centile for gestational age, Soothill et al. (1986) reported an association between reduced flow in the fetal aorta, as ascertained by Doppler ultrasound, and hypoxia, acidaemia and hyperlacticaemia at cordocentesis. In a study of 59 fetuses with AC < 5th centile for gestational age and absent end diastolic frequencies in the umbilical artery Doppler waveforms, Nicolaides et al. (1988) found the majority of these fetuses to be hypoxic and acidaemic. Subsequent studies on the aortic, common carotid (Bilardo et al. 1990), renal (Vyas et al. 1989) and middle cerebral Doppler waveforms (Vyas et al. 1990) have all suggested that fetal Doppler waveforms which reflect the brain-sparing effect of hypoxia (Peeters et al. 1979) are useful in the identification of small fetuses with hypoxia.

Other studies based on venous blood obtained at cordocentesis have shown that fetuses with an AC < 2.5th centile are more likely to be hypoinsulinaemic (Economides et al. 1989b), hypertriglyceridaemic (Economides et al. 1988) and have raised non essential / essential amino acid levels, when compared with normal sized fetuses (Economides et al. 1989a). Cetin et al. (1988) also reported differences in the amino acid profiles of small and normal sized fetuses in umbilical arterial and venous blood obtained at delivery. However, although Cetin et al. (1988) classified fetuses to be small or normal sized by serial ultrasound assessment, they did not report which ultrasound parameters were used for

diagnosis.

Hawdon et al. (1992b) evaluated the ability of umbilical artery Doppler waveform indices to predict neonatal hypoglycaemia in 25 fetuses with an EFW > -2 SD for gestational age. Fetuses were classified into those with (n = 14) or without end diastolic flow (n = 11) in the umbilical artery Doppler waveform. Although umbilical venous glucose levels obtained at delivery were not significantly different between the two groups, glucose levels obtained at heel-prick samples at 6 hours after birth were significantly lower in neonates with absent end diastolic flow.

2.3.4.2 Limitations of previous studies

1. Inspection of the data from all the cordocentesis studies reviewed showed that not all small fetuses had biochemical indices suggestive of asphyxia or malnutrition. This is not surprising as a substantial proportion of small fetuses are not growth retarded (Patterson and Pouliot 1987a) and a single estimate of smallness would not discriminate those with or without IUGR.

2. Many of the studies reviewed divided fetuses into those who were small and those who were normal sized. Although biochemical differences were demonstrated between these two groups, there are numerous inherent problems of using size to define IUGR. The biochemical profile of those fetuses purported to reflect malnutrition may just be a manifestation of smallness rather than impaired fetal growth per se.

3. The assessment of umbilical and fetal Doppler waveforms represented a step forward in subdividing small fetuses into those with or without IUGR. However, the numbers in each of these studies were small ($n \leq 56$) and results were presented in a form which precluded any calculations of sensitivities or ORs.

4. The use of data obtained at cordocentesis to confirm IUGR is unlikely to be applicable in general clinical practice. Only two ultrasound studies have used biochemical indices of malnutrition obtained at delivery to define IUGR (Cetin et al. 1988, Hawdon et al. 1992b). The numbers involved in these two studies were also small (n = 25 in each study).

5. No study has addressed the ability of serial ultrasound estimates of fetal size to predict biochemical indices at delivery indicative of IUGR.

CHAPTER 3

FACTORS INFLUENCING STUDY DESIGN

3.1 Need for further study

3.1.1 Diagnosis of IUGR: Serial ultrasound of AC and EFW

It is apparent from Chapter 1 that not all small fetuses are growth retarded (Patterson and Pouliot 1987a, Villar et al. 1990, Fay et al. 1991a). A proportion of small fetuses are constitutionally small with no evidence of wasting or malnutrition (Walther and Raemaker 1982, Hill et al. 1984). Neonatal morphometric indices of malnutrition such as ponderal index, MAC / HC ratio and skinfold thickness have been shown to be the best method of confirming IUGR in the newborn (Section 2.2.5). Successful antenatal separation of small fetuses into those with evidence of growth failure and those with normal growth would allow a rational approach to antenatal surveillance and possible need for intervention.

It is apparent from the literature review in Section 2.3.1 that AC and EFW are superior to all other ultrasound parameters in the prediction of SGA. By inference, serial ultrasound assessment of AC and EFW may be useful in the diagnosis of IUGR within a group of small fetuses. The antenatal diagnosis of IUGR requires serial measurements of fetal size to demonstrate that a fall-off in the growth trajectory has occurred (Altman and Hytten 1989, Deter et al. 1990, Deter and Harrist 1992). The review in Section 2.3.2 showed that only eight studies have reported the use of ultrasound to predict abnormal neonatal morphometry, all of which were confined to single ultrasound assessments of size. None reported the use of serial values of AC or EFW. Similarly, no study has previously reported the use of serial ultrasound estimates of fetal size to predict adverse perinatal outcome (Sections 2.3.3) or biochemical indices associated with IUGR (Section 2.3.4).

In view of this, a study was constructed to evaluate the ability of serial values of AC and EFW to predict abnormal neonatal morphometry in a group of small fetuses. The usefulness of serial values of AC and EFW was to be compared with single estimates of fetal size and umbilical artery Doppler waveform indices, standard ultrasound parameters currently used to evaluate small fetuses with suspected IUGR. Finally, the clinical significance of separating small fetuses antenatally into those with ultrasonic normal growth vs. IUGR would be assessed by comparing perinatal and biochemical outcome in the two groups postnatally.

3.1.2 Derivation of reference standards for AC and EFW

3.1.2.1 Introduction

The use of serial ultrasound values of AC and EFW to describe fetal growth requires reference standards which are appropriately constructed, preferably from longitudinal data (Deter et al. 1982a, Evans et al. 1990, Deter and Harrist 1992). However, it will be apparent from the following critical review of published reference standards for AC and EFW that no single reference standard has been optimally derived using appropriate statistical methods.

3.1.2.2 Limitations of previous reference standards for AC

Despite the many reference standards published for AC (Campbell 1976, Hoffbauer et al. 1979, Weinraub et al. 1979, Tamura and Sabbagha 1980, Meire and Farrant 1981, Deter et al. 1982a, Deter et al. 1982b, Fescina et al. 1982, Hadlock et al. 1982a, Deter et al. 1984, Jeanty et al. 1984a, Woo et al. 1984, Larsen et al. 1990), the British Medical Ultrasound Society (BMUS) Bulletin (Evans et al. 1990) recently highlighted the lack of a single optimally derived reference standard for this ultrasound parameter. The limitations with these studies are summarized below;

1. Data should be used in the construction of reference standards only if there is reliable information on gestational age assessment, either by menstrual dating or by early ultrasound assessment (Geirsson 1991). Two of the studies did not report the method of assessing gestational age (Campbell 1976, Hoffbauer et al. 1979).

2. The method of measuring abdominal circumference should be standardised and clearly reported in the literature, as it has been shown that measurements obtained by tracing the circumference are different from those obtained using abdominal diameters (Woo et al. 1984). No information was given on the method of measurement in two studies (Meire and Farrant 1981, Hadlock et al. 1982a). In three studies, the AC was determined from the diameters rather than by direct measurement (Fescina et al. 1982, Jeanty et al. 1984a, Larsen et al. 1990).

3. Ultrasound data for reference standards should either be collected cross-sectionally or longitudinally, but not a mixture of the two. In five studies (Hoffbauer et al. 1979, Weinraub et al. 1979, Tamura and Sabbagha 1980, Meire and Farrant 1981, Fescina et al. 1982), data from the same fetus were used more than once during cross-sectional data

collection. In the study of Hadlock et al. (1982a), insufficient information was given about the method of data collection.

5. Of the three longitudinal studies on AC (Deter et al. 1982a, Jeanty et al. 1984a, Larsen et al. 1990), two (Jeanty et al. 1984a, Larsen et al. 1990) reported data which were inappropriately analysed by regression analysis as though the data had been obtained cross-sectionally. The use of regression analysis in the analysis of longitudinal data is inappropriate as it does not take into account the inter-dependence of repeated measurements on the same fetus (Deter et al. 1982a). Such a method of analysis would result in an inappropriately low estimate of residual variance.

6. The reference ranges reported by all the above studies show remarkable similarities from 10 to 38 weeks gestation. However, there is much discrepancy of measurement means beyond 38 weeks gestation. The two most commonly used cross-sectional standards in clinical practice (Deter et al. 1982b, Hadlock et al. 1982a) demonstrate significant differences at term. The curve for Hadlock et al. (1982a) shows a plateau whilst that for Deter et al. (1982b) does not, primarily related to the fact that the former used a quadratic model to describe the data whilst the latter used a linear model. Similar discrepancies were evident when comparing AC curves derived from longitudinal data. Jeanty et al. (1984a) and Larsen et al. (1990) found that their AC data were best described by a cubic and fourth order polynomial model respectively while Deter et al. (1982a) reported similar R^2 values using a linear model.

3.1.2.3 Limitations of previous reference standards for EFW

Although EFW is commonly used to predict fetal size, there has been much confusion as to the appropriate reference standards to use (Gardosi et al. 1992). Until recently, many "intrauterine growth charts" used as reference standards for ultrasonic EFW were not derived from intrauterine data, but were in fact reference standards based on birthweight data. Such reference standards differ markedly from those derived from ultrasonic EFW data. This is because derivation of standards using birthweight data necessitates the use of birthweights of neonates who deliver prematurely. The limitations of such birthweight charts and other EFW charts derived from ultrasound measurements are discussed:

1. The inappropriateness of using birthweight data to define fetal growth is highlighted by many studies which have shown that a higher proportion of preterm infants are growth retarded compared to infants born at term. In about 50% of premature labours, underlying conditions exist which cause uteroplacental insufficiency (Adelstein and Fedrick 1976).

Many obstetric risk factors associated with preterm delivery are also associated with IUGR (Kaminski and Papiernik 1974). This is also supported by the observation that birthweight charts based on data from selected low-risk pregnancies show higher weights for gestation than those based on the high risk population (Ulrich 1982). At least four studies have reported significant discrepancies between ultrasonically determined EFW standards and those determined from birthweight data, particularly at gestational ages less than 37 weeks (Tamura et al. 1984, Weiner et al. 1985a, Bottoms et al. 1992, Bernstein et al. 1992). All the above data suggest that pregnancies ending in preterm deliveries cannot be considered normal, and therefore EFW charts should be derived using ultrasonically determined EFW data.

2. Most ultrasound-derived reference standards for EFW have been derived from longitudinally collected data. Apart from the study of Deter et al. (1982a), other longitudinal studies have analysed their data inappropriately as though collected cross-sectionally (Jeanty et al. 1984b, Persson and Weldner 1986, Larsen et al. 1990). Deter et al. (1982a) are the only workers who have analysed their longitudinal data using least squares fitting to obtain growth curves for individual fetuses. However, sample sizes in this study and the study of Person and Weldner (1986) were small (n = 20 and n = 19 respectively). There was a paucity of data after 38 weeks gestation in these studies as a result of earlier delivery in many fetuses, thereby limiting the usefulness of these reference standards.

3. There were significant discrepancies in the reference standards reported in different studies. These were most marked between the reference ranges reported by Deter et al. (1982a) and Jeanty et al. (1984b). Although mean birthweight of the infants in the study of Deter et al. (1982a) was similar to that in the study of Jeanty et al. 1984a), mean fetal weights after 36 weeks gestation reported by the former were consistently higher than those reported by the latter. This is all the more surprising as study of Jeanty et al. (1984b) excluded data from SGA infants. No differences were noted between the fetal weight curves of Larsen et al. (1990) and Persson and Weldner (1986) but the mean fetal weights between 30 and 34 weeks obtained by Larsen et al. (1990) were 1.5% lower those that obtained by Deter et al. (1982a).

5. As has already been observed, the formulae used to describe EFW in all four longitudinal studies of EFW were non-linear; Deter et al. (1982a) and Larsen et al. (1990) concluded that a quadratic formula best described changes in EFW whilst Jeanty et al. (1984b) and Persson and Weldner (1986) used a cubic formula. These results contradict the AC reference standards reported by Deter et al. (1982a) who reported growth to be linear at all gestational ages.

3.1.2.4 Need for appropriately derived reference ranges for AC and EFW

It was therefore apparent from the preceding sections that not a single reference standard for either AC or EFW had been derived using appropriate statistical methods in a sizeable group of fetuses. The linearity of growth of AC reported by Deter et al. (1982a) is in disagreement with all other cross-sectionally and longitudinally derived reference standards for AC and EFW. Despite the small number of fetuses studied (n = 20), these reference standards have been recommended by the BMUS Bulletin (Evans et al. 1991) as the reference standard of choice. Further verification is therefore needed on the growth profile of fetuses especially towards term.

Whether reference standards for evaluation of serial ultrasound measurements should be derived from cross-sectionally or longitudinally collected data remains a matter of considerable debate (Bland and Altman 1992). Reference standards derived from cross-sectional data may be adequate for assessment of fetal size (Sparks and Cetin 1991, Bland and Altman 1992). However, there are theoretical disadvantages with the use of cross-sectional data in the assessment of growth (Deter et al. 1982a, Evans et al. 1990, Deter and Harrist 1992). Although cross-sectional AC and EFW data can be used to establish centiles at a given gestational age, it may not be legitimate to establish a dynamic measurement of rate of change of AC and EFW based on differences between static measurements of different fetuses (Sparks and Cetin 1991). The advantage with longitudinally collected data is that it is a true representation of fetal growth, with dynamic changes in each individual fetus being assessed over time. Longitudinal data, collected at regularly spaced intervals, should be subjected to a statistical analysis which takes into account the inter-dependence of measurements on the same fetus (Deter et al. 1982a, Bland and Altman 1992). This is because longitudinal data cannot be considered to be a set of independent observations, as repeated measurements on the same fetus would introduce correlation between measurements (Deter et al. 1982a).

There was therefore a need for reference standards of AC and EFW to be produced based on the appropriate statistical analysis of longitudinal ultrasound data in a sizeable group of fetuses.

3.1.3 Description of serial values of AC and EFW

3.1.3.1 Limitations of previous studies

Before serial ultrasound values of AC and EFW can be used in the diagnosis of IUGR, these serial changes need to be quantified. With the exception of the study of Deter et al. (1982a), previous studies have demonstrated the non-linearity of fetal growth towards term. If growth is confirmed to be non-linear, gestational age-independent changes in absolute measurements of AC or EFW cannot be used to quantify growth.

Few studies have addressed the quantification of serial ultrasound values of AC or EFW. One previous study evaluated fetal growth using simple increments in AC values with gestation (Divon et al. 1986). In this study of 90 fetuses, an abnormal rate of AC growth was defined as ≤10 mm / 14 days. The choice of 10 mm / 14 days in this study was an arbitrary cut-off level based on the ability to distinguish between small and normal sized fetuses rather than on published standards of rate of growth of AC. The major problem with this study was the quantification of growth using a change in absolute measurements of AC, inappropriate in view of the non-linearity of fetal growth at term.

In order to overcome the problems of quantifying serial ultrasound measurements given that fetal growth is not linear, other workers have reported the derivation of individual growth curves using the Rossavik Growth Model (Deter et al. 1984, Deter and Rossavik 1987, Deter et al. 1990, Deter and Rossavik 1992). The general equation for this model is: $P = c\,(t)^{k + s\,(t)}$ where P is the ultrasound parameter, k is a fixed coefficient determined by the anatomical parameter, c is related to genetic regulators of growth, s is an unknown regulatory system that modifies genetically determined growth and t is the duration of growth of the parameter (Rossavik and Deter 1984). The value k is suggested to be a fixed value for a specific ultrasound parameter. Appropriate values for k have been established for AC and EFW by regression analysis from serial scans obtained every 2 weeks from 12 to 26 weeks. The advantage of this model is that each fetus can act as its own control and that by performing 2 ultrasound scans before 27 weeks gestation, the coefficients c and s can be determined for that individual fetus (Simon et al. 1987). The individual growth curve for that particular fetus can then be derived and any subsequent deviation from that curve is regarded as failure to achieve its growth potential (Deter et al. 1989a, Deter et al. 1990). It also has the potential advantage that each fetus acts as its own control and that assessment of growth is therefore "individualised".

The Rossavik model is the only statistical method that has been reported for the quantification of serial fetal measurements. However, there are marked limitations with model which severely limit its use in routine clinical practice:

1. The model depends on a constant value for the coefficient k for a particular ultrasound parameter. These values of k were derived from a small group of 20 middle-class mothers of different ethnic backgrounds (17 Caucasian, 2 Blacks, and 1 Hispanic) in Houston, Texas, USA (Deter et al. 1987). The women were chosen because they delivered at term and the infants had no abnormalities on paediatric and morphometric assessment. It is of considerable doubt as to whether such values of the coefficient k would be similar in other obstetric populations in other countries. As yet, no other group has derived separate values for k based on their own indigenous obstetric populations. The only other group of workers who have independently verified the use of this model reported a significant systematic over-prediction of AC and EFW in the third trimester using this method (Simon et al. 1987). It appears that further work needs to be done to verify the validity of this model, especially in the derivation of the coefficient k.

2. The derivation of growth curves for individual fetuses requires the data from two scans performed before 27 weeks gestation. The inherent assumption with this model is that fetal growth is normal before 27 weeks gestation, and that any growth deviation occurs subsequently (Deter et al. 1990). Whilst such a simplistic assumption may be true when using this model to detect late onset (ie. third trimester) IUGR, this will not be applicable to fetuses with early onset IUGR. Furthermore, the model requires the results of two scans performed before 27 weeks, information rarely available in clinical practice.

3. Possibly the most important limitation with the Rossavik Growth Model is the lack of any prospective data on the degree of deviation from the growth curve before fetuses are deemed to be growth retarded. To date, the model has not been validated against standard neonatal morphometry or measures of perinatal outcome. Whilst two studies (Deter et al. 1990, Ott 1990a) have used the model to predict morphometric evidence of IUGR in the neonate, neither study defined the antenatal criteria for abnormal growth nor used accepted morphometric criteria to define IUGR postnatally. Further prospective studies need to be carried out to evaluate this model against standard outcome measures of IUGR.

3.1.3.2 Need to quantify serial values of AC and EFW

In view of the paucity of data on how best to describe serial values of AC and EFW, numerous statistical methods of quantifying change in AC / EFW with gestation have to be evaluated against neonatal morphometry, the "gold standard" for IUGR. The best statistical method of quantifying serial measurements would then be determined. This optimal measure of serial AC / EFW data would then be compared with other standard ultrasound measures, such as umbilical artery Doppler waveform indices, currently used to evaluate such fetuses.

3.2 Outline of studies to be described

3.2.1 Studies

Study 1 Derivation of normal reference ranges for AC and EFW (Chapter 5)

Study 2 Reproducibility of ultrasound and neonatal morphometric measurements (Chapter 6)

Study 3 Diagnosis of IUGR using serial ultrasound measurements (Chapter 7):
a) Quantification of serial ultrasound measurements of AC and EFW.
b) Comparison of the optimal measure of serial AC / EFW data with umbilical artery PI and estimates of fetal size in the prediction of abnormal neonatal morphometry and adverse perinatal outcome in small fetuses.
c) Ability of the optimal measure of serial AC / EFW data to separate small fetuses into two groups with distinctly different perinatal outcome and biochemical data at delivery.

CHAPTER 4

PRINCIPLES AND METHODOLOGY

4.1 Ultrasound Measurements

4.1.1 B-mode Ultrasound Imaging

4.1.1.1 History of B-Mode Ultrasound Imaging

The use of ultrasound as a diagnostic tool in medicine began after the end of the Second World War. Donald and Brown (1961) were the first to report the use of ultrasound in obstetrics to demonstrate tissue interfaces within the uterine cavity. Campbell (1969) first reported the use of ultrasound to evaluate fetal anthropometry when he used measurements of BPD to evaluate gestational age. The use of serial ultrasound measurements of BPD were first introduced by Campbell and Newman (1971) to evaluate fetal growth. Subsequently measurements of other parts of the fetus were reported; Campbell and Wilkins (1975) first reported the calculation of estimated fetal weight from measurements of abdominal circumference.

4.1.1.2 Physical properties and principles

A sound wave is a series of compressions and rarefactions. The combination of one compression and one rarefaction is one cycle and the distance between one cycle to the next is the wavelength. The velocity represents the speed with which sound waves travel through a particular medium and it is equal to the product of frequency and wavelength. Thus frequency and wavelength are inversely related. Ultrasound is sound with a frequency greater than 20,000 cycles / second or Hertz (Hz) which is above the audible range. Most ultrasound instruments in clinical use employ frequencies in the range of 1 - 10 MHz.

Sound waves travel in tissue at a speed which depends on the physical properties of the tissue. The velocity of sound is fairly constant for human soft tissue, being approximately 1,540 m / sec (Eldridge et al. 1983). However, there is a significant difference in velocity if sound passes through solid structures such as bone. When a ultrasound reaches an interface between two tissues of different acoustic properties (impedances), the beam undergoes reflection and refraction. The amount of speed that reflected depends on the degree of acoustic mismatch between the two tissues, the angle the beam strikes the interface and the relative size of the mismatched tissue and the wavelength. The total thickness presented to the ultrasound beam must be at least a quarter of the wavelength of the ultrasound beam for ultrasound to be reflected. Ultrasound with a high frequency can reflect sound from smaller objects and therefore has a high resolving power. Sound with a frequency of 2 MHz permits the recording of distant echoes from

interfaces that are approximately 1 mm apart.

The amplitude (intensity) of an ultrasound wave is reduced (attenuated) as the distance of travel increases. The extent of this depends on the transmitted frequency, energy absorption by the tissues and the amount of reflection by the tissue interfaces. Very high frequency ultrasound is reflected by many small interfaces and therefore less energy is available to penetrate deeply into the body. Thus, the penetration of the beam decreases as the frequency increases.

In obstetric ultrasound, structures in the fetus can be measured using dynamic ultrasound imaging from two-dimensional, or B-mode images. Ultrasound energy is generated by a piezoelectric transducer placed on the maternal abdomen. As the piezoelectric transducer expands and contracts, it produces compressions and rarefactions or sound waves. The crystal also acts as a receiver and is able to detect a signal even if less than 1% of the ultrasound energy is reflected. The acoustic resolution depends on the space occupied by the pulse as it propagates. In most obstetric imaging, the pulse shape produces an axial resolution that is superior to the lateral resolution over most of the beam path. Another approach is to use a large-aperture, single-element transducer in order to achieve a lengthy focal zone relatively far away from the transducer. Recent ultrasound machines have combined relatively large apertures, multielement arrays and computer processing so that the focal zone can be changed dynamically by the form and timing of element excitation during the cycle and by the way the element inputs are combined during reception.

The transducers used in obstetric ultrasound are either linear, curvilinear or sector scanners. As long as the part of the fetus to be measured, for example the femur length, is oriented horizontally with the ends in the focal zone and with a system gain set relatively low, measurements made by different transducers should not differ (Birnholz 1986). Once the optimal image is obtained on the screen, measurements can be made using electronic calipers integrated into the system.

4.1.1.3 Method of measurement

The same methodology of measurement was consistently used throughout all the studies and are described here in detail. All these measurements were performed on an Accuson KXP 128 / 1 ultrasound machine using a 3.5 MHz linear transducer (L312) with an Aperture size of 120 mm. Measurements were performed with the patient in a comfortable, semi-recumbent position.

Acoustic coupling gel was applied to the maternal abdomen and the transducer placed on the abdomen. **Biparietal diameter** was measured along a transverse plane of the fetal head, with the thalami, mid-line and cavum septum pellucidum all displayed in the same plane (Hadlock et al. 1982b). Having identified an appropriate section, the image was frozen and the measurement made from the outer edge of the proximal skull surface to the inner edge of the distal skull surface using electronic calipers on the screen. Three measurements were made, and the average of these was taken as the BPD.

Head circumference was measured on the same frozen image used for the measurement of BPD. This was measured directly on screen by tracing around the outer edge of the circumference of the image using the tracker-ball. The average of three measurements was taken as the HC.

Abdominal circumference was measured by first obtaining a longitudinal section through the fetal spine and aorta. The transducer was then rotated through 90 degrees to obtain a transverse image of the fetus at the level where the umbilical vein entered the portal system of the liver. The transverse section should be circular in outline; the outline of the aorta and fetal spine should also be circular to confirm that the plane was perpendicular to the long axis of the fetus (Deter et al. 1982b). Once the correct plane was identified, the image was frozen and AC was measured directly on the screen by tracing around the outer edge of the image using the tracker-ball. The average of three measurements was then used to calculate AC.

Femur length was measured between the two ends of the femoral diaphysis (Warda et al. 1985). To ensure that the whole of the femur was measured and that it was not foreshortened, the transducer was rotated until the longest possible image of the femur was obtained and the transducer was along the long axis of the femur, thereby producing a femur image with clear blunt ends. The blunt ends correspond to the femoral diaphyses. The distal femoral epiphysis, which ossifies late in pregnancy, was not included in the ultrasonic measurement. The image was then frozen and a straight line measurement was made between the two ends of the femoral diaphysis. This was repeated three times, the average of which was taken to be the FL.

Estimated fetal weight was calculated using the four-parameter formula of Hadlock et al. (1985); $Log_{10}(EFW) = 1.5115 + 0.0436\ (AC) + 0.1517\ (FL) - 0.00321\ (AC.FL) + 0.0006923\ (BPD.HC)$. This calculation was made using the average of each of three measurements of BPD, HC, AC and FL.

4.1.1.4 Discussion of methods

To standardise ultrasound measurements, all ultrasound scans were performed on an Accuson KXP 128 / 1 ultrasound machine. This machine was chosen because it provided high quality B-mode ultrasound imaging and Doppler facilities. A 3.5 MHz linear transducer was chosen to perform all ultrasound measurements as it gave a good image quality of fetal anthropometry commensurate with fetal size in the third trimester.

At least three methods of measuring BPD have been described in the literature. Hadlock et al. (1982b) and Sabbagha (1989) measured BPD from the outer- to the inner-table, whilst Shepard (1982a) measured BPD from the middle of the proximal table to the middle of the distal table. Others measured BPD from the outer- to the outer-table. The methodology reported by Hadlock et al. (1982b) was chosen as it was the most widely accepted method of measuring BPD and recommended by BMUS (Evans et al. 1990) for general use. More importantly, this method of measuring BPD was the same used to measure BPD for the purposes of calculating EFW using the formula of Hadlock et al. (1985).

There is much less discrepancy in the literature with regards to the methodology of ultrasound measurements of HC, AC and FL. In view of the discrepancies noted between abdominal circumferences measured directly on screen and those calculated from abdominal diameters (Woo et al. 1984), the methodology of measurement of AC used by Hadlock et al. (1985) in the calculation of EFW was adopted.

At least three measurements were made, as this was an important means of reducing random measurement errors in obstetrics (Birnholz 1986). This was especially important with certain structures such as the femur length where a slight misalignment of the scan plane would have introduced a substantial error such as a shortening of the femur length.

Numerous formulae based on different ultrasound parameters have been reported for the calculation of EFW (Table 4.1). The initial formulae incorporated BPD and AC measurements (Warsof et al. 1977 and Shepard et al. 1982b). Subsequent formulae have additionally incorporated FL and HC and improved the predictive accuracy in studies of unselected fetuses (Hadlock et al. 1984, Ott and Doyle 1984, Hadlock et al. 1985). This was because FL was an indirect measurement of fetal crown-heel length, and the addition of HC reduced errors of head measurement due to altered head shape, like dolichocephaly or brachycephaly (Hadlock et al. 1984). Of these studies, only the study of Hadlock et al. (1985) tested the formulae prospectively in another group of fetuses separate from the

group used to derive the formula.

The theoretical disadvantage with the above formulae was their derivation from the data of fetuses of all sizes, taking no account of whether the fetuses were small or normal sized. However, changes in fetal morphologic characteristics are dynamic; the fetal HC / AC ratio is purported to increase progressively in fetuses with asymmetric growth retardation (Campbell and Thom 1977). When these general formulae were applied to a population of small fetuses, systematic and random errors were large (Simon et al. 1987). In that study, formulae which incorporated BPD and AC (Shepard et al. 1982b), BPD, AC and FL (Hadlock et al. 1985) and BPD, HC, AC and FL (Hadlock et al. 1985) all produced a systematic error in overestimating birthweight in SGA infants, with the formula of Shepard et al. (1982b) producing the smallest systematic error. Such non-targeted formulae may be less accurate than formulae targeted at small fetuses because they do not take into account changing fetal morphology with gestational age and size / growth deviations (Sabbagha et al. 1989). At least 5 studies have evaluated the use of targeted formulae in estimating fetal weight in preterm fetuses with birth-weight less than 2000g.The formula of Weiner et al. (1985b) resulted in the smallest mean errors (Pielet et al. 1987). However, preterm fetuses are not similar to small fetuses and formulae derived from such fetuses are therefore not applicable to the SGA population.

Only one study reported a targeted formula derived solely from small fetuses. In this prospective study (Sabbagha et al. 1989), the targeted formula was compared with the formula of Hadlock et al. (1985) in a group of 70 small fetuses in the estimation of birthweight. Although the use of the formula of Hadlock et al. (1985) resulted in a significant systematic error (systematic over-estimation of birthweight by not more than 5%), the random errors associated with the two formulae were not statistically different. There are, however, two particular limitations with the targeted formula reported by Sabbagha et al. (1989). First, unlike the study of Hadlock et al. (1985), the birth weight distribution of the fetal population used to derive the targeted formula for small fetuses was not reported. Ideally a stratified sample should be used with approximately equal numbers of observations in each weight class over the range for which the formula is intended (Persson 1989). Second, Sabbagha et al. (1989) compared birth weight predictions using the mean percentage error and the absolute 2 SD values. They claimed the latter value reflected the random variation of the percentage errors. However this was inappropriate as deletion of the sign of the errors reduced the variance and hence underestimated the spread of the errors. The variance of the (signed) percentage errors is better summarized by calculating the 95% prediction intervals of the mean percentage error. A prospective study evaluating these two as well as other formulae in small fetuses is needed.

The optimal formula for use in small fetuses should have the lowest mean percentage error and residual random variation with no systematic error. Neither the 4-parameter formula of Hadlock et al. (1985) nor the targeted formula of Sabbagha et al. (1989) fulfilled all these criteria. The four-parameter formula of Hadlock et al. (1985) was chosen to calculate EFW in preference to that reported by Sabbagha et al. (1989) in view of its simplicity of use.

Table 4.1 Various formulae used to calculate EFW

Measurements	**Reference**	**Equation**
GA, AC, HC, FL	Sabbagha et al. (1989)	$EFW = 1849.4 - 47.13\ (SUM) + 0.37721\ (SUM)^2$ where $SUM = GA + HC + 2AC + FL$
BPD, HC, AC, FL	Hadlock et al (1985)	$Log_{10}(EFW) = 1.5115 + 0.0436\ (AC) + 0.1517\ (FL) - 0.00321\ (AC.FL) + 0.0006923\ (BPD.HC)$
BPD, AC	Warsof et al. (1977)	$Log_{10}(EFW) = -1.599 + 0.144\ (BPD) + 0.032\ (AC) - 0.111\ (BPD^2.AC)/1000$
BPD, AC	Shepard et al. (1982b)	$Log_{10}(EFW) = -1.7492 + 0.166\ (BPD) + 0.046\ (AC) - 2.646\ (AC.BPD)/1000$
HC, AC, FL	Weiner et al. (1985b)	$Log_{10}(EFW) = 1.6961 + 0.02253\ (HC) + 0.01645\ (AC) + 0.06439\ (FL)$

4.1.2 Umbilical artery Doppler waveform indices

4.1.2.1 History of Doppler ultrasound

The Doppler effect is named after its first describer, Christian Johann Doppler (1803 - 1853), an Austrian physicist. In 1842, he described how the colour of the light emitted from a star changes depending on the direction and velocity of movement of the star relative to the observer on earth. The Doppler principle was first introduced into medicine for the measurement of velocity of blood flow in the evaluation of cardiac function (Satomura (1957). By emitting sound towards a vessel, it was possible to elicit information about velocity distribution. The first crude blood velocity results were obtained using Doppler instruments in the continuous mode (Satomura 1957). Later, Baker (1970) introduced the pulsed Doppler mode.

The first attempts to analyse human fetal blood flow by Doppler ultrasound were made by Fitzgerald and Drumm (1977). McCallum (1978) reported that the flow pattern in the umbilical artery of fetuses in normal pregnancies differed from the flow pattern in fetuses of pre-eclamptic mothers. In 1979, Gill and Kossoff quantified blood flow in the umbilical vein using a B-mode scanner. Eik-Nes et al. (1980) published a method which used a linear B-mode scanner in combination with pulsed Doppler fixed to the linear transducer to quantify blood flow.

4.1.2.2 Physical properties and principles

The Doppler principle is that the frequency of oscillation an observer measures is affected by relative movement between the observer and the source of the oscillation. Doppler ultrasound in obstetrics is a slight modification of that principle in that the source and receiver of the ultrasound (the transducer) are stationary and the ultrasound is bounced off the red cells in the vessels. The random changes in the density of groups of red cells in the plasma scatter the ultrasound. The difference between the transmitted and received frequencies is termed the Doppler shift (f_d) and is related to the velocity (v) of the reflector by the equation;

$f_d = (2.f.v. \cos \theta) / c$

where f = frequency of the transmitted ultrasound

c = velocity of sound in the medium being examined

θ = angle between the ultrasound beam and the moving target.

As f and c are constants in any given situation, it follows that the Doppler shift is directly proportional to the velocity of the reflector and the cosine of the intercept of the angle. The maximum Doppler-shifted frequency will occur at angles of 0 and 180 degrees. At an

angle of 90 degrees, no Doppler shifted-frequencies will be recorded as cosine 90 degrees is equal to zero. Doppler-shifted frequencies obtained from moving blood in the fetal circulations lie in the audible range (up to 12 KHz) and can therefore be monitored by speakers or stored on magnetic audiotape.

The Doppler signals are processed by performing a Fourier transform on samples of the Doppler-shifted signal to produce a series of Doppler spectra (Evans 1992). Arterial blood flow is pulsatile and examination of the spectral characteristics of the Doppler-shifted signal at a single instant in time gives little information. It is therefore necessary to calculate new spectra at frequent intervals, usually 80 - 200 times a second. This makes it impossible to examine each spectrum in detail, so the data is usually presented in the form of a sonogram to allow interpretation. In this type of data display, the Doppler-shifted frequencies is plotted along the vertical axis and time along the horizontal axis. The sonogram allows volume flow measurements or waveform analyses to be made. The former has not achieved widespread popularity in obstetrics because the technique is cumbersome and there is a large inherent error in the measurements of volume (Burns 1992). Most Dopplerstudies in obstetrics have reported indices based upon the maximum velocity envelope. In the assessment of umbilical artery blood flow, it has the advantage that the maximum frequency envelope is usually recorded even when the vessel is insonated in a non-uniform manner. Three indices have been used to describe the maximum velocity envelope; the systolic / diastolic ratio, pulsatility index and resistance index.

Two types of Doppler ultrasound equipment are used in obstetrics; continuous (CW) and pulsed wave (PW) systems. Continuous wave equipment continually transmit a beam of ultrasound into the tissue and, at the same time, detects the echoes. The disadvantage of CW equipment is that as transmission is continuous, the devices have little or no range resolution and are sensitive to any movements within the ultrasound beam. The advantage with CW equipment is power levels are lower than those obtained from PW systems. Pulsed wave systems have the advantage over CW in that the operator may choose the direction of the Doppler beam and determine the depth from which the signals are gathered. This is done by allowing the receiver to gather signals only from a particular time-window. In obstetrics, PW systems are best used with a coexistent real-time imaging system (duplex system) which overlay a colour-coded map of Doppler-shifted frequencies on a part of real-time imaging. However, PW systems suffer from velocity limitations. The maximum frequency shift that can be detected by a PW system is determined by the Nyquist limit. This states that the maximum Doppler-shifted frequency that can be detected is one-half of the pulse repetition frequency. If the Doppler frequency shift exceeds this limit, the phenomenon of aliasing occurs (Goldberg et al. 1988). Here, the high frequency part of the waveform is cut off and reinserted into the display but in the opposite direction channel. This is not usually a problem in the Doppler assessment of the

umbilical artery waveform.

4.1.3.3. Method of measurement

Umbilical artery Doppler waveforms were obtained using PW Doppler with real-time ultrasound imaging (duplex system). The waveforms were obtained during the same scanning session at which other ultrasound measurements of BPD, HC, AC and FL were made. All Doppler assessments were performed by the author. The patient remained in the similar semi-recumbent position as described previously. The same transducer, a 3.5 MHz linear transducer (L-312) on the same ultrasound machine, Accuson KXP 128 / 1 was used to obtain the umbilical artery Doppler waveforms. The middle portion of the umbilical artery was sampled, away from both placental and fetal cord insertion (Arduini and Rizzo 1990). The umbilical cord was visualised along its longitudinal axis and the the range gate placed over the entire width of the cord in the similar plane, to ensure that all Doppler frequencies were recorded. The Doppler mode on the machine was engaged and the resulting signals were obtained in the absence of gross body movements and during fetal apnoea, as the latter had a marked effect on the waveform (Marsal 1978). The angle of insonation was kept below 55 degrees to avoid artefactual loss of end-diastolic frequencies. The vessel wall filter frequency was always set at the lowest level to ensure that absence of end-diastolic flow was not related to filter frequency.

The image was frozen on screen once satisfactory Doppler waveforms were obtained. The optimal pictures were based on an optimal waveform with a sharp cut-off between the maximum frequency envelope and the blank space above, and the presence of constant venous flow on the negative axis. Three consecutive cardiac cycles were assessed to calculate the pulsatility index. This index was chosen, instead of the resistance index or S / D ratio, as it gave the best indication of the degree of absence of end-diastolic flow. For each cardiac cycle, the maximum systolic and minimum diastolic frequencies were recorded using a cursor on the screen; the cursor was then used to trace the outline of the cardiac cycle from the beginning of the wave to the end. The pulsatility index (Gosling and King 1976) was then calculated automatically by the computer of the ultrasound machine using the formula; Pulsatility index (PI) = (maximum systolic frequency - minimum diastolic frequency) / mean frequency. The resultant three PI's were then averaged.

4.1.2.4 Discussion of methods

The Accuson KXP 128 / 1 machine was chosen to perform all Doppler measurements as it produced good quality flow velocity waveforms for analysis. The duplex facility on the ultrasound machine allowed the portion of the umbilical artery

sampled to be visualised, and ensured that the gate was placed over the entire width of the cord.

Analysis of the flow velocity waveform was undertaken instead of measurement of blood flow velocity in view of the numerous methodological problems and sources of error inherent in the volumetric estimation of the latter (Eik Nes et al. 1984b). The umbilical artery flow velocity waveform can be characterised by three indices; the pulsatility index (Gosling and King 1976), resistance index or the S / D ratio. The latter two have the advantage of being more easily derived from the Doppler spectrum. However, as the diastolic velocity approaches zero, both these indices lose their resolution (Marsal 1989). For this reason, the pulsatility index is superior, and was chosen to characterise the umbilical artery waveform.

4.1.3 Safety of ultrasound measurements

Ultrasound is used routinely in most obstetric departments for the assessment of fetal well-being. The safety of ultrasound has been endorsed by a number of expert bodies, including the safety committees of the European Federation of Societies of Ultrasound in Medicine and Biology (EFSUMB) (1992) and the American Institute of Ultrasound in Medicine (1988). No study has proved that ultrasound, at diagnostic intensities as used to date, has led to any deleterious effect on fetus or mother (EFSUMB 1992).

Any hazard from diagnostic ultrasound exposures comes from the potential of ultrasound to produce change in biological tissues. The biological effects in vitro that would be relevant to its safe use in vivo include inherited changes, increased sister chromatid changes and membrane permeability changes (Dyson 1986). Ultrasound does not appear to damage DNA of somatic cells which might otherwise lead to chromosome aberrations, cell death or mutagenesis. Although MacIntosh and Davey (1970, 1972) reported that ultrasound was capable of causing chromosome aberrations under experimental conditions, others have failed to confirm their findings under more vigorous conditions in-vitro (Buckton and Baker 1972, Abdulla et al. 1972) and in-vivo (Abdulla et al. 1971, Lucas et al. 1972).

Ultrasound can exert a biological effect by heating tissue due to absorption of sound in the medium. The temperature rise depends on the heat generated at a particular site (related to the absorption coefficient) and heat conduction away from the tissues (related to tissue vascularity). Ultrasound of diagnostic intensity did not significantly raise the temperature of mammalian tissue (Nyborg et al. 1983). In vivo exposure of a thermocouple during a second trimester termination did not increase amniotic fluid temperature (Soothill et al. 1987b).

Ultrasound can cause mechanical effects on tissue, including cavitation (the oscillation of gas bubbles due to the pressure of the acoustic wave), microstreaming (an eddying effect) or radiation force (a steady force on cells that moves them). These do not seem to occur to a significant degree in human tissues with diagnostic ultrasound (Kremkau 1983). However, pre- and post- delivery red cell osmotic fragility testing showed a marginal increase in fragility in women exposed to continuous heart monitoring for over seven hours with Doppler ultrasound (Bause et al. 1983).

Ultrasound exposure during pregnancy of 1114 women did not lead to an increase in fetal abnormalities (Hellman et al. 1970). A one-year follow-up of 297 fetuses exposed in-utero to ultrasound and amniocentesis revealed no difference in neurological and

physical examination compared to groups having amniocentesis alone or no testing (Scheidt et al. 1978). A study of 381 exposed fetuses and unexposed matched controls found no difference in a variety of birth or behavioural variables (Stark et al. 1984). In a recent published study of a randomised controlled trial of Norwegian children who were scanned between 1979 and 1981, there were no significant differences in the subsequent vision and hearing at primary school age (Salvesen et al. 1992).

Although no adverse effects arising from the use of diagnostic ultrasound have been identified in practice, epidemiological investigations are not able to give unequivocal reassurance. This is because of difficulty in matching exposed and non-exposed controls and that small changes in the rate of occurence of a common abnormality may be missed. Furthermore, the peak acoustic pressures from today's pulse echo transducers are greater than those used in the early 1980's. Whilst the amount of heating and cavitation induced by current pulse-echo fields in vivo is likely to be biologically insignificant (Starritt and Duck 1992), the same may not be true for pulse Doppler fields (Ter Haar 1992).

It is therefore prudent to minimise the exposure of the fetus and mother to ultrasound energy during an examination. At the same time, a necessary level of acoustic output must be used to achieve the penetration and sensitivity needed to obtain the desired information. Constant checks should be made on the ultrasound equipment in use and the operator should be aware of the power levels at which the ultrasound scans are performed.

All ultrasound measurements were performed on an Accuson 128 / 1 machine using a 3.5 MHz linear transducer (L312) with an Aperture size of 120 mm. The acoustic intensity generated by an ultrasound machine is usually described by the following terms:
1. I-SPTA: Spatial Peak Time Averaged Intensity is the highest intensity within the field, averaged over an entire scan frame period. In considering possible bio-effects, I-SPTA is a measure of thermal effects, such as the heating of tissue.
2. I-SPPA. Spatial Peak Pulse Average Intensity is the highest intensity along the beam path averaged over the duration of the pulse. As the pulse Doppler mode uses short pulses, the average intensity during the pulse may be a fairly large quantity. In speculation about possible bio-effects, I-SPPA is a measure of possible non-thermal mechanisms, such as cavitation.
3. I-m. Maximum intensity is the average intensity during the half-cycle with greatest amplitude during the pulse. Like I-SPPA, I-m is a measure of possible non-thermal mechanisms.

In determining the possible effect of ultrasound on tissue, the intensity encountered at the tissue site must be calculated. Because of attenuation of the beam within the body, the intensity at the tissue in-situ may be 10 to 100 times less than would be measured at the

same location in water. Therefore, to calculate an in-situ intensity that is meaningful for bio-effect considerations, intensity measurements made in water in the laboratory must be adjusted to reflect the effects of attenuation. The amount of attenuation of an ultrasound beam as it travels through the body is determine by the type of tissue along the beam path, frequency of the ultrasound energy and the distance travelled by the beam. The Food and Drugs Administration (FDA) in the USA has produced a formula to calculate a conservative estimate of attenuation due to these factors;

$I_t = I_w .\exp\ (-0.23.a\ .f\ .z)$

where I_t is an estimated in-situ intensity at the tissue site, I_w is the intensity measured in water at a distance of z in cm, a is the attenuation coefficient expressed in dB / cmMHZ, and f is the acoustic frequency in Mhz of the ultrasound beam. The FDA has specified that the value of the attenuation coefficient a, to be used in the above formula is 0.3 dB / cmMhz. The use of such a low attenuation coefficient is far below the typical values for fat, muscle and liver and generally guarantees an over-estimation of the in-situ intensity that the tissues would experience.

The FDA has established guide-lines for ultrasound limits for fetal ultrasound work. The limits set are expressed as in-situ estimated levels, assuming the attenuated model as has already been described. Table 4.2 lists the limits of the 3 measures of intensity as defined by the FDA. The intensities of the L-312 transducer, the only transducer used in all the ultrasound work performed by the author, are also listed both for B-Mode and also for pulsed Doppler. It can be seen that the maximum estimated in-situ intensities for the L-312 transducer on the Accuson machine were less than the acoustic output guide-lines for "Fetal Imaging" as issued by the FDA.

Table 4.2 Maximum estimated in-situ intensities for L312 Transducer and guideline levels issued by FDA (USA).

Ultrasound mode	**I-SPTA** (mW/cm^2)	**I-SPPA** (W/cm^2)	**I-m** (W/cm^2)
B-mode	15	185	277
PW Doppler (medium)	93	39	72
PW Doppler (low)	49	19	36
FDA guidelines	94	190	310

4.2 Neonatal morphometric indices

4.2.1 History of neonatal morphometric measurements

Parizkova (1961) and Forbes (1962) were the first to report quantitative methods of assessing the amount of fat in children. Gruenwald (1963) in his classic paper on " Chronic fetal distress and placental insufficiency" first highlighted the morphometric differences between different infants of the same birthweight. These differences in body composition and nutritional status of infants with the same birthweight were further explored by Scott and Usher (1966).

Miller and Hassanein (1971) were the first to report reference ranges for a morphometric measurement (ponderal index) for the diagnosis of "wasting" or malnutrition in the newborn. Standards for skinfold measurements in the neonate were first reported by Gampel (1965). Subsequent reference standards for children derived from a British population were reported by Tanner (1975) and the first British reference ranges for subscapular and triceps skinfold thickness in neonates were reported by Oakley et al. (1977b).

4.2.2 Method of measurement

Measurements of weight, crown-heel length, mid-arm circumference (MAC) and HC, and subscapular and triceps skinfolds were made within the first two days of life. **Weight** was measured using the standard weighing machine on the postnatal ward (Marsden's, London W9). **Crown-heel length** was measured by placing the infant in a specially-designed shallow box or "infantometer", a measuring scale with a fixed head plate and a movable foot plate (Colley et al. 1991). Two observers (the author and an independent observer) were required to perform this measurement. The baby was placed with the head touching the centre of the fixed end of the box; the tonic neck reflex was used to ensure that the knees were fully extended during the measurement of length (Miller and Hassanein 1971). The foot plate was then brought up by the assistant to touch the sole of the foot in its entire length. The crown-heel length was read off directly from a scale on the right side of the box to the nearest mm. The average of 3 readings was then taken as the final crown-heel length. **Ponderal index** was then calculated using the formula: Ponderal index = weight / length^3 (g / cm^3) x 100.

Mid-arm circumference and **head circumference** were measured using a tape measurer to the nearest mm.The largest occipitofrontal diameter was measured three times and averaged to obtain the HC. To measure the MAC, the mid-point of the left upper

arm was identified by measuring the distance between the acromion and the olecranon with the arm in a horizontal position. The MAC was measured at this mid-point with the arm held in extension. The average of three measurements was used.

Subscapular and triceps skinfold thicknesses were measured using Holtain calipers (Crymych, Wales, United Kingdom). The dial of the caliper was calibrated to 0.2 mm, but measurements could be made accurately to the last 0.1 mm. The calipers were calibrated at regular intervals using a micrometer at the Department of Growth and Development at the Middlesex Hospital, London. During the measurement, the right hand was used to hold the caliper, the left hand maintained a hold on the skinfold throughout the measurement. The jaws of the caliper were applied to the skinfold under the pinch point and the right hand was allowed to relax its grip on the handle so that the jaws could exert their full pressure. A reading was made once the caliper reading was stable (Oakley et al. 1977b). By convention, all measurements were all made on the left arm and the left side of the body. Triceps skinfold was measured half-way down the back of the arm, half-way between the acromiom and the olecranon, on a line passing upwards from the olecranon in the axis of the limb. The arm was held by the side of the body with the elbow extended.The subscapular skinfold was measured immediately below the angle of the left scapula with the fold either in a vertical position or slightly inclined, in the natural cleavage of the skin. The average of three readings was used to calculate each of the skinfold measurements.

4.2.3 Discussion of Methods

Particular care was taken in the measurement of neonatal morphometry as different methodologies have been described for each of these measurements. The same methodology was used in all measurements to achieve consistency of results.

The Holtain caliper was chosen in preference to the Harpenden caliper for the measurement of subscapular and triceps skinfold thickness as it was easier to operate in a confined space such as an incubator. It exerts a pressure of 10 g / mm^2 over the whole range of openings and has a small enough surface area to be used on newborns. Different methods of measuring skinfold thicknesses have been described (Brans et al. 1974, Oakley et al. 1977b). Brans et al. (1974) suggested that skinfold thickness readings should be recorded 60 seconds after application of the calipers in view of the phenomenon of skin compressibility. This method was not used by Oakley et al. (1977b) who recorded measurements once the dial on the caliper achieved a stable reading. As the latter was less likely to cause discomfort and as the reference standards of Oakley et al. (1977b) were to be used in this study, this was the preferred method of choice.

In view of the possible errors involved in the measurement of crown-heel length, an "infantometer" was used to obtain accurate measurements. Most importantly, the methods used to derive the ponderal index and MAC / HC ratio were as defined in the studies which reported reference ranges for these morphometric indices.

In order to reduce potential bias due to prior knowledge of the antenatal ultrasound results of the infants studied in Study 3 (all of which were performed by the author), an independent observer (SCR), who was unaware of the ultrasound data, was recruited to perform the neonatal morphometric measurements. Of all infants who had morphometric measurements made, 86% were assessed by the independent observer (SCR) and the remainder (14%) by the author. Both underwent training in the morphometric assessment of the neonate by attending Growth Clinics at the Department of Growth and Development, Middlesex Hospital, London.

4.3 Measures of perinatal morbidity

Perinatal outcome was assessed in all cases using the following criteria:

i) **Emergency Caesarean section for fetal distress in labour.** Fetal distress was diagnosed by the obstetrician on-duty in the labour ward to be an ominous cardiotocographic trace warranting immediate delivery.

ii) **Umbilical arterial and venous pH and base excess (BE).** Umbilical arterial and venous blood were obtained at delivery by the author or attending mid-wife using the triple-clamping procedure. Following the delivery of the infant and prior to the expulsion of the placenta, three clamps were placed on the umbilical cord. One was placed close to the vaginal orifice and the other two close together near the infant end of the cord. The cord was cut between the latter two clamps and the infant handed over to the paediatrician or assisting midwife. Prior to the expulsion of the placenta, cord blood was obtained separately from the umbilical artery and vein between the two remaining clamps for pH and BE analysis. One ml of umbilical arterial blood was collected into a heparinised Steriseal syringe for analysis of arterial pH and BE. One ml of umbilical venous blood was also collected using a two way tap. Arterial and venous samples were analysed immediately using a blood gas analyser (ABL 300 Acid-base Laboratory, Radiometer Copenhagen).

iii) **Acidaemia at delivery.** Acidaemia was defined as an umbilical venous or arterial pH < 10th centile (< 7.23, <7.14 respectively according to the standards of Eskes et al. 1983).

iv) **Apgar at 5 minutes ≤ 7.** The Apgar score at 5 minutes was recorded by the paediatrician or midwife in the delivery room. Scores of 0, 1 or 2 were assigned to five vital signs; heart rate, respiration rate, muscle tone, reflex irritability and colour of the newborn to give a total score ranging from 0 to 10.

v) **Admissions to NICU related to IUGR.** The number of neonates admitted to NICU was recorded. Admission to NICU was only included as a measure of adverse perinatal outcome if the reasons for admission were complications associated with IUGR. These included hypoglycaemia, necrotising enterocolitis and neurological abnormalities such as hypotonia, irritability or neonatal convulsions. **Hypoglycaemia** was defined as a heel prick glucose level (BM stix testing) ≤ 2 mmol / l on Day 1, refractory to oral feeding and requiring intravenous glucose treatment. **Necrotising enterocolitis** was defined as the occurrence of abdominal distension, bilious vomiting and treatment with parenteral nutrition and antibiotics, together with radiological features of pneumatosis coli or perforation (Malcolm et al. 1991). All neonates were examined in the first three days of

life by a dedicated paediatric neonatologist (Dr. Simon Roth) who performed a neurological examination (Amiel-Tison et al. 1982) to ascertain any neurological deficits with particular regard to tone and evidence of irritability. **Neurological deficits** were defined as seizures, hypotonia or irritablity requiring admission to NICU for further observations.

4.4 Biochemical indices of IUGR at delivery

Umbilical venous samples were also obtained at delivery to evaluate the biochemical profile of the neonates. At the same time as obtaining paired umbilical arterial and venous samples for pH and BE, the author also collected umbilical venous blood for analysis of other biochemical indices suggestive of IUGR. Umbilical venous blood was obtained by the triple clamping of the cord after the birth of the baby and before expulsion of the placenta (see Section 4.3). A separate 10 ml syringe was fitted onto the two way tap and **10 mls** of umbilical venous blood collected. This was divided into the following amounts:

4 mls into a heparinised tube (Vacutainer PST lithium heparin, Rutherford, N.J., USA) for insulin assay.

3 mls into a plain Vacutainer tube for IGF-1 assay.

2 mls into a heparinised tube (Li- Heparin LH/1.3) for triglycerides assay

1 ml into a sodium fluoride bottle (Fluoride / Heparin Alpha Laboratories) for glucose levels.

With the exception of the sample for glucose levels, all remaining samples were centrifuged within 5 minutes of collection by the author using a IEC Centra 4-B centrifuge (International Electric Company, USA) at 3500 rpm for 10 minutes. Samples of plasma (for insulin and triglycerides) and serum (for IGF-1) were then pipetted off and stored immediately at -30° Centigrade until further analysis.

Glucose samples were sent on the same day of collection to the Department of Biochemistry, Middlesex Hospital, London for further analysis. Glucose levels were measured using a glucose oxidase analyser (Yellow Springs Intrument, Ohio, USA).

Plasma triglyceride concentrations were measured using fully enzymatic colorimetric procedures (triglycerides N and NEFA C, Wako Pure Chemicals, FRG) on a discrete automated analyser (Cobas MIRA, Roche Diagnostics, Basel, Switzerland). The intra- and inter-assay coefficients of variation for triglycerides were less than 2%. These plasma samples were analysed by Dr. David Crook at the Cavendish Laboratories, Wynn Institute, London NW2.

Plasma insulin was measured using radioimmunoassay as described by Albano et al. (1972). The method of measuring plasma insulin was based on the use of activated charcoal for the separation of free and unbound fractions. The intra- and inter-assay coefficients of variation were less than 6%. Plasma samples were analysed by Dr. Anthony Proudler at the Cavendish Laboratories, Wynn Institute, London NW2.

Serum IGF-1 levels were acid / ethanol extracted and IGF-1 concentrations

determined by radioimmunoassay according to the method described by Taylor et al. (1988). IGF-1 antiserum was raised in rabbits by immunization with 125 micrograms of 20% pure IGF-1 conjugated to ovalbumin. The sensitivity of the assay was 0.07 U / ml. Intra-assay coefficients of variation were 11.3% at 0.23 U / ml and 6.5% at 1.23 U / ml. Inter-assay coefficients of variation were 10.5% at 0.38 U / ml and 12.1% at 0.99 U / ml. Serum samples were analysed by Dr. David Morrell and Dr. Jennifer Jones at the Department of Growth and Development, Institute of Child Health, London WC1.

4.5. Statistical Methods

4.5.1 Derivation of reference ranges for AC and EFW

In Study 1, reference standards for AC and EFW were constructed to provide the basis for evaluating fetal growth in Study 3. An analysis was performed to determine a suitable class of mathematical models which could be used to describe the growth of individual fetuses. As the variance of AC and EFW is known to increase with gestational age (Deter et al. 1982), a logarithmic transformation was made to the data. Four mathematical models were then investigated as candidates for approximating $\log_{10}$(AC) and $\log_{10}$(EFW) over the range of gestational ages available;

1. Linear model; $\log_{10}(AC) = a + b\,(GA)$
2. Quadratic model; $\log_{10}(AC) = a + b\,(GA) + c\,(GA)^2$
3. Gompertz model; $\log_{10}(AC) = a - \exp(b + c\,(GA)$
4. Rossavik model; $\log_{10}(AC) = a + (b + c\,GA)\log(GA)$

where a, b and c represent constants whose values need to be determined for each individual fetus using least squares fitting. Similar models were also used to describe $\log_{10}$(EFW). The linear and quadratic models were chosen on the grounds of simplicity. The Gompertz model has been used in other biological contexts, for example the modelling of tumour growth (Day 1966). The Rossavik model has previously been used to evaluate fetal growth (Rossavik and Deter 1984, Deter and Harrist 1992).

The different models were evaluated by 3 methods. The residual fitting errors were first visually inspected by plotting them against gestational age. The standard deviations of the residual fitting errors with each model were then compared. The final test involved omitting the final AC and EFW measurements from the fitting process for each fetus. To assess how well the resulting calibrated formula predicted the final measurement, the prediction errors (expressed as a 95% confidence interval) were compared.

Having determined the most appropriate model for describing $\log_{10}$(AC) and $\log_{10}$(EFW), the individual values of a, b and c were used to produce individual growth curves for $\log_{10}$(AC) and $\log_{10}$(EFW). These were then used to interpolate $\log_{10}$(AC) and $\log_{10}$(EFW) values for each fetus at a range of exact gestational ages between 20 and 40 weeks. Values were never extrapolated beyond the gestation of a fetus's final scan. The mean and standard deviation of the resulting interpolated values were then used to derive centile ranges for AC and EFW, taking account of the sample size available using the Student-t correction.

Reference standards for growth velocity of EFW and AC were also

constructed by using measurements of AC and EFW collected at two weekly intervals. Fetal growth velocity was then expressed as the increment in AC (or EFW) per week. The original data within ± 2 days of specific two weekly gestations were used for this calculation. Smoothed data were not used for the derivation of these reference ranges as the resultant SD would otherwise be inappropriately small (Laird and Ware 1982). The mean (and SD) growth velocities for AC and EFW at weekly intervals were then calculated.

4.5.2 Reproducibility of ultrasound and morphometric measurements

In Study 2, ultrasound measurements and measurements of neonatal morphometry were subjected to tests of reproducibility. Intra-observer and inter-observer variability were assessed using one-way analysis of variance (Healy 1989) and the limits of agreement method (Bland and Altman 1986) respectively.

For each observer, the variability was estimated from the residual variance (Healy 1989). Assuming an observer makes p replicate measurements on each of m fetuses, the data form a one-way classification. The statistical model $x_{ik} = \alpha i + \beta + \varepsilon_{ik}$ was used where x_{ik} is the k-th reading on the i-th fetus, αi is the true value for this fetus, β is the clinician's fixed measurement bias and εik is a normally distributed error term with variance σ^2. The standard analysis of variance table is shown in Table 4.3.

Table 4.3 Table of One-way Analysis of Variance

Source of variation	Degrees of freedom	Expected mean square
Between fetuses	m - 1	
Within fetus	m (p -1)	σ^2
Total	mp - 1	

The intra-observer variance was thus estimated from the within-fetus mean square.

Inter-observer variability was assessed, using the mean of each observers three measurements, by the limits of agreement method (Bland and Altman 1986). Plots of the

differences between the observers against their means were constructed for each measured variable. This was done for both absolute and percentage differences; the latter was calculated as the difference between observers expressed as a percentage of the mean. Having determined that there was no significant correlation between the differences and the means, the mean difference between observers, expressed both as absolute and percentage values, were calculated to determine if there was any systematic bias between the observers. The standard deviations of the mean absolute and percentage inter-observer differences were then used to calculate the respective 95% prediction intervals (mean ± 2SD).

4.5.3 Quantification of serial values of AC and EFW

In Study 3a, three methods of quantifying serial measurements were evaluated. Each of these were measures of growth which involved a different statistical assessment of the fetal growth trajectory.

The first method of quantifying serial measurements was the use of change in standard deviation scores (Δ SDS) of AC and EFW over gestation (Δ AC.SDS and Δ EFW.SDS). For each fetus, the standard deviation score (SDS) for AC at any gestational age was calculated using the formula: AC.SDS = (Measured AC - Mean AC at same gestation) / SD of AC at same gestation). Mean (and SD) values of AC at gestational ages between 26 and 40 weeks were obtained from the reference standards of Study 1. Standard deviation scores for AC at the first ($AC.SDS_1$) and last ($AC.SDS_2$) scans were then used to calculate the change in SDS (Δ AC.SDS = $AC.SDS_1$ - $AC.SDS_2$). Similar calculations were performed to derive Δ EFW.SDS.

The second method of quantifying serial measurements involved the fitting of a log quadratic model [$\log_{10}(AC) = a + b\,(GA) + c\,(GA)^2$] to all the AC data for each fetus. Values of the quadratic coefficients *a*, *b* and *c* were then calculated for each individual fetus. These coefficients were descriptive of different parts of the quadratic curve. The *a* coefficient represented the intersect on the y-axis whilst the *c* coefficient described the departure of the curve from linearity towards term. The quadratic coefficient *b* was a measure of growth velocity. Comparisons of the value of *b* -coef.AC (*b* -coefficient for AC) with the mean (and standard deviation) values of *b* -coef.AC of normal fetuses allowed the degree of deviant growth in a fetus to be quantified. Similar *b* -coefficients were also calculated for serial EFW data in each fetus (*b* -coef.EFW), and compared with reference values from the normal group.

The third method involved the calculation of fetal growth velocity (FGV). This

was obtained by dividing the increment in AC by the duration between scans, standardized here to be two weeks. A SDS of FGV (FGV.SDS) was calculated to enable comparisons to be made independent of gestational age. This was calculated using the formula; FGV.AC.SDS = [(Calculated FGV.AC) - (Mean FGV.AC over same period)] / (SD of FGV.AC over same period). The mean (and SD) values for AC growth velocity were obtained from reference standards in Study 1. The smallest value of FGV for each fetus was used to define abnormal growth. Likewise, change in SDS of the growth velocity of EFW (FGV.EFW.SDS) was also calculated for each fetus.

4.5.4 Comparison of ultrasound parameters

Linear regression analysis was performed to ascertain the relationship between ultrasound variables and neonatal morphometric indices. Significance was assumed if $p < 0.05$. Receiver operating characteristic curves (Richardson et al. 1985) were used to compare different ultrasound parameters in Studies 3a and 3b. Plots of sensitivity vs. (1 - specificity) were generated for each ultrasound and outcome measure. Overall test performance of each ultrasound measure was assessed by calculating the area under the ROC curve (Hanley and MacNeil 1982) and standard error (Hanley and MacNeil 1983). Calculation of the standard errors allowed different areas under the ROC curves to be compared using the Wilcoxon test (Hanley and MacNeil 1983, Hanley 1989), with differences being statistically significant if $p < 0.05$.

For each ultrasound measure, the sensitivity, specificity, OR (together with 95% CI) and Cohen's Kappa index were also calculated using the optimal cut-off as determined from the ROC curve, as well as using a standard cut-off of 2 SD. The optimal cut-off point on the ROC curve was defined as the point on the ROC curve closest to a sensitivity of 100% and (1 - specificity) of 0%. The OR and 95% CI were calculated according to the formula described by Kahn and Sempos (1989). Cohen's Kappa index was calculated using the formula as described by Grant and Mohide (1982). The formulae for these calculations are shown in Table 4.4.

Table 4.4 Statistical evaluation of a diagnostic test

	Actual condition	
Test result	Present	Absent
Positive	True positive (TP)	False positive (FP)
Negative	False negative (FN)	True negative (TN)

Sensitivity = TP / (TP + FN)

Specificity = TN / (FP + TN)

Odds ratio = (TP.TN) / (FN.FP)

Cohen's Kappa index = $(P_o - P_e) / (1 - P_e)$
where $P_o = (TP + TN) / (TP + TN + FP + FN)$
and $P_e = [(TP + FP)(TP + FN) + (FN + TN)(FP + TN)] / (TP + TN + FP + FN)^2$

4.5.5 Comparison between variables

In Studies 1 and 3c, continuous variables between groups were compared using the Student - t test for normally distributed data or Mann-Whitney U test for non-parametric data. Categorical variables were compared using the χ^2 test or Fisher's exact test, depending on the number size in the contingency tables. A p value < 0.05 was considered statistically significant.

4.5.6 Discussion of methods

In the construction of reference standards for AC and EFW, a logarithmic transformation was made to the data to reduce the variability of the variance, as the latter is known to increase with gestational age (Deter et al. 1982a). A ?-Shapiro test was used to confirm that the transformed data was normally distributed at all gestational ages.

One-way analysis of variance was used to assess the intra-observer variability of measurements in preference to calculation of correlation coefficient or coefficient of variation in view of the limitations of the latter two methods in assessing the repeatability of measurements (Healy 1989, Bland and Altman 1992). The limits of agreement method, first reported by Bland and Altman (1986), was used to assess inter-observer variability as it allowed differences between observers to be expressed as a 95% prediction interval (Brennan and Silman 1992).

Although the three different methods of quantifying serial ultrasound measurements were by no means exhaustive, they represented a variety of statistical measures which could be used to quantify serial non-linear numerical data. Receiver operating characteristic curves were used to compare different ultrasound parameters for three reasons. First, the overall test performance of each ultrasound measure could be assessed by a single statistic, the area under the ROC curve (Hanley and MacNeil 1982). A non-discriminatory ultrasound test would detect the same proportion of correct cases whether in infants with or without IUGR and would therefore be useless (line of equality, area under the curve = 0.5). The ultrasound test with the greatest area under the curve (area under the curve closest to 1.0) would be the superior test. Second, different ultrasound tests could be compared by calculating the standard errors of the respective areas under the ROC curves (Hanley and MacNeil 1983, Hanley 1989). Third, ROC curves enabled different ultrasound tests to be compared using their respective optimal cut-offs as determined from the ROC curve. This method has previously been used to compare different ultrasound tests in the prediction of SGA and obviates the use of an arbitrary antenatal cut-off (Miller and Gabert 1992).

Contingency tables were constructed using the optimal cut-off for each ultrasound parameter. The best ultrasound measure of fetal growth should have the highest sensitivity, specificity, predictive odds and test efficiency. Three methods of expressing the predictive odds are the positive predictive value (PPV), OR and the odds of being affected given a positive result (OAPR). The OR was used in preference to the others as it is prevalence-independent (Kahn and Sempos 1989). In contrast, both the PPV and OAPR are prevalence-dependent (Villar and Belizan 1986, Wald and Cuckle 1989). The OR is the ratio of the odds that a fetus diagnosed to be growth retarded by ultrasound will actually have neonatal morphometric evidence of IUGR, to that of a fetus with ultrasonic evidence of normal growth having normal neonatal morphometric indices. The 95% confidence intervals for each OR were also calculated according to the method described by Kahn and Sempos (1989). A discriminatory test would have a 95% confidence interval greater than zero. The 95% confidence intervals could also be used to determine if particular ultrasound criterion was significantly better than others.

The efficiency of each ultrasound measure was also assessed by calculating the Cohen's kappa index (Grant and Mohide 1982). The Kappa index is a ratio of the observed accuracy beyond chance to the maximum achievable accuracy beyond chance: [Kappa = $(P_o - P_c)/(1-P_c)$] where P_o is the observed proportion of the total number of patients who have correct test results, and P_c is the proportion of results expected to be correct on the basis of chance alone. The Kappa index has a maximum value of 1. Large Kappa values reflect optimal agreement between ultrasound test and true IUGR and values between 0.2 and 0.8 indicate fairly good agreement. Values between 0 and 0.2 reflect agreement only by chance whilst values below zero are associated with disagreement (Grant and Mohide 1982).

4.6 Ethical Permission

Permission for all investigations was obtained from the Ethics Committee of the University College and Middlesex Schools of Medicine, London. Informed consent was obtained from each subject.

CHAPTER 5

REFERENCE RANGES FOR ABDOMINAL CIRCUMFERENCE AND ESTIMATED FETAL WEIGHT

5.1 Introduction

Numerous workers have reported reference charts relating AC and EFW to gestational age (GA). There are advantages in the use of longitudinal data in preference to cross-sectional data for the construction of growth standards (Evans et al. 1990, Spark and Cetin 1991). However with the exception of one small study (Deter et al. 1982a), the data have been analysed as if collected cross-sectionally. There was therefore a need for reference standards for AC and EFW to be produced based on the appropriate statistical analysis of longitudinal ultrasound data in a sizeable group of fetuses.

5.2 Subjects

Seventy Caucasian women were recruited from the antenatal clinics of University College Hospital and Queen Charlotte's and Chelsea Hospital, London after a routine 18-20 week anomaly scan. The criteria for recruitment were that the women had regular menstrual cycles and that ultrasonic measurements of BPD and FL at the level II scan were less than 7 days discrepant from menstrual dates. Gestational age was calculated according to menstrual dates. Forty women recruited from University College Hospital were scanned by the author at approximately 2 week intervals from 26 weeks gestation until delivery. The remaining 30 women from Queen Charlotte's and Chelsea Hospital were scanned by another obstetrician (SCR) at similar intervals. Three fetuses were subsequently excluded from the final analysis because of delivery less than 37 weeks gestation, leaving a study group of 67 fetuses. To ensure that the group of fetuses did not represent a supra-normal sample, women who had risk factors associated with IUGR (like smoking, pregnancy-induced hypertension) were not excluded.

The demographic and delivery details of the study group are shown in Table 5.1. No significant differences in demographic or delivery details were found between the two groups of fetuses. The data from these two populations were therefore combined.

Table 5.1 Demographic and delivery details of fetuses for construction of references ranges

	University College Hospital	Queen Charlotte's Hospital
No. of patients	37	30
GA at delivery (days)	281 (7)	278 (10)
Birthweight (g)	3497 (346)	3410 (511)
Maternal height (cm)	162 (5.9)	160 (6.1)
Maternal weight (kg)	61.6 (7.3)	60.0 (7.6)
No. of smokers	3	4
Social class I and II	15	16
III and IV	17	15
V	0	1

Figures are mean (SD).

5.3 Results

The results of 67 fetuses were used in the final analysis. The mean EFW at the final scan was 3523 (SD 470) g. The median interval between the last scan and delivery was 7 [range 0-13] days. The mean gestational age at delivery was 279.9 (SD 8.6) days and the mean birthweight was 3474 (SD 427) g. The mean difference between birthweight and the final EFW was -48.5 g [95% CI (-105.7, 8.7 g)].

Preliminary analysis of the raw data from the 67 fetuses showed that the variances of both AC and EFW increased with gestational age. Regression analysis of EFW data available within 3 days of 28 and 36 weeks gestation (n = 47) showed a significant correlation between the increment in EFW between 28 and 36 weeks and the mean value of EFW at 28 and 36 weeks, suggesting that small fetuses have a smaller increment in EFW (Figure 5.1). Thus it was inappropriate to analyse the entire data set as though it were independent cross-sectional data.

Inspection of the individual plots of AC and EFW against gestational age suggested that growth was approximately linear until 36 weeks gestation. Visual inspection of the plots of the residual errors $\log_{10}$(AC) against gestational age showed that the errors for the linear model varied systematically with gestation (Figure 5.2), $\log_{10}$(AC) being overestimated at either end of the gestational range. The 3 other models showed no obvious systematic bias. A similar pattern was evident for $\log_{10}$(EFW). The standard deviation of the residual errors for the 4 models are shown in Table 5.2. The values for the linear model for both $\log_{10}$(AC) and $\log_{10}$(EFW) were significantly greater than for the other 3 models, confirming the poorness of fit of the linear model.

Table 5.2 Standard deviation of residual errors after least squares fitting of growth models.

	Standard deviation of residual errors	
Formula	**Log(AC)**	**Log(EFW)**
Linear	0.0251	0.0634
Quadratic	0.0087	0.0179
Gompertz	0.0085	0.0181
Rossavik	0.0086	0.0183

Table 5.3 show the 95% confidence intervals of the prediction errors for the final $\log_{10}$(AC) and $\log_{10}$(EFW) in the 67 fetuses, having omitted this data point from the fitting process. In contrast to the quadratic, Gompertz and Rossavik models, the linear model overestimated $\log_{10}$(AC). The linear, Gompertz and Rossavik formulae also significantly overestimated the final $\log_{10}$(EFW) although the error was substantially greater with the linear model. The quadratic model was therefore used to describe changes in both $\log_{10}$(AC) and $\log_{10}$(EFW) with gestational age.

Table 5.3 95% confidence intervals of the prediction errors for the final log10(AC) and log10(EFW) data.

Formula	**Log(AC)**	**Log(EFW)**
Linear	(-0.0501, -0.0424)	(-0.1351,-0.1186)
Quadratic	(-0.0007, 0.0085)	(-0.0086, 0.0105)
Gompertz	(-0.0060, 0.0015)	(-0.0227,-0.0059)
Rossavik	(-0.0061, 0.0022)	(-0.0230,-0.0051)

The mean (SD) of the constants *a, b* and *c* for the quadratic formula for $\log_{10}$(AC) were 0.3356 (0.1808), 0.0544 (0.0121) and 0.0006 (0.0002) respectively. Corresponding values for $\log_{10}$(EFW) were 0.2508 (0.3333), 0.1458 (0.0231) and -0.0016 (0.0004) respectively. For each gestational age between 20 and 40 weeks, the 67 interpolated values for $\log_{10}$(AC) and $\log_{10}$(EFW) were approximately normally distributed. Figures 5.3 and 5.4 show the mean (±2 SD) of AC and EFW plotted against gestational age. The widening of the normal range with gestational age reflected the increasing variance of both AC and EFW. This widening also increased beyond 37 weeks due to a reduction in the available sample size. The normal range, expressed as percentiles, are shown in Tables 5.4 and 5.5 respectively.

No significant differences in the coefficients *a, b* or *c* were found between male and female fetuses, or between primiparity and multiparity. Comparison of unsmoothed AC and EFW values at 28, 32 and 36 weeks gestation in male vs. female fetuses and primiparity vs. multiparity revealed no significant differences.

Forty three fetuses with AC and EFW values to within ± 2 days of exact gestational ages at fortnightly intervals between 26 and 40 weeks were used to construct

reference standards for growth velocity. The mean (± 2 SD) values for fetal growth velocity (FGV.AC and FGV.EFW) from 26 weeks to 40 weeks gestation are shown in Table 5.6. It can be seen that values for FGV decreased towards term. For example, mean (SD) values of FGV.AC and FGV.EFW decreased from 11.3 (SD 6.2) mm / week and 219 (SD 86) g / week respectively at 30 weeks to 7.1 (SD 3.9) mm / week and 194 (SD 99) g / week at 38 weeks gestation. The reference ranges (10th, 50th and 90th centile values) for the growth velocities of AC and EFW are shown in Figures 5.5 and 5.6 respectively.

Table 5.4 Centiles for abdominal circumference (mm)

GA	1st	5th	10th	50th	90th	95th	99th
20	127	136	140	153	167	171	180
21	141	147	151	163	177	181	190
22	150	158	162	174	188	192	200
23	164	170	173	186	199	203	211
24	175	182	185	197	211	215	223
25	187	193	197	209	222	226	234
26	198	205	208	221	234	238	246
27	209	216	220	233	246	251	259
28	221	228	231	245	258	263	271
29	232	239	243	256	270	275	283
30	243	250	254	268	282	286	295
31	254	261	265	279	293	298	306
32	265	272	276	290	305	309	317
33	275	282	286	303	315	320	328
34	284	292	296	310	326	330	339
35	292	300	304	320	336	341	350
36	299	308	312	329	346	351	361
37	306	315	320	337	356	361	372
38	315	324	329	345	363	368	378
39	319	329	334	352	371	377	388
40	324	334	339	355	373	379	390

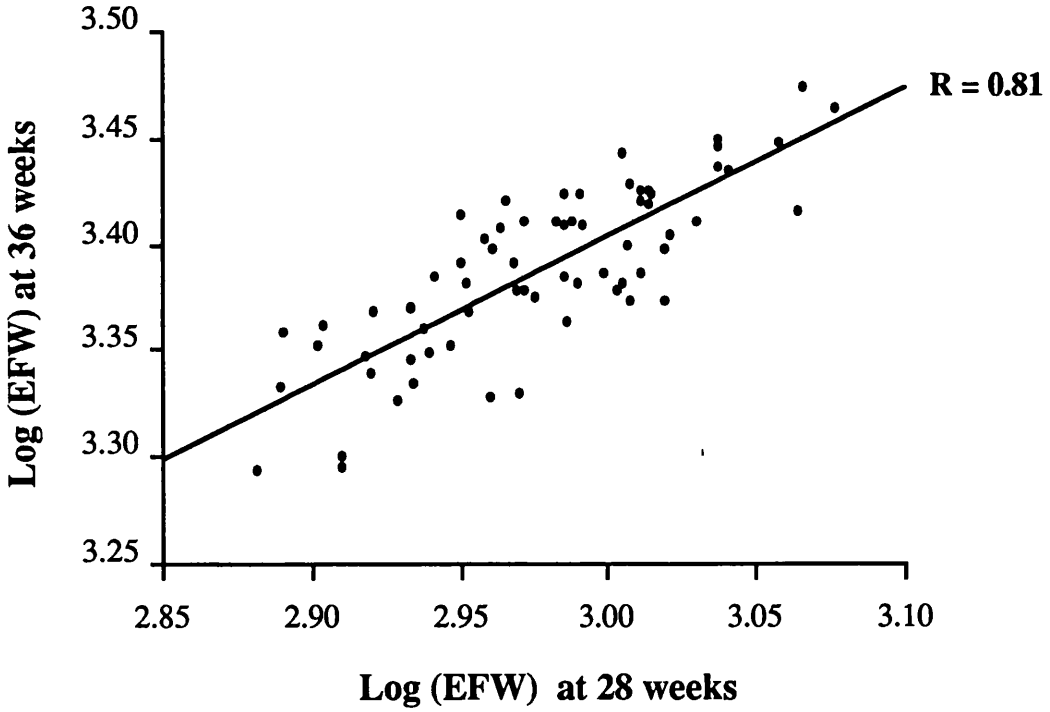
Table 5.5 Centiles for estimated fetal weight (g)

GA	1st	5th	10th	50th	90th	95th	99th
20	250	275	289	344	410	431	474
21	307	336	353	415	489	512	561
22	374	407	426	497	580	607	661
23	449	488	509	591	685	716	777
24	534	579	603	697	805	839	909
25	630	681	709	816	940	979	1058
26	736	794	827	949	1090	1134	1224
27	853	920	957	1096	1255	1305	1407
28	983	1058	1099	1256	1435	1491	1605
29	1123	1208	1254	1429	1628	1691	1818
30	1275	1369	1421	1614	1834	1903	2043
31	1438	1541	1598	1810	2051	2127	2279
32	1609	1722	1784	2016	2278	2360	2525
33	1785	1908	1976	2228	2512	2601	2781
34	1962	2096	2170	2445	2754	2851	3046
35	2134	2281	2362	2663	3003	3109	3323
36	2202	2457	2546	2880	3259	3378	3617
37	2745	2650	2747	3108	3517	3645	3904
38	2697	2873	2968	3320	3714	3837	4086
39	2843	3033	3135	3509	3928	4061	4332
40	2924	3131	3240	3633	4074	4216	4515

Table 5.6 Mean (and SD) values for fetal growth velocity of AC and EFW.

Gestation	FGV.AC (mm / wk)		FGV.EFW (g / wk)	
	Mean	SD	Mean	SD
26	9.8	5.1	169	54
28	10.5	5.1	181	61
30	11.3	6.2	219	86
32	9.7	4.9	213	81
34	8.1	4.4	207	89
36	8.5	4.7	198	85
38	7.1	3.9	194	99

Figure 5.1 Relationship between log (EFW) at 28 weeks and log (EFW) at 36 weeks.




Residual
0.1
0.05
0
-0.05
-0.1
15
17
19
21
23
25
27
29
31
33
35
37
39
41
43
45
Gestational Age (wks)

Figure 5.3 Reference Range for Abdominal Circumference (Mean ± 2 SD)

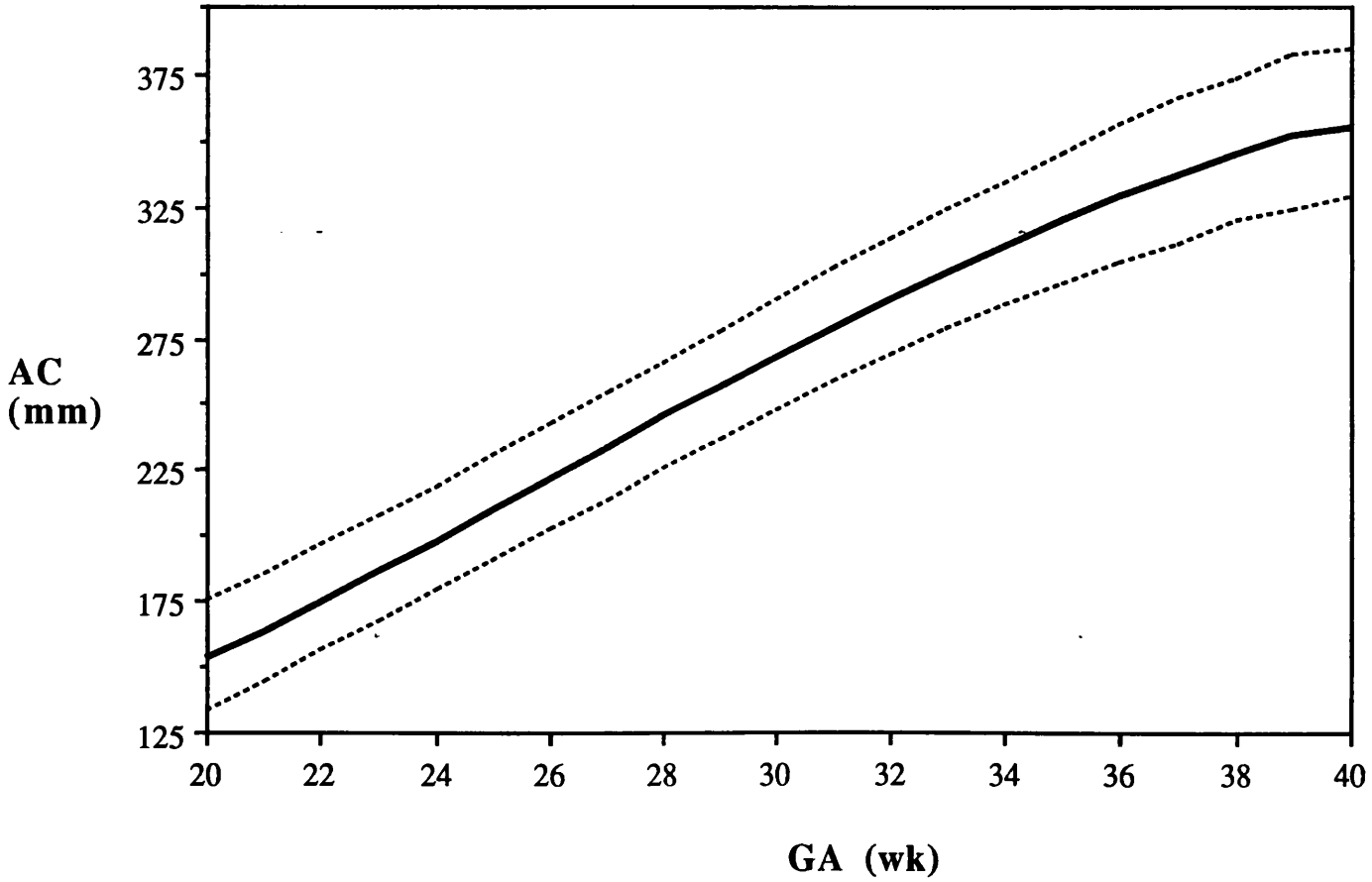


Figure 5.4 Reference Range for Estimated Fetal Weight (Mean ± 2 SD)

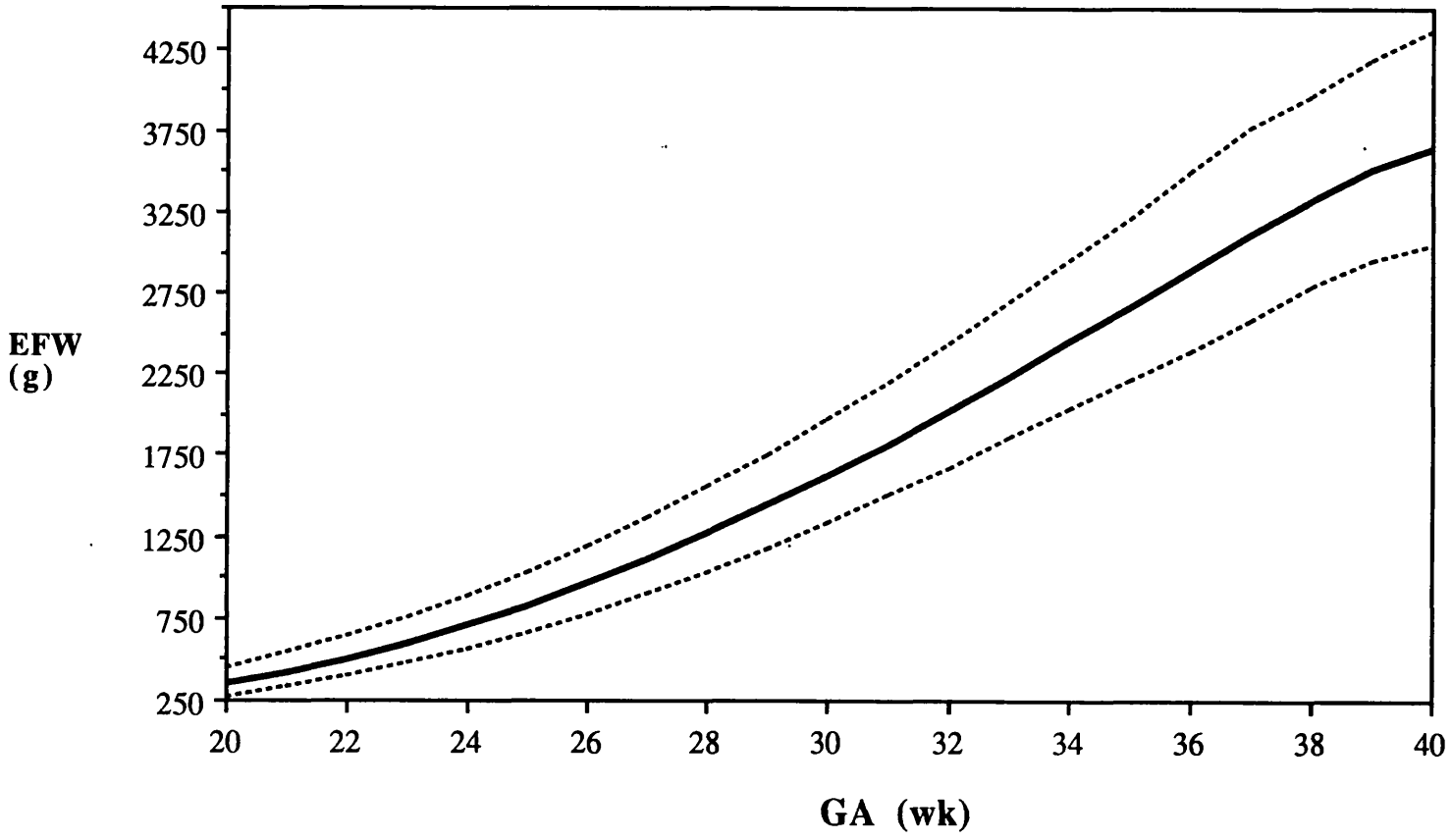


Figure 5.5 Reference range for growth velocities of AC (mean ± upper and lower 10th centiles)

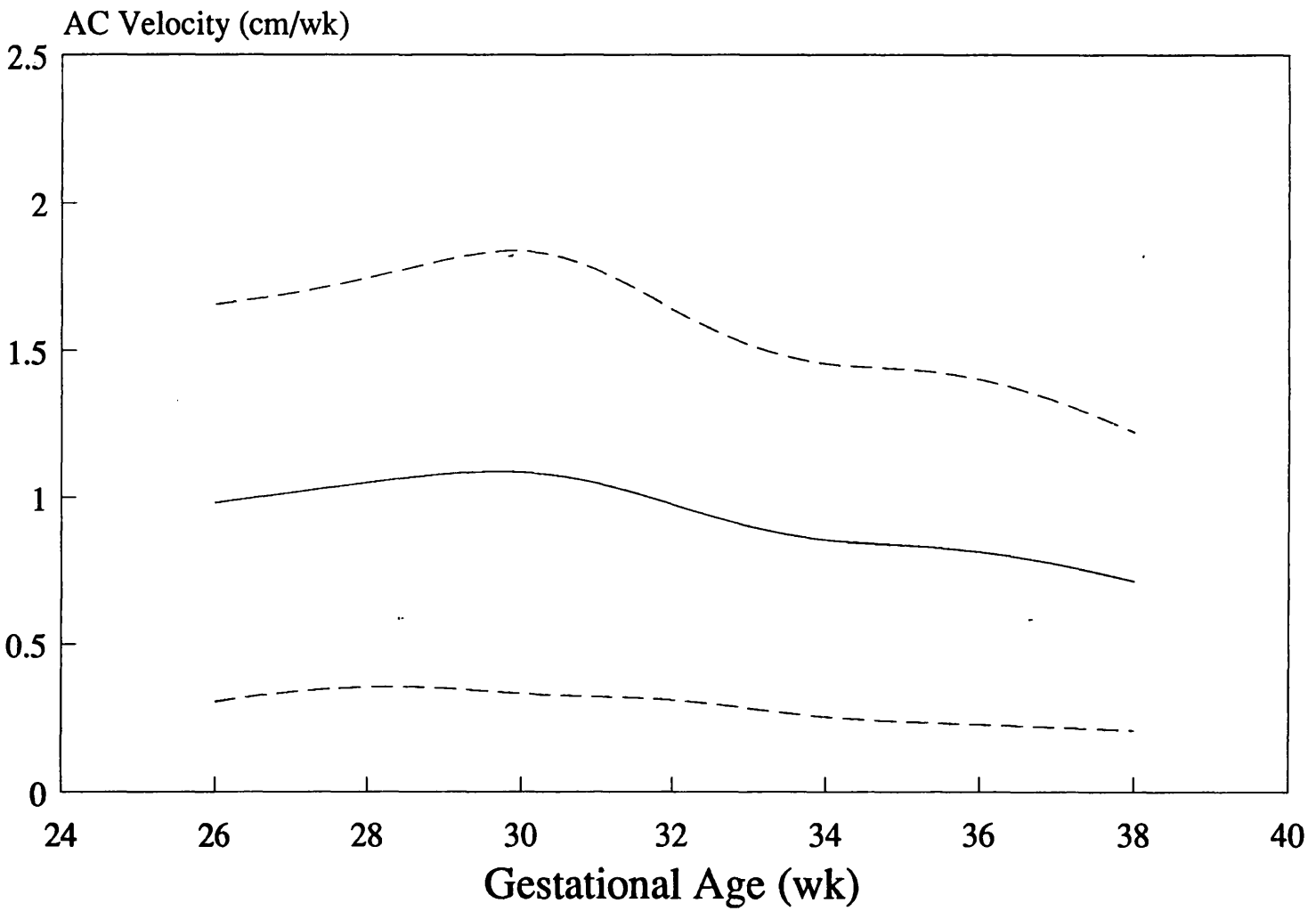
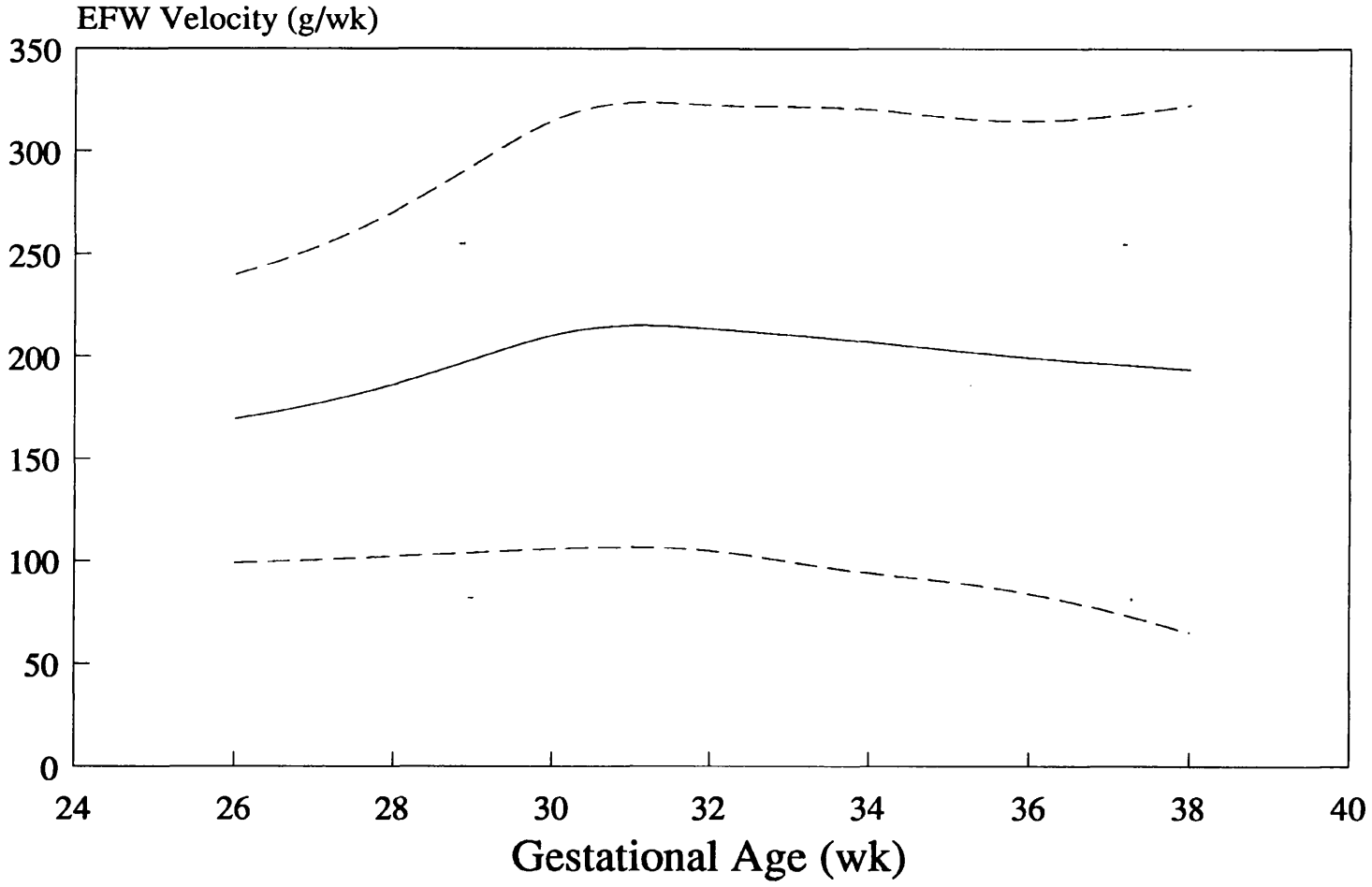


Figure 5.6 Reference range for growth velocities of EFW (mean ± upper and lower 10th centiles)



5.4 Discussion

The BMUS bulletin (Evans et al. 1990) highlighted the lack of appropriate reference standards for AC. To address that, reference ranges for AC and EFW were derived from a Caucasian population using longitudinal ultrasound data. A log quadratic formula best described serial changes in both AC and EFW and the quadratic coefficients were used to establish normal ranges for the population.

The subjects for the study were Caucasian women chosen at random from the two populations. No differences in demography or delivery details were found between the two groups and therefore the data were combined. The study was restricted to Caucasian women to exclude any effect of race on fetal size and growth (Meire and Farrant 1981, Tamura and Sabbagha 1980). Women were included only if menstrual and ultrasonic dates agreed to within 7 days. Although Geirsson (1991) suggested that ultrasonic dates should be used in preference to menstrual dates, the size of the difference was so small in the study population that we did not feel this was justified. Following recruitment, the only fetuses excluded from the final analysis were those in which data were not available beyond 36 weeks gestation (n = 3). Specifically smokers, women with pregnancy-induced hypertension or fetuses that were subsequently small-for-gestational age at delivery were not excluded. Thus, the data derived from the study population was representative of fetal growth for the Caucasian population studied.

Several previous studies of serial AC and EFW have evaluated models for describing fetal growth (Deter et al. 1982a, Jeanty et al. 1984a, Jeanty et al. 1984b, Persson and Weldner 1986, Persson 1989, Larsen et al. 1990). Regression analyses have been used to compare the goodness of fit of various models; the use of the R^2 value and the coefficient of variation, however, conveys little information concerning the size of the residual error. For this reason, different models were compared using the standard deviation of the residual error in this study. In addition, the prediction error for the final AC and EFW was calculated having omitted these data from the fitting process; this was particularly relevant in view of the conflicting data regarding fetal growth in late pregnancy. The results suggested that changes in AC and EFW over gestation were not linear and were best described using a quadratic formula. Jeanty et al. (1984a) and Larsen et al. (1990) found that their AC data were best described by a cubic and fourth order polynomial model respectively while Deter et al. (1982a) reported similar R^2 values using a linear or a cubic model. In agreement with the present results, Deter et al. (1982a) and Larsen et al. (1990) concluded that a quadratic formula best described changes in EFW although Persson and Weldner (1986) and Jeanty et al (1984b) reported a higher correlation coefficient with a cubic formula. The results of this study confirmed that fetal

growth, whether assessed by serial changes in AC or EFW, decreased near term.

A comparison of the mean (and -2 SD) values of AC at 28, 34 and 40 weeks gestation with those from previous serial studies are shown in Table 5.7. Mean values of AC were comparable with the exception of those reported by Jeanty et al. (1984a). In that study mean AC was significantly lower, even though the authors excluded measurements from fetuses with a birth weight below the third percentile. The size of the SD's varied between different studies; this probably reflected variations in sample size and the statistical methods used. Larsen et al. (1990) analysed their longitudinal data by regression analysis which assumes that all data are independent. Use of several measurements from the same fetus will result in an inappropriate estimate of residual variance. Deter et al. (1982a) were the only other authors who had analysed their longitudinal data using least squares fitting to obtain growth curves for individual fetuses. However, in contrast to the methodology used here, they derived their summary growth curve and normal ranges using the mean value and standard deviation of each coefficient. This would have led to an inappropriate estimate of the standard deviation.

Table 5.7 Comparison of mean (and -2 SD) values of AC from published data.

Reference	N	GESTATIONAL AGE (in weeks)		
		28	34	40
Jeanty et al. (1984a)	48	225 (203*)	279 (256*)	316 (294*)
Deter et al. (1982a)	20	240 (218)	308 (250)	377 (341)
Larsen et al. (1990)	35	247 (224)	314 (286)	369 (340)
Present study	67	245 (224)	310 (288)	355 (330)

* 5th centile criteria
All other values are mean (-2 SD).

Corresponding values of EFW at 28, 34 and 40 weeks gestation in the different studies are shown in Table 5.8. The results reported by Jeanty et al. (1984b) were again much lower. The mean birthweight in that study was not significantly different from the final EFW, indicating that the Hadlock formula did not systematically over- or under-

estimate weight. The mean EFW at 40 weeks gestation reported by Larsen et al. (1990) was also very close to their mean birthweight (3660 g) (mean gestation 282 days). In contrast, the mean EFW at 40 weeks gestation in the study of Deter et al. (1982a) was 471 g more than the mean birthweight (mean gestation 282 days). The reason for this difference is uncertain; it is unlikely to be related to the Warsof formula used by Deter et al. (1982a) since prospective studies have not demonstrated an overestimation of birthweight with this formula. Furthermore the same formula was used by Larsen et al. (1990).

Table 5.8 Comparison of mean (and -2 SD) values of EFW from published data.

Reference	N	EFW Formula	GESTATIONAL AGE (in weeks) 28	34	40
Jeanty et al. (1984b)	48	Shepard	1288 (802)	2369 (1460)	3131 (1887)
Persson and Weldner (1986)	19	Persson	1288 (880)	2351 (1811)	3589 (2819)
Deter et al. (1982a)	20	Warsof	1239	2363	3863
Larsen et al. (1990)	35	Warsof	1282 (981)	2454 (1856)	3584 (2794)
Present study	67	Hadlock	1256 (1011)	2445(2014)	3633 (3031)

All values are mean (-2 SD).

The finding that small fetuses at 28 weeks gestation had a smaller increase in EFW at 36 weeks supported the findings of Persson (1989) who found that, between 32 and 37 weeks gestation, daily weight increment was lower (19.4g) in fetuses born SGA compared with those born AGA (29.2g). A comparable reduction in AC growth in SGA fetuses was reported by Divon et al. (1986). This phenomenon may partly explain the increasing variance of both AC and EFW with gestational age.

Two previous groups of workers have reported reference ranges for rate of growth of AC in the second and third trimesters (Fescina et al. 1982, Deter and Harrist 1992). Both these studies derived their reference ranges using smaller cohorts of fetuses (n = 30

and n = 20 respectively) compared with the present study (n = 43). Unlike the present study, Fescina et al. (1982) used smoothed data to within ± 7 days of defined gestations. This would have led to inappropriate estimations of standard deviation (Laird and Ware 1982). Nevertheless, mean rates of growth of AC reported by Fescina et al. (1982) were comparable to those in this study up to 36 weeks gestation. Thereafter mean rates of growth reported by Fescina et al. (1982) were much smaller than those in the present study. Deter and Harrist (1992) derived reference ranges for rate of growth of AC using the Rossavik Growth Model and Hotellings T^2-statistic, a multivariate form of the t-test. Mean rates of growth in their study were higher than those reported in the present study from 32 weeks gestation, with no evidence of any decrease in growth velocity towards term. The SD reported by Deter and Harrist (1992) were also considerably smaller than those reported in the present study at all gestational ages. This may be due to their smaller sample size (n = 20) compared to the present study (n = 43).

The reference ranges presented in this study were derived from the largest group of fetuses used to collect serial ultrasound data. Reference ranges for AC and EFW were constructed using the quadratic coefficients of individual fetuses, thereby avoiding the inappropriate use of regression analysis. Reference ranges for rates of growth were derived from the data of fetuses to within ± 2 days of specific gestational ages. Such reference ranges provided suitable standards for the evaluation of serial ultrasound measurements in Study 3 (described in Chapter 7).

5.5 Summary

Serial ultrasonic measurements were performed in 67 Caucasian fetuses from 20 weeks gestation until term to derive reference standards for AC and EFW. The variances of both AC and EFW increased with gestational age. Four mathematical models (linear, quadratic, Gompertz and Rossavik) were fitted to the $\log_{10}$(AC) and $\log_{10}$(EFW) data from each fetus using least squares regression analysis. The standard deviations of the residual error were greatest for the linear model. The linear model also overestimated the final $\log_{10}$(AC) while the linear, Gompertz and Rossavik models all overestimated the final $\log_{10}$(EFW) when these data were omitted from the fitting process. The individual quadratic constants were therefore used to produce 67 individual growth curves. Values of $\log_{10}$(AC) and $\log_{10}$(EFW) for each fetus at exact gestational ages between 20 and 40 weeks were derived by interpolation; the mean and standard deviation values were then used to derive percentile ranges for AC and EFW. Reference ranges for rate of growth of AC and EFW were also derived from the unsmoothed data of 43 fetuses to within ± 2 days of exact gestational ages. These reference ranges provided suitable standards for the subsequent evaluation of serial ultrasound values of AC and EFW in the diagnosis of IUGR.

CHAPTER 6

REPRODUCIBILITY OF ULTRASOUND AND NEONATAL MORPHOMETRIC MEASUREMENTS

6.1 Reproducibility of ultrasound measurements

6.1.1 Introduction

The appropriate interpretation of fetal growth requires an understanding of the inherent variability of ultrasonic measurements (Sarmandal et al. 1989). A number of previous studies have reported the reproducibility of standard ultrasound measurements using a correlation coefficient or a coefficient of variation (Weiner et al. 1981, Fescina et al. 1982, Larsen et al. 1990). Several workers have highlighted the limitations of these methods of statistical analysis when assessing the repeatability of clinical measurements (Bland and Altman 1986, Brennan and Silman 1992). Analysis of variance (Healy 1989) is the preferred method of evaluating intra-observer variability of ultrasound measurements (Bland and Altman 1992). The limits of agreement method, first reported by Bland and Altman (1986), is a more appropriate method of assessing inter-observer variability (Bailey et al. 1988, Sarmandal et al. 1989). No study has investigated the reproducibility of ultrasonic EFW using these methods. There was therefore a need to assess the intra- and inter-observer variability of standard ultrasound measurements and EFW in third trimester fetuses using analysis of variance and the limits of agreement method respectively.

6.1.2 Subjects and methods

Forty fetuses in the third trimester of pregnancy (mean gestational age 34.5 (SD 3.3) weeks) were used to evaluate the reproducibility of ultrasound measurements. Thirty fetuses were referred from the ultrasound department in view of an AC less than the 10th centile according to the charts of Deter et al. (1982a) and all subsequently had a birthweight less than the 10th centile (Thomson et al. 1968). The remaining 10 fetuses were part of the study group used to derive normal reference ranges for AC and EFW; all subsequently had a birthweight between the 10th and 90th centiles.

Ultrasonic measurements were performed by two observers, the author and an independent observer (SCR). Each observer measured BPD, HC, AC, and FL three times. Measurements were blinded to the observer by covering the measurement display on the screen. Each measurement was then stored in the instrument computer and printed out after completion of all measurements. Neither observer was aware of the others measurements. Estimated fetal weights were calculated with each set of measurements using the formula of Hadlock et al. (1985).

6.1.3 Results

Satisfactory measurements were obtained by both observers in all 40 fetuses. Mean (SD) gestational age at ultrasound assessment was 238 (8.0) days. The intra-observer standard deviations, calculated from the residual variance, are shown in Table 6.1. For each observer the values for BPD and FL were < 1 mm and for AC and HC < 5 mm. Corresponding values for EFW were < 75 g.

Table 6.1 Intra-observer variability of ultrasound measurements (expressed as SD's)

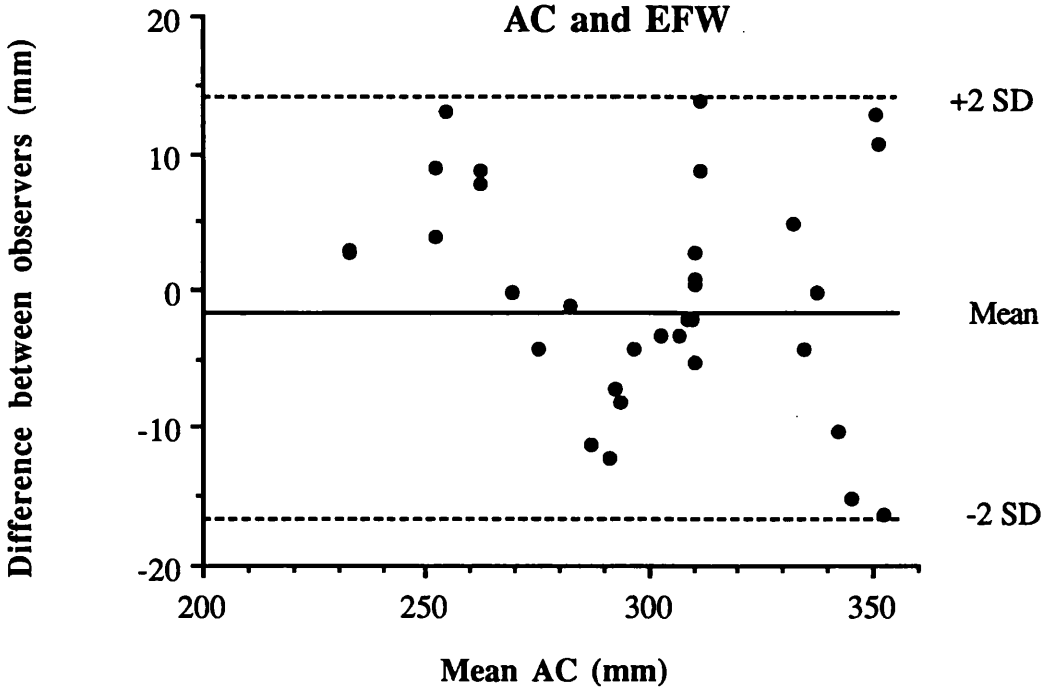
	Observer 1	**Observer 2**
BPD (mm)	0.69	0.67
HC (mm)	4.1	3.58
AC (mm)	4.29	4.29
FL (mm)	0.95	0.83
EFW (g)	64.1	62.71

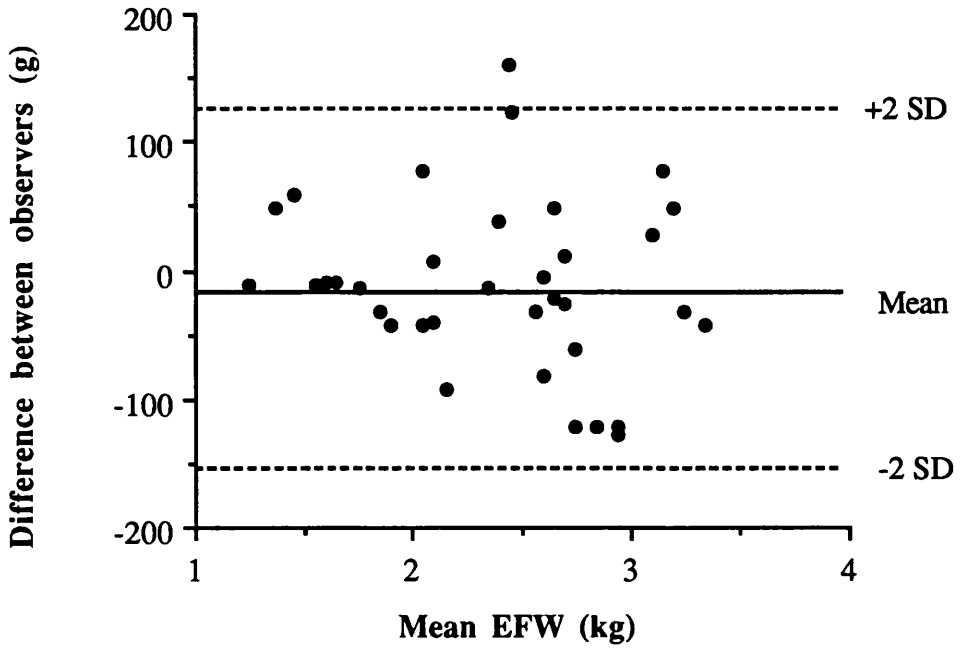
The mean (±2 SD) inter-observer difference for each ultrasound measurement was as follows: BPD -0.6 (-1.2, 0.0) mm, HC -1.6 (-3.5, 0.3) mm, AC -1.9 (-4.4, 0.5) mm, FL -0.1 (-0.6, 0.4) mm and EFW -17.8 (-50.4, 3.0) g. The mean difference for BPD was significantly different from zero, suggesting a systematic under-estimation (by 0.6 mm) by one of the observers. No systematic bias was observed for any of the other measurements. The mean absolute and percentage differences between observers, together with the 95% prediction intervals for all parameters, are shown in Table 6.2. The absolute limits of agreement for inter-observer comparisons of AC and EFW are shown in Figure 6.1. The 95% prediction interval for AC was (-16.8 to 13.0 mm) and for EFW (-159.9 to 124.3 g).

Table 6.2 Inter-observer variability of ultrasound measurements: mean absolute and percentage differences with 95% limits of agreement

	Mean Difference		Limits of agreement	
	Absolute	**Percentage**	**Absolute**	**Percentage**
BPD (mm)	-0.6	-0.68	(-4.1, 2.9)	(-4.7, 3.4)
HC (mm)	-1.6	-0.51	(-13.2, 10.1)	(-4.3, 3.3)
AC (mm)	-1.9	-0.54	(-16.8, 13.0)	(-5.4, 4.3)
FL (mm)	-0.1	-0.15	(-2.9, 2.8)	(-4.6, 4.3)
EFW(g)	-17.8	-0.69	(-159.9, 124.3)	(-6.3, 4.9)

Figure 6.1 Inter-observer variability: Limits of agreement for AC and EFW





6.1.4 Discussion

The use of serial ultrasound values of AC and EFW to evaluate fetal growth necessitates the variability of such measurements to be calculated. The results of the present study suggest that intra-observer variability was similar for each of the two observers and was consistently less than inter-observer variability. The magnitude of inter-observer variability has considerable implications on the interpretation of serial ultrasound measurements in clinical practice.

Deter et al. (1986) reviewed 139 studies of ultrasound assessment of fetal growth; in only 19 of them was inter-observer variability assessed, and none by the limits of agreement method. Previous studies on the reproducibility of AC have generally reported results either as a coefficient of variation (standard deviation of the differences between 2 sets of measurements expressed as a percentage of the mean of the measurements) or a correlation coefficient. All previous studies have reported high correlation coefficients ($\geq$ 0.95) (Clement et al. 1981, Weiner et al. 1981, Sarmandal et al. 1989) and low coefficients of variation (< 5%) (Campbell 1976, Hadlock et al. 1982a, Hadlock et al. 1982b, Larsen et al. 1990), suggesting AC to be a highly reproducible ultrasound measurement. The only previous study on the reproducibility of EFW also reported a low coefficients of variation ($\leq$ 7%) (Larsen et al. 1990). Analysis of reproducibility using either the correlation coefficient or coefficient of variation would appear to confirm that AC and EFW are reproducible ultrasound parameters.

However, neither of these methods actually assesses the degree of agreement between measurements (Bland and Altman 1986, Breenan and Silman 1992, Bland and Altman 1992). Measurements which are in poor agreement with one another can still show a misleadingly high degree of correlation (Bland and Altman 1986). This is because the correlation coefficient simply assesses the association between two sets of measurements (Brennan and Silman 1992). For example, one set of measurements consistently recorded to be twice the value of another set of measurements would result in a high correlation coefficient. For this reason, the limits of agreement method (Bland and Altman 1986) has been suggested as a superior method of assessing the repeatability of measurements between observers. This allows the calculation of 95% limits of inter-observer differences, within which 95% of all differences between observers are expected to lie.

The superiority of the limits of agreement method for the analysis of inter-observer variability was shown by Sarmandal et al. (1989) who studied the inter-observer variability of ultrasound measurements using all three methods. They reported the coefficient of variation and correlation coefficient for AC to be 2.7% and 0.98 respectively; however the the limits of agreement were wide (-22, 20 mm). Only one other

study has published the actual ultrasound measurements of AC used to assess reproducibility, such that correlation coefficients, coefficients of variation and limits of agreement could be calculated from the original data (Tamura et al. 1980). In that study, the correlation coefficient and coefficient of variation were calculated to be 0.97 and 4.2% respectively; yet the limits of agreement were wide (-34, 28 mm). The size of the limits of agreement in these studies can be appreciated by referring these limits to the centile charts reported in Study 1 (Table 6.3). It is evident that despite the favourable correlation coefficients and coefficients of variation reported by Tamura et al. (1981) and Sarmandal et al. (1989), the limits of agreement when referred to mean values of AC at 34 weeks gestation were wider than the centile values at the 5th and 95th centiles. Mean values at 34 weeks were chosen as an illustration as this was the mean gestational age for measurements of reproducibility in this study. In view of this and other studies which have reported similar discrepancies (Bailey et al. 1988), inter-observer variability was reported in this study using only the limits of agreement.

Table 6.3 Effect of limits of agreement of AC at 34 weeks gestation

Centiles *	**AC (mm)**	**Limits of agreement**		
		Tamura et al. (1980)	Sarmandal et al. (1989)	Present study
1st	284			
5th	292			
10th	296			
50th	310	(-34, 28mm)	(-22, 20 mm)	(-17, 13 mm)
90th	326	= 276 to 338 mm	= 288 to 330 mm	= 293 to 323 mm
95th	330			
99th	339			

* Reference ranges derived in Study 1 (see Section 5.3).

In the present study, the limits of agreement for inter-observer comparisons were all less than ± 8%. Sarmandal et al. (1989) are the only other group who have reported comparable data. Their limits of agreement [BPD (-8.0, 7.0 mm); HC (-24, 24 mm); AC (-22, 20 mm); FL (-7.4, 4.9 mm)] were much wider than those reported in this study, especially with regards to values for AC. The magnitude of the difference between their results and those reported in this study for AC can be appreciated by referring the respective limits of agreement to mean value of AC at 34 weeks gestation (Table 6.3).

These differences are unlikely to be related to observer experience but may be partly explained by the fact that they averaged 2, rather than 3 measurements.

Sarmandal et al. (1989) suggested that the use of multiple parameters for calculating EFW would compound the inaccuracies of each ultrasonic parameter. However, they did not report any data to support this. No previous workers have investigated the reproducibility of EFW using the limits of agreement method. The results in this study showed that although percentage differences for inter-observer comparisons of EFW were larger than those for the individual ultrasonic parameters, the differences were not great. Table 6.4 shows the effects of referring the limits of agreement for EFW to the mean value of EFW at 34 weeks gestation reported in Study 1. The limits of agreement for EFW were narrower relative to the respective centiles than the corresponding limits of agreement for AC (Tables 6.3 and 6.4).

Table 6.4 Effect of limits of agreement of EFW at 34 weeks gestation

Centiles *	EFW (g)	Limits of agreement
1st	1962	
5th	2096	
10th	2170	
50th	2445	(-160, 124 g)
90th	2754	= 2285 to 2569 g
95th	2851	
99th	3046	

* Reference ranges derived in Study 1 (see Section 5.3).

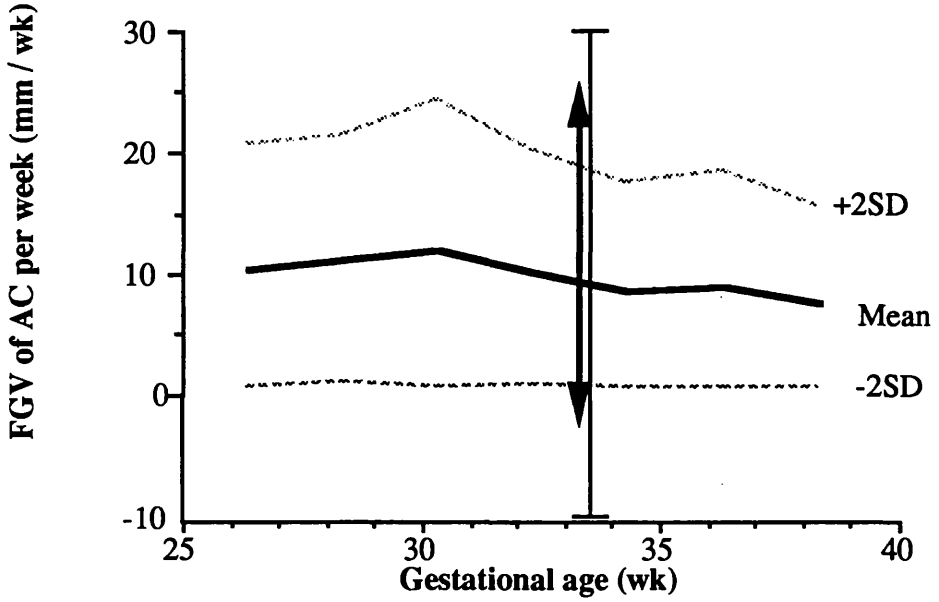
The magnitude of the 95% prediction intervals needs to be taken into consideration when making inferences regarding serial changes in AC and EFW. For any ultrasound measurement, changes greater than the 95% prediction interval cannot be explained by measurement variability and therefore are likely to represent true fetal growth. The limits of agreement for AC reported in this study (-17, 13 mm) and by Sarmandal et al. (1989) [-22, 20 mm] were greater than the ± 2 SD limits (± 8.8 mm) for weekly increment in AC at 34 weeks gestation reported in Study 1 (Figure 6.2). The limits of agreement for EFW (-160, 124 g) were of the same magnitude as the ± 2 SD limits (± 178g) for weekly increment in EFW at 34 weeks gestation (Figure 6.3). These results suggest that the serial ultrasound assessment of AC and EFW at weekly intervals is unlikely to reflect accurate

growth. In the light of these findings, serial values of AC and EFW were quantified in Chapter 7 using ultrasound data obtained at intervals of at least two weeks.

6.1.5 Summary

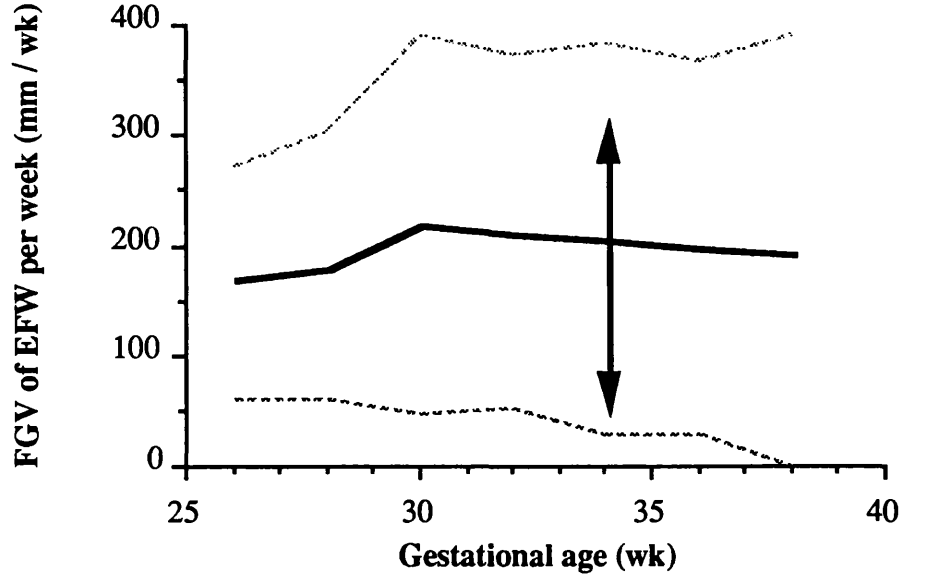
Standard ultrasound measurements were performed by 2 observers in 40 third trimester fetuses. Observers were blinded to the results of the measurements. Estimated fetal weight was calculated using the formula of Hadlock et al. (1985). The intra-observer standard deviation for EFW, assessed using one-way analysis of variance, was < 75 g for both observers. The 95% prediction intervals for inter-observer comparisons of AC and EFW, calculated using the limits of agreement method, were (-16.8 to 13.0 mm) and (-159.9 to 124.3 g) respectively. The results suggest that the reproducibilities of AC and EFW are clinically acceptable. Comparison of the limits of agreement of AC and EFW with reference ranges for increment in AC and EFW showed that an interval of not less than two weeks has to lapse between ultrasound scans before any meaningful interpretation about fetal growth can be made.

Figure 6.2 Limits of agreement relative to reference range for fetal growth velocity of AC



Limits of agreement reported by Sarmandal et al. (1989)
Limits of agreement reported in present study

Figure 6.3 Limits of agreement referred to reference ranges for fetal growth velocity of EFW



Limits of agreement reported in the present study

6.2 Reproducibility of neonatal morphometric measurements

6.2.1 Introduction

The clinical value of neonatal morphometric measurements for the definition of IUGR depends to a large extent on the reproducibility. Studies which have published reference standards for ponderal index (Miller and Hassanein 1971), MAC / HC ratio (Sasanow et al. 1986) and subscapular and triceps skinfold measurements (Oakley et al. 1977b) have reported reproducibility of such measurements using either the correlation coefficient and / or coefficient of variation. However, there are many limitations with the use of such statistical methods in the assessment of agreement between measurements (Bland and Altman 1986, Brennan and Silman 1992). There was therefore a need to determine the intra- and inter-observer variability of standard neonatal morphometric measurements using analysis of variance (Healy 1989) and limits of agreement (Bland and Altman 1986) respectively.

6.2.2 Subjects and Methods

Thirty neonates were studied between days 1-3 after birth. All were part of the study group with ultrasound evidence of smallness [AC < 10th centile (Deter et al. 1982b)]. The mean gestational age at delivery was 39.6 (SD 2.0) days and the mean birthweight was 2586 (SD 265) g. All had a birthweight below the tenth centile (Thomson et al. 1968).

Morphometric measurements were performed by the author and an independent observer (SCR), both of whom have been trained in the anthropometric assessment of neonates. Measurements of weight, crown-heel length, MAC, HC and subscapular and triceps skinfold thickness were made using techniques as described in Section 4.2.1. With the exception of weight, three readings of each of the above measurements were made by each observer. Neither observer was aware of the other's readings.

The 95% inter-observer prediction intervals for ponderal index, MAC / HC ratio and subscapular and triceps skinfold thickness were then super-imposed on the respective reference ranges (Miller and Hassanein 1971, Sasanow et al. 1986, Oakley et al. 1977b).

6.2.3 Results

A complete set of readings was obtained in all 30 neonates. The intra-observer SD's for each morphometric measurement are shown in Table 6.5. The variability of each measurement was small and similar for each of the two observers.

Table 6.5 Intra-observer variability of morphometric measurements as assessed by one-way analysis of variance.

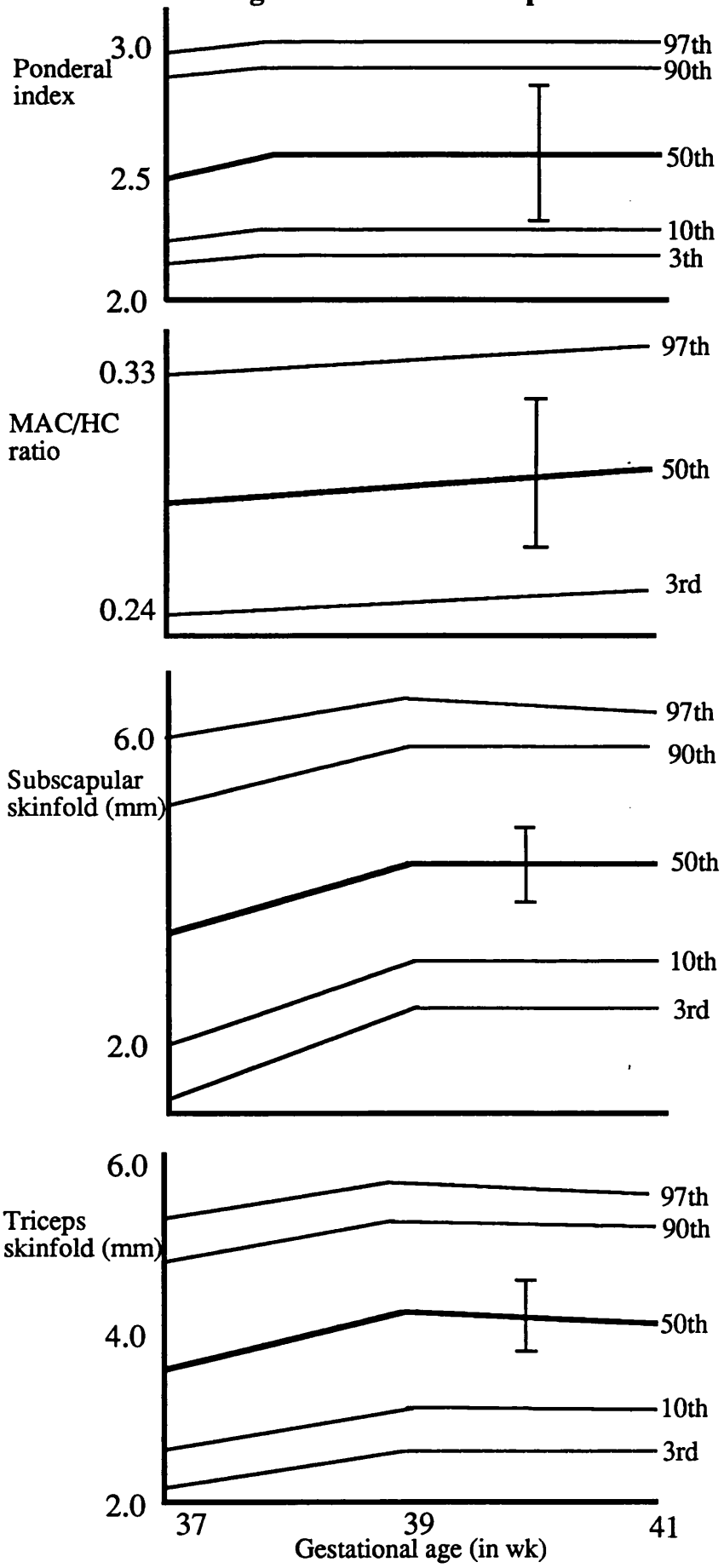
	Intra-observer variability (Standard deviation)	
Measurement	**Observer 1**	**Observer 2**
Length (mm)	2.1	2.5
Ponderal index	0.03	0.04
Head circumference (mm)	1.7	2.1
MAC (mm)	1.8	1.3
MAC / HC	0.006	0.005
Subscapular skinfold (mm)	0.12	0.12
Triceps skinfold (mm)	0.13	0.17

The mean inter-observer differences, together with the 95% prediction intervals are shown in Table 6.6. Significant differences between observers were found for triceps skinfold, length and ponderal index. Observer 1 underestimated triceps skinfold and overestimated length and ponderal index with respect to observer 2. The 95% prediction intervals were plotted with respect to the centile reference ranges for ponderal index, MAC / HC ratio and subscapular and triceps skinfold thickness in Figure 6.4. It is evident that all three neonatal morphometric indices had narrow limits of agreement with respect to the reference ranges.

Table 6.6 Inter-observer variability of morphometric measurements, as assessed by the limits of agreement method.

Measurement	Mean difference (95% CI)	95% prediction interval
Length (mm)	3.40 (0.52, 6.88)	(-12.14, 18.9)
Ponderal index	0.05 (0.01, 0.09)	(-0.23, 0.28)
Head circumference (mm)	-0.78 (-2.48, 0.92)	(-9.96, 8.40)
MAC (mm)	-0.33 (-2.20, 1.54)	(-10.39, 9.73)
MAC/HC	0.00 (0.00, 0.00)	(-0.02, 0.02)
Subscapular skinfold (mm)	0.08 (-0.02, 0.19)	(-0.51, 0.68)
Triceps skinfold (mm)	-0.18 (-0.09, -0.27)	(-0.65, 0.29)

Figure 6.4 Limits of agreement relative to reference ranges for neonatal morphometric indices



Reference range derived from Miller and Hassanein (1971).

Reference range derived from Sasanow et al. (1986).

Reference range derived from Oakley et al. (1977b) (measurements in mm).

Reference range derived from Oakley et al. (1977) (measurements in mm).

6.2.4 Discussion

Intrauterine growth retardation is best defined using neonatal morphometric indices of malnutrition (Section 2.2.5). The reproducibility of ponderal index, MAC / HC ratio, subscapular and triceps skinfold thicknesses is of crucial importance in the choice of these indices as the "gold standard" for IUGR. Previous studies have analysed the reproducibility of these indices using inappropriate methods of analysis. The present study showed that intra-observer variability of all morphometric measurements was smaller than corresponding inter-observer differences. The inter-observer variabilities of all neonatal morphometric indices were clinically acceptable.

Care was taken to perform these morphometric measurements in a standardized fashion to reduce methodological errors. This applied particularly to the use of the Holtain calipers where the timing of measurement, with respect to caliper placement, affects skinfold thickness (Brans et al. 1974, Oakley et al. 1977b). The sites of skinfold measurements were not marked as this has not been shown to reduce variability (McGowan et al. 1975). Although each observer was aware of his own measurements, this information was not available to the other observer. Thus whilst intra-observer variability may have been reduced by such knowledge, no such bias was possible with inter-observer comparisons.

Few studies have reported data on the reproducibility of the ponderal index. In three of the most important studies on the use of the ponderal index in the diagnosis of IUGR, no information was given about any reproducibility of length measurements (Miller and Hassanein 1971, Patterson and Pouliot 1987a, Fay et al. 1991a). Two studies which presented data using the coefficient of variation suggested that the inter-observer variability of ponderal index was low (Wolfe et al. 1990, Catalano et al. 1992). The coefficients of variation were reported to be 11% and < 6% respectively in these two studies. One study reported the intra-observer coefficient of variation for length measurements to be 0.75% (Colley et al. 1991). However, these statistical tests present no information on the clinically relevant limits of agreement (Bland and Altman 1986, Brennan and Silman 1992). In the present study, the intra-observer variability of length and ponderal index were both low. The limits of agreement for ponderal index (-0.23, 0.28) were narrow relative to the centiles for ponderal index (Figure 6.4), confirming the reproducibility of this index.

Three studies have previously published data on the reproducibility of measurements of MAC and HC in the neonate. In the definitive study where reference standards of MAC / HC at various gestational ages were reported, the inter-observer coefficient of variation of MAC measurements was reported to be 2%. No data on the

reproducibility of HC was reported. Meadows et al. (1986) reported intra-observer and inter-observer coefficients of variation for the MAC / HC ratio to be 2.0% and 2.5% respectively. A similarly low inter-observer coefficient of variation of 2.46% was reported by Excler et al. (1985). In the present study, the intra-observer standard deviation for MAC, HC and the ratio were low. The narrow limits of agreement for the MAC / HC ratio (Figure 6.4) confirmed the reproducibility of this morphometric index.

Some of the most well-known studies on skinfold thickness measurements in the neonate did not report data on the reproducibility of measurements (Dauncey et al. 1977, Udall et al. 1978, Whitelaw 1979). Two studies assessed intra-observer variability using analysis of variance. In the study of Oakley et al. (1977b), the ± 2 SD for intra-observer variability of subscapular and triceps skinfold measurements were ± 4.5% and ± 2.5% respectively. In another study, the ± 2 SD limits for intra-observer variability were 19% and 24% at the triceps and subscapular sites respectively (McGowan et al. 1975). The ± 2 SD limits were also presented as absolute values (1.04 and 1.37 mm respectively). These results suggest a much greater degree of intra-observer variability than those reported by Oakley et al. (1977b). The results of the present study showed that the intra-observer standard deviation for subscapular and triceps skinfolds were low (≤ 0.17 mm). These values were considerably smaller than those reported by McGowan et al. (1975). Percentages for standard deviations were not calculated for comparisons with the results of McGowan et al. (1975) or Oakley et al. (1977b) as it was unclear how these were calculated in those two studies. All other studies have reported data on intra-observer variability using either a correlation coefficient or coefficient of variation (Branson et al. 1982, Weile et al. 1986).

Three previous studies have reported data on the inter-observer variability of skinfold measurements. Oakley et al. (1977b) reported that differences between observers for subscapular and triceps skinfolds were not more than 0.3 mm. In the study by Branson et al. (1982), inter-observer correlation coefficients and coefficients of variation for triceps and subscapular skinfolds were reported to be 0.84 and 0.77, and 13.0% and 16.4% respectively. In the study by Weile et al. (1986), the inter-observer coefficients of variation for subscapular and triceps skinfolds were 8.1 and 7.7% respectively. However as none of these studies included the standard deviation of the difference between observers, their results cannot be compared with those from the present study. The limits of agreement for subscapular and triceps skinfold thickness reported in this study were narrow compared with the centiles of the reference ranges (Figure 6.4), confirming the reproducibility of these measurements.

The advantage of quantifying inter-observer variability by the limits of agreement method was that the results could be directly compared with appropriate reference charts.

The results in Figure 6.4 demonstrated the reproducibility of all neonatal morphometric indices investigated and confirmed their use for defining IUGR.

6.2.5 Summary

Measurements of triceps and subscapular skinfold thickness, MAC, HC and crown-heel length were performed by two observers in 30 neonates. The intra-observer standard deviation for all measurements, calculated using one-way analysis of variance, was small and similar for each of the two observers. Inter-observer variability was assessed using the limits of agreement method. There were small systematic differences between the observers for triceps skinfold, length and ponderal index. The reproducibility of these morphometric measurements was confirmed by comparing the 95% prediction intervals for subscapular thickness (-0.51, 0.68 mm), triceps thickness (-0.65, 0.29 mm), MAC / HC ratio (-0.02, 0.02) and ponderal index (-0.23, 0.28) with the respective reference ranges.

CHAPTER 7

DIAGNOSIS OF INTRAUTERINE GROWTH RETARDATION USING SERIAL ULTRASOUND VALUES OF ABDOMINAL CIRCUMFERENCE AND ESTIMATED FETAL WEIGHT

7.1 Introduction

Although serial ultrasound values of AC and EFW are used in the everyday clinical management of fetuses suspected of IUGR, the usefulness of such a practice has rarely been subjected to any vigorous quantitative evaluation. There was a need to evaluate the predictive ability of serial values of AC and EFW in the diagnosis of IUGR. The reference ranges for AC and EFW reported in Chapter 5 provided the standards for such a study..

The most appropriate method of quantifying serial ultrasound values of AC and EFW was determined. This optimal measure of serial AC or EFW data was then compared with umbilical artery PI and single estimates of fetal size in the prediction of abnormal neonatal morphometry and adverse perinatal outcome in small fetuses. Finally, the ability of serial ultrasound data to separate small fetuses into two groups with distinctly different perinatal outcome and biochemical indices was assessed.

7.2 Subjects

A study was constructed to evaluate the growth profile of small fetuses in the third trimester. The entry criteria into this study were a) the women must have had certain menstrual dates b) all fetuses must have had an anomaly scan between 18 to 20 weeks gestation (level II scan) during which gestational age assessment using measurements of BPD and FL must not be more than 7 days discrepant of menstrual dates c) all fetuses had an AC < 10th centile for gestational age according to the charts of Deter et al. (1982a) at entry into the study in the third trimester.

One hundred and forty eight women in the third trimester of pregnancy were referred to the author for antenatal surveillance in the Fetal Medicine Unit at University College Hospital, London between January 1991 and July 1992 because of a suspected small fetus on ultrasound. All had been scanned by radiographers in the Ultrasound Department and found to have an AC < 10th centile for gestational age according to the charts of Deter et al. (1982a). All were subsequently rescanned by the author on the same day; in 113 cases, the same findings were confirmed by the author. Thirty five women in whom the author found the AC to be > 10th centile (Deter et al. 1982a) were therefore excluded from the analysis. Five women were excluded because ultrasound dates at the level II scan were more than 7 days discrepant of menstrual dates. Four women were excluded because of delivery before 36 weeks gestation. The data from the remaining 104 third trimester fetuses were subsequently used in the final analysis.

The demographic and delivery data are shown in Table 7.1. Fifty four of the subjects were primigravidae. Of the 50 multigravidae, 34 (68%) had a previous history of a SGA infant (birthweight < 10th centile according to the charts of Thomson et al. 1968). A quarter of all subjects were smokers and 19 (18.3%) had either pregnancy-induced hypertension or pre-eclampsia in the present pregnancy.

Ninety four (90.4%) of the neonates had a birthweight < 10th centile and 69 (66.7%) of the neonates had a birthweight < 5th centile (Thomson et al. 1968). Forty nine (47.1%) of the pregnancies were terminated by induction of labour, either by administration of vaginal prostaglandin pessaries or by artificial rupture of membranes. Of the 104 fetuses, 15 (14.3%), 13 (12.2%), 28 (26.5%) and 33 (31.6%) had neonatal morphometric evidence of IUGR as defined by abnormal PI, MAC / HC ratio and subscapular and triceps skinfold thickness respectively. Twenty four (23.5%) had a total morphometric score ≥ 2. Twenty eight (26.9%) infants had adverse perinatal outcome, as defined by the presence of one or more of the following outcomes; acidaemia at birth (n = 24), emergency Caesarean section for fetal distress (n = 13) and admission to NICU for complications related to IUGR (n = 8). There were no perinatal deaths in the group.

Table 7.1 Demographic and delivery details of study group.

No. of fetuses	104
Maternal age (yr)	28.6 (4.9)
Primigravidae	54 (51.9%)
History of previous SGA†	34 (32.7%)
Present history of PIH	10 (9.6%)
Present history of PET	9 (8.6%)
Smokers	26 (25.0%)
Male infants	46 (44.2%)
Birthweight (g)	2550 (354)
Gestational age at delivery (days)	274 (8)
Vaginal deliveries	66 (63.5%)
Operative delivery for fetal distress	24 (23.1%)
Umbilical arterial pH	7.23 (0.08)
Umbilical arterial base excess	- 5.2 (3.3)
Apgar < 7 at 5 minutes	5 (4.8%)
NICU admissions	8 (7.7%)

Figures are mean (SD) or n (%).

†Birthweight < 10th centile (Thomson et al. 1968).

Pregnancy-induced hypertension (PIH) was defined as BP > 140 / 90; pre-eclamptic toxaemia (PET) was defined as BP > 140 / 90 and ≥ 2+ proteinuria.

7.3 Study design

After recruitment into the study, ultrasound scans were performed at weekly intervals. Each fetus was assessed at least three times; 88 (85%) of the 104 fetuses were assessed on at least four occasions. At each visit, ultrasound measurements of BPD, HC, AC, FL and umbilical artery PI were made and EFW calculated (see Section 4.1)

Three statistical methods of quantifying serial values of AC and EFW were evaluated in the prediction of abnormal neonatal morphometry (see Section 4.5.3). The best measure of serial AC or EFW data was then compared with umbilical artery PI and estimates of fetal size in the prediction of abnormal neonatal morphometry and adverse perinatal outcome. The umbilical artery PI at the last ultrasound scan prior to delivery was expressed as a SDS; PI.SDS = (Measured PI - Mean PI at same gestation) / (SD of PI at same gestation). The mean and SD at different gestations were obtained from the regression equation reported by Pearce et al. (1988). Estimates of fetal size were obtained by calculating the SDS of the last EFW prior to delivery; Last EFW.SDS = (Last EFW - Mean EFW at same gestation) / (SD of EFW at same gestation).

Values of neonatal ponderal index, MAC / HC ratio and subscapular and triceps skinfold thickness more than 2 SD below the respective means (Miller and Hassanein 1971, Sasanow et al. 1986, Oakley et al. 1977b) were used to define IUGR. A total neonatal morphometric score was also calculated by according values of 0 or 1 to individual morphometric indices if they were within 2 SD or > -2 SD respectively, resulting in a total morphometric score ranging from 0 (no evidence of IUGR using any of the four morphometric indices) to 4 (evidence of IUGR with all four morphometric indices). A total morphometric score of 2 or more was regarded as abnormal for the purposes of defining IUGR. Adverse perinatal outcome was defined as the presence of one or more of the following: acidaemia at birth, emergency Caesarean section for fetal distress in labour and admission to NICU because of morbidity associated with IUGR.

The best measure of serial AC or EFW data was then used to divide the small fetuses into two groups, those with ultrasonic evidence of normal growth and those with evidence of impaired growth. These neonates were classified using the optimal cut-off level as determined from the ROC curve. Perinatal morbidity and biochemical indices of IUGR were then compared in these two groups.

All measurements of AC were reported to the clinicians in charge of the patients using the standard ultrasound reporting sheets used in the Fetal Medicine Unit. Subsequent management decisions were made in the light of revealed data.

7.4 Results

7.4.1 Quantification of serial values of AC and EFW

A total of 557 scans were performed on 104 fetuses with a median of 5.0 (range 3.0 to 13.0) scans per fetus. Satisfactory measurements were obtained in all cases. The median gestational age at the first ultrasound assessment was 220 (range 182 to 270) days. The distribution of gestational ages at the first ultrasound assessment is shown in Figure 7.1. The median gestational age at the last scan was 269 (range 238 to 294) days. The median interval between the first and the last scan was 28 (range 14 to 91) days. A distribution plot of these intervals is shown in Figure 7.2. The median interval between the last scan and delivery was 5 (range 0 to 14) days. A regression analysis of the difference between birthweight and last EFW, and the mean of the two revealed no significant relationship.The mean difference between the final EFW and birthweight was -105.8 (SE 31.4 g). The distribution of birthweights and gestational ages at delivery are shown in Figures 7.3 and 7.4 respectively.

Correlations between the different ultrasound measures and neonatal morphometry were assessed using linear regression analysis and the results are shown in Table 7.2. Change in the SDS of AC and EFW (Δ AC.SDS and Δ EFW.SDS) were the only ultrasound measures which showed a significant relationship ($p < 0.05$) with all four neonatal morphometric indices, with R-values ranging from 0.24 to 0.40 for Δ AC.SDS and 0.30 to 0.40 for Δ EFW.SDS. In contrast, serial measurements of AC as assessed by the *b* -coefficient or by change in the SDS of the fetal growth velocity (FGV.AC.SDS) showed no correlation with any of the neonatal morphometric indices. Serial values of EFW as assessed by the *b* -coefficient and FGV.EFW.SDS demonstrated a significant relationship with some but not all of the morphometric indices.

Receiver operating characteristic curves were constructed for each ultrasound measure in the prediction of PI, MAC / HC ratio, subscapular skinfold thickness and abnormal total morphometric score (Figures 7.5 to 7.8 respectively). The areas under the ROC curves and respective standard errors are shown in Table 7.3. The ultrasound measures Δ AC.SDS and Δ EFW.SDS resulted in the largest area under the ROC curves in the prediction of all morphometric indices. In particular, the areas under the ROC curves resulting from the use of Δ EFW.SDS were significantly larger than those achieved by *b* -coef.EFW or FGV.EFW.SDS, in the prediction of each of the four outcomes. Comparisons of Δ AC.SDS, *b* -coef.AC and FGV.AC.SDS also showed similar trends although the differences only reached statistical significance in the prediction of MAC / HC ratio.

The superiority of Δ AC.SDS and Δ EFW.SDS over the other ultrasound measures was further confirmed when sensitivities, specificities, Cohen's kappa values (Table 7.4) and odds ratios (Table 7.5) were calculated using the optimal cut-offs derived from the ROC curves. Although the sensitivities were broadly similar with different ultrasound measures, the specificities, Cohen's Kappa and ORs were greater using Δ AC.SDS and Δ EFW.SDS. Odds ratios using these two ultrasound measures were significantly greater than zero, irrespective of the outcome criteria used. In contrast, none of the odds ratios for FGV.AC.SDS or FGV.EFW.SDS were significantly different from zero and all Kappa values were ≤ 0.13. All Kappa values for *b* -coef.AC and all but one for *b* -coef.EFW were < 0.2.

Similar results were obtained using a standard ultrasound cut-off level of ≥ -2 SD (Tables 7.6 and 7.7). The ultrasound measures with the highest sensitivities were *b* -coef.AC and *b* -coef.EFW (Table 7.6). However, these high sensitivities were achieved at the expense of unacceptably low specificities (all <37%). Although Δ AC.SDS and Δ EFW.SDS had lower sensitivities, the specificities were considerably higher (all > 82%). The superior diagnostic value of Δ AC.SDS and Δ EFW.SDS was further confirmed by the higher Kappa values (all ≥ 0.35) (Table 7.6) and higher odds ratios (Table 7.7), compared with all other ultrasound measures. The *b* -coefficients and FGV.SDS all had Kappa values below 0.2 and odds ratios which were not significantly different from zero.

These results show that Δ EFW.SDS and, to a lesser extent Δ AC.SDS, were consistently superior to other methods of quantifying serial ultrasound data in the prediction of abnormal neonatal morphometry. A histogram of the distribution of Δ EFW.SDS values is shown in Figure 7.9.

Table 7.2 Linear regression analysis of relationship between ultrasound measures and neonatal morphometry.

Ultrasound measure (y)	Morphometric index (x)	Equation	R-value	p
Δ AC.SDS	Ponderal index	$y = -4.56 + 1.57x$	0.32	0.001
	MAC / HC	$y = -7.7 + 25.2x$	0.40	0.000
	Subscapular	$y = -2.57 + 0.623x$	0.24	0.015
	Triceps	$y = -2.95 + 0.754x$	0.30	0.002
b -coef.AC	Ponderal index	-	-	NS
	MAC / HC	-	-	NS
	Subscapular	-	-	NS
	Triceps	-	-	NS
FGV.AC.SDS	Ponderal index	-	-	NS
	MAC / HC	-	-	NS
	Subscapular	-	-	NS
	Triceps	-	-	NS
Δ EFW.SDS	Ponderal index	$y = -4.71 + 1.67x$	0.37	0.000
	MAC / HC	$y = -7.14 + 23.5x$	0.40	0.000
	Subscapular	$y = -3.27 + 0.882x$	0.38	0.000
	Triceps	$y = -2.73 + 23.5x$	0.30	0.002
b -coef.EFW	Ponderal index	$y = 0.03 + 0.005x$	0.18	0.043
	MAC / HC	$y = 0.02 + 0.077x$	0.24	0.011
	Subscapular	$y = 0.035 + 0.0026x$	0.18	0.045
	Triceps	-	-	NS
FGV.EFW.SDS	Ponderal index	-	-	NS
	MAC / HC	$y = -4.04 + 10.1x$	0.28	0.008
	Subscapular	-	-	NS
	Triceps	-	-	NS

Table 7.3 Serial ultrasound: Comparison of areas under the ROC curves in the prediction of neonatal morphometry.

Ultrasound criteria	Outcome criteria	Area	Standard error
Δ AC.SDS	**Ponderal index ≥ -2.0 SD**	0.60	0.02
b-coef. AC		0.56	0.02 *
FGV.AC.SDS		0.53	0.02 *
Δ EFW.SDS		0.66	0.02
b-coef. EFW		0.58	0.02 *
FGV.EFW.SDS		0.55	0.03 *
Δ AC.SDS	**MAC / HC ≥ -2.0 SD**	0.67	0.02
b-coef. AC		0.53	0.02 *§
FGV.AC.SDS		0.51	0.03 *§
Δ EFW.SDS		0.66	0.02
b-coef. EFW		0.57	0.02 *§
FGV.EFW.SDS		0.49	0.03 *§
Δ AC.SDS	**SS ≥ -2.0 SD**	0.59	0.02
b-coef. AC		0.55	0.02 *
FGV.AC.SDS		0.50	0.03 *§
Δ EFW.SDS		0.66	0.03
b-coef. EFW		0.54	0.02 *
FGV.EFW.SDS		0.47	0.03 *§
Δ AC.SDS	**Total score ≥ 2**	0.62	0.02
b-coef. AC		0.56	0.02 *
FGV.AC.SDS		0.50	0.03 *§
Δ EFW.SDS		0.66	0.02
b-coef. EFW		0.56	0.02 *
FGV.EFW.SDS		0.48	0.03 *§

* $p < 0.05$ vs. Δ EFW.SDS.
§ $p < 0.05$ vs. Δ AC.SDS.

Abbreviatios; SS, subscapular skinfold.

Table 7.4 Serial ultrasound: Sensitivities, specificities and Cohen's kappa indices in the prediction of neonatal morphometry using optimal ultrasound cut-off criteria.

Ultrasound criteria	Outcome	Sensitivity	Specificity	Kappa
Δ AC.SDS ≥ -1.0	Ponderal index	69.2	67.1	0.22
Δ AC.SDS ≥ -1.5	MAC / HC	90.0	76.0	0.37
Δ AC.SDS ≥ -1.5	SS	72.0	56.3	0.22
Δ AC.SDS ≥ -1.0	Total score	59.1	68.7	0.23
b-coef. AC ≥ -8.0 SD	Ponderal index	64.3	59.1	0.12
b-coef. AC ≥ -8.0 SD	MAC / HC	58.3	57.7	0.07
b-coef. AC ≥ -3.0 SD	SS	84.0	41.7	0.17
b-coef. AC ≥ -3.0 SD	Total score	82.6	40.5	0.15
FGV.AC.SDS ≥ -1.0	Ponderal index	66.7	57.1	0.13
FGV.AC.SDS ≥ -1.0	MAC / HC	50.0	53.8	0.01
FGV.AC.SDS ≥ -1.0	SS	50.0	54.3	0.03
FGV.AC.SDS ≥ -1.0	Total score	56.2	55.9	0.08
Δ EFW.SDS ≥ -1.5	Ponderal index	76.9	80.3	0.41
Δ EFW.SDS ≥ -1.5	MAC / HC	80.0	78.5	0.35
Δ EFW.SDS ≥ -1.5	SS	72.0	64.1	0.30
Δ EFW.SDS ≥ -1.5	Total score	63.6	83.6	0.33
b-coef. EFW ≥ -8.0 SD	Ponderal index	60.0	77.1	0.29
b-coef. EFW ≥ -8.0 SD	MAC / HC	75.0	58.9	0.16
b-coef. EFW ≥ -8.0 SD	SS	56.0	58.3	0.11
b-coef. EFW ≥ -8.0 SD	Total score	65.2	60.8	0.20
FGV.EFW.SDS ≥ -1.0	Ponderal index	66.7	42.9	0.04
FGV.EFW.SDS ≥ -1.0	MAC / HC	70.0	43.1	0.05
FGV.EFW.SDS ≥ -1.0	SS	61.1	42.1	0.02
FGV.EFW.SDS ≥ -1.0	Total score	62.5	42.4	0.03

Abbreviation; SS, subscapular skinfold.

Table 7.5 Serial ultrasound: Odds ratios in the prediction of neonatal morphometry using optimal ultrasound cut-off criteria.

Ultrasound criteria	Outcome criteria	Odds ratio (and 95% CI)	
Δ AC.SDS ≥ -1.0	Ponderal index	4.6	(1.3, 16.6)
Δ AC.SDS ≥ -1.5	MAC / HC	28.4	(3.3, 244.2)
Δ AC.SDS ≥ -0.5	SS	3.3	(1.2, 9.1)
Δ AC.SDS ≥ 1.0	Total score	3.2	(1.2, 8.7)
b-coef. AC ≥ -8.0 SD	Ponderal index	2.6	(0.8, 8.5)
b-coef. AC ≥ -8.0 SD	MAC / HC	1.9	(0.6, 6.6)
b-coef. AC ≥ -3.0SD	SS	3.7	(1.1, 12.2)
b-coef. AC ≥ -3.0SD	Total score	3.2	(1.0, 10.6)
FGV.AC.SDS ≥ -1.0	Ponderal index	2.7	(0.7, 9.9)
FGV.AC.SDS ≥ -1.0	MAC / HC	1.2	(0.3, 4.5)
FGV.AC.SDS ≥ -1.0	SS	1.2	(0.4, 3.5)
FGV.AC.SDS ≥ -1.0	Total score	1.6	(0.5, 5.0)
Δ EFW.SDS ≥ -1.5	Ponderal index	13.6	(3.3, 56.2)
Δ EFW.SDS ≥ -1.5	MAC / HC	14.6	(2.8, 76.5)
Δ EFW.SDS ≥ -1 .5	SS	4.6	(1.6, 12.7)
Δ EFW.SDS ≥ 1.5	Total score	6.1	(2.0, 18.8)
b-coef. EFW ≥ -8.0 SD	Ponderal index	6.1	(1.8, 20.5)
b-coef. EFW ≥ -8.0 SD	MAC / HC	4.3	(1.1, 17.2)
b-coef. EFW ≥ -8.0 SD	SS	1.8	(0.7, 4.5)
b-coef. EFW ≥ -8.0 SD	Total score	2.9	(1.1, 7.8)
FGV.EFW.SDS ≥ -1.0	Ponderal index	1.5	(0.4, 5.6)
FGV.EFW.SDS ≥ -1.0	MAC / HC	1.8	(0.4, 7.5)
FGV.EFW.SDS ≥ -1.0	SS	1.1	(0.4, 3.4)
FGV.EFW.SDS ≥ -1.0	Total score	1.2	(0.4, 3.9)

Abbreviation; SS, subscapular skinfold.

Table 7.6 Serial ultrasound: Sensitivities, specificities and Cohen's kappa indices in the prediction of neonatal morphometry using standard antenatal cut-off criteria.

Ultrasound criteria	Outcome	Sensitivity	Specificity	Kappa
Δ AC.SDS ≥ -2.0	Ponderal index	61.5	82.9	0.35
	MAC / HC	70.0	82.2	0.35
	SS	48.0	85.9	0.36
	Total score	54.5	86.6	0.42
b-coef. AC ≥ -2.0 SD	Ponderal index	100.0	36.1	0.14
	MAC / HC	75.0	31.8	0.02
	SS	84.0	36.1	0.13
	Total score	82.6	35.1	0.11
FGV.AC.SDS ≥ -2.0	Ponderal index	8.3	92.1	0.01
	MAC / HC	10.0	92.3	0.03
	SS	11.1	92.9	0.18
	Total score	18.7	94.9	0.18
Δ EFW.SDS ≥ -2.0	Ponderal index	69.2	86.8	0.47
	MAC / HC	70.0	84.8	0.39
	SS	52.0	90.6	0.46
	Total score	59.1	91.0	0.52
b-coef. EFW ≥ -2.0 SD	Ponderal index	92.9	33.7	0.10
	MAC / HC	83.3	31.8	0.05
	SS	80.0	33.3	0.08
	Total score	78.3	32.4	0.06
FGV.EFW.SDS	Ponderal index	25.0	88.9	0.15
	MAC / HC	10.0	86.1	0.00
	SS	11.1	86.0	0.00
	Total score	12.5	86.4	0.00

Abbreviation; SS, subscapular skinfold.

Table 7.7 Serial ultrasound: Odds ratios in the prediction of neonatal morphometry using standard antenatal cut-off criteria.

Ultrasound parameter	Outcome	Odds ratio	(95% CI)
Δ AC.SDS ≥ -2.0	Ponderal index	7.7	(2.2, 27.9)
	MAC / HC	10.8	(2.5, 47.9)
	SS	5.6	(1.9, 16.4)
	Total score	7.7	(2.6, 23.4)
b-coef. AC ≥ -2.0 SD	Ponderal index	-	-
	MAC / HC	1.4	(0.3, 5.6)
	SS	2.9	(0.9, 9.7)
	Total score	2.6	(0.8, 8.5)
FGV.AC.SDS ≥ -2.0	Ponderal index	1.1	(0.1, 10.1)
	MAC / HC	1.3	(0.1, 13.0)
	SS	1.7	(0.3, 10.1)
	Total score	4.3	(0.8, 26.2)
Δ EFW.SDS ≥ -2.0	Ponderal index	14.8	(3.8, 58.2)
	MAC / HC	13.0	(2.9, 58.4)
	SS	10.5	(3.3, 33.5)
	Total score	14.7	(4.4, 49.1)
b-coef. EFW ≥ -2.0 SD	Ponderal index	6.6	(0.8, 54.3)
	MAC / HC	2.3	(0.5, 11.5)
	SS	2.0	(0.7, 6.3)
	Total score	1.7	(0.6, 5.3)
FGV.EFW.SDS ≥ -2.0	Ponderal index	2.7	(0.6, 12.4)
	MAC / HC	0.7	(0.1, 6.3)
	SS	0.8	(0.1, 4.1)
	Total score	0.9	(0.2, 4.9)

Abbreviation; SS, subscapular skinfold.

Figure 7.1 Distribution of gestational ages at first ultrasound assessment

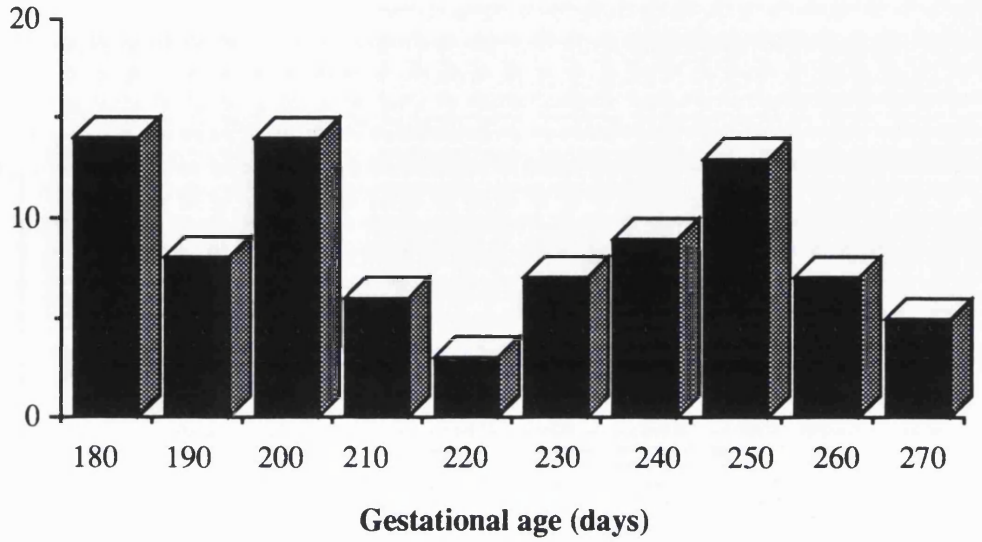


Figure 7.2 Distribution of intervals between first and last ultrasound assessments

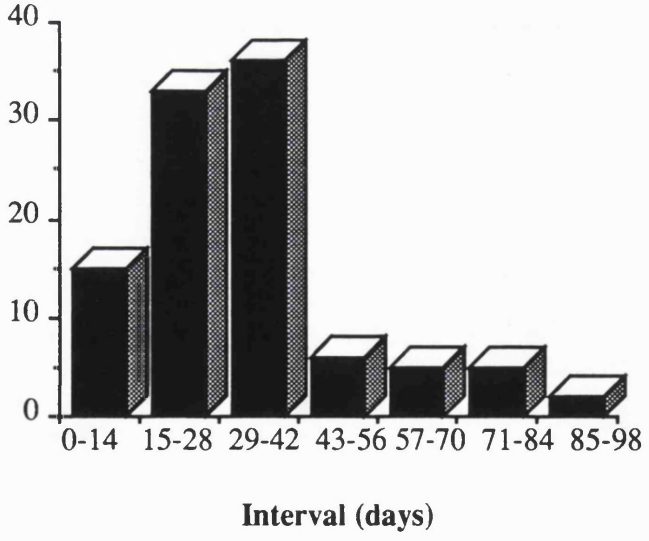


Figure 7.3 Distribution of Birthweights

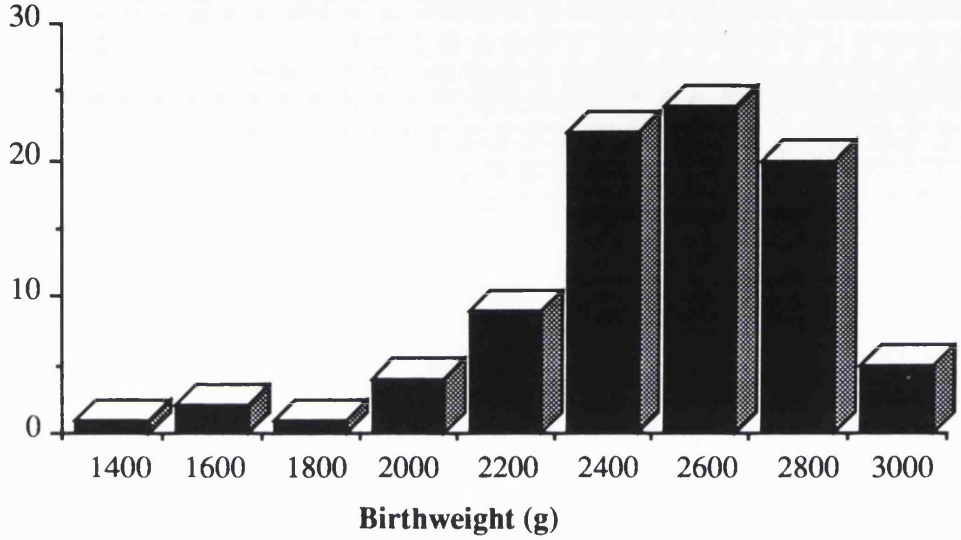


Figure 7.4 Distribution of gestational ages at delivery

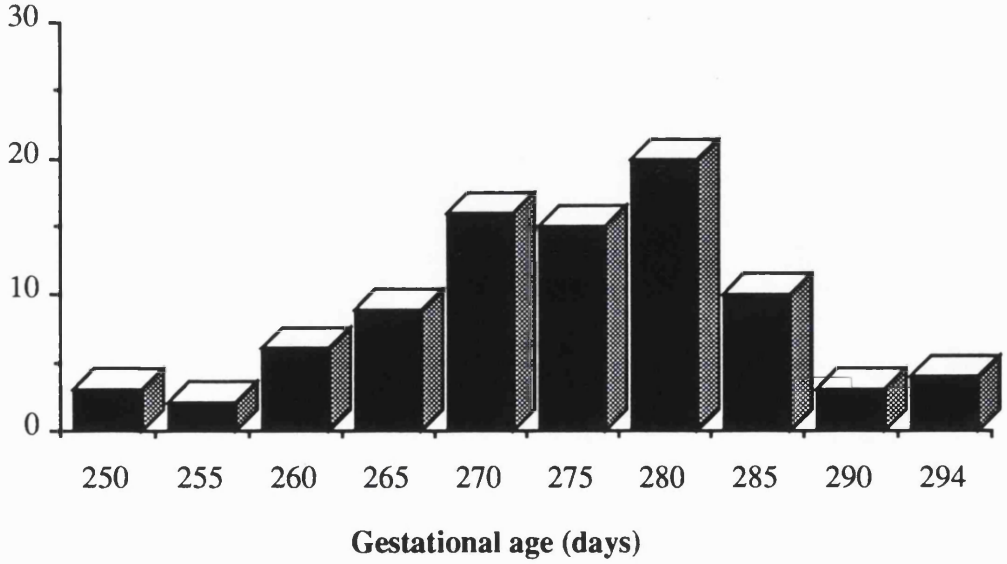
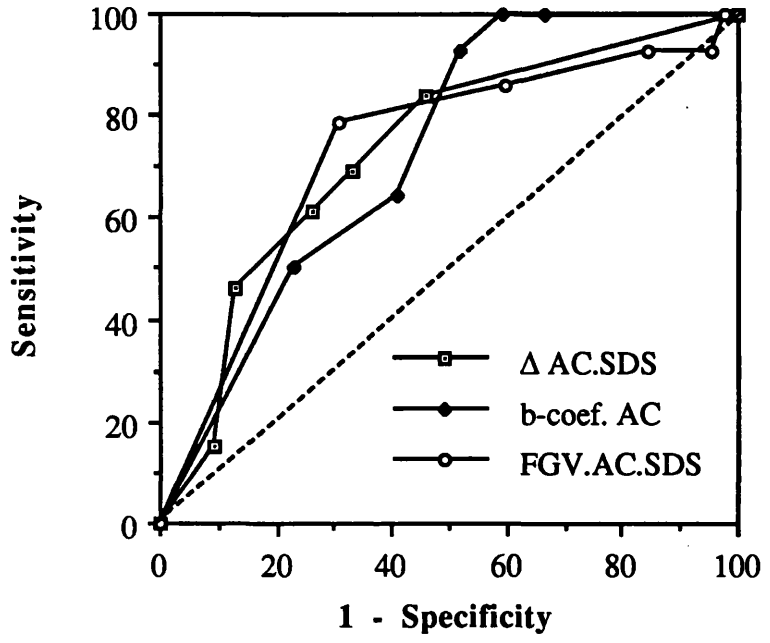


Figure 7.5 Receiver operating characteristic curves in the prediction of ponderal index



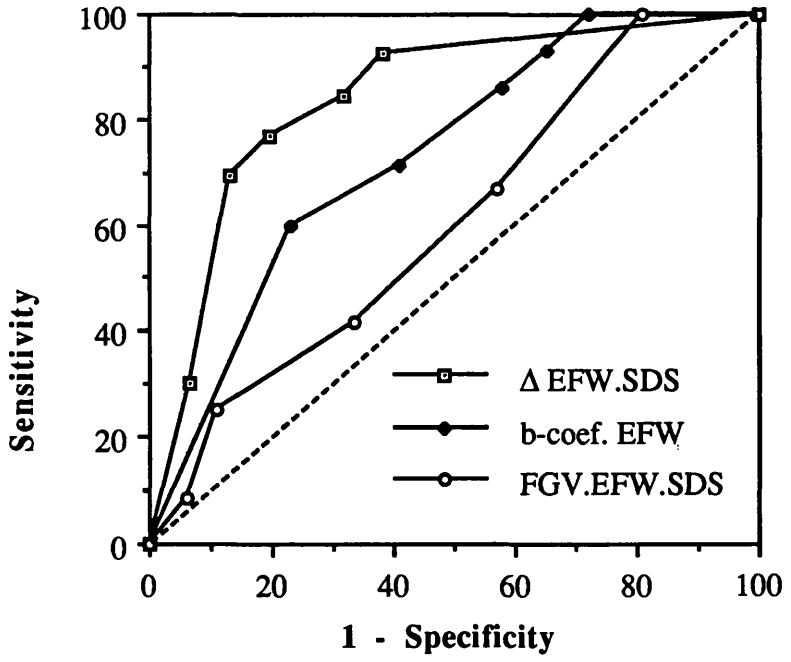
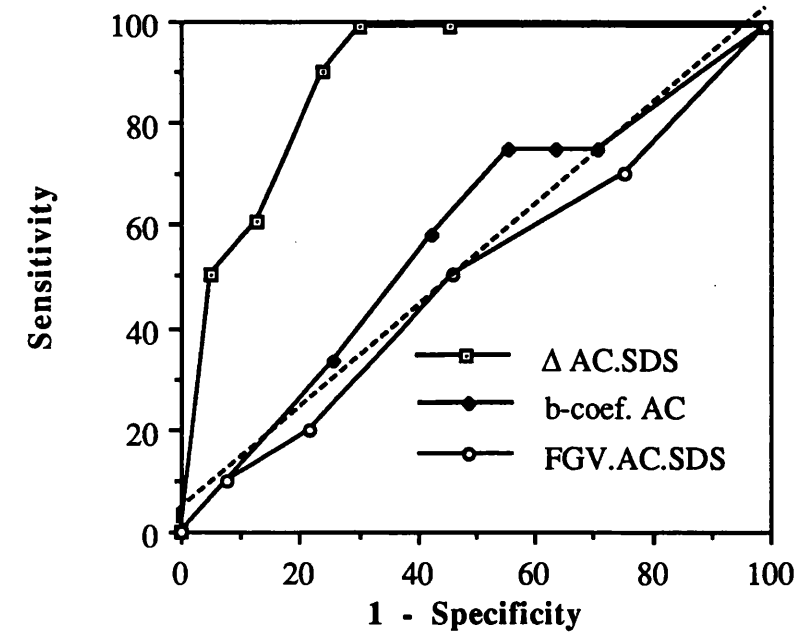


Figure 7.6 Receiver operating characteristic curves in the prediction of MAC / HC ratio



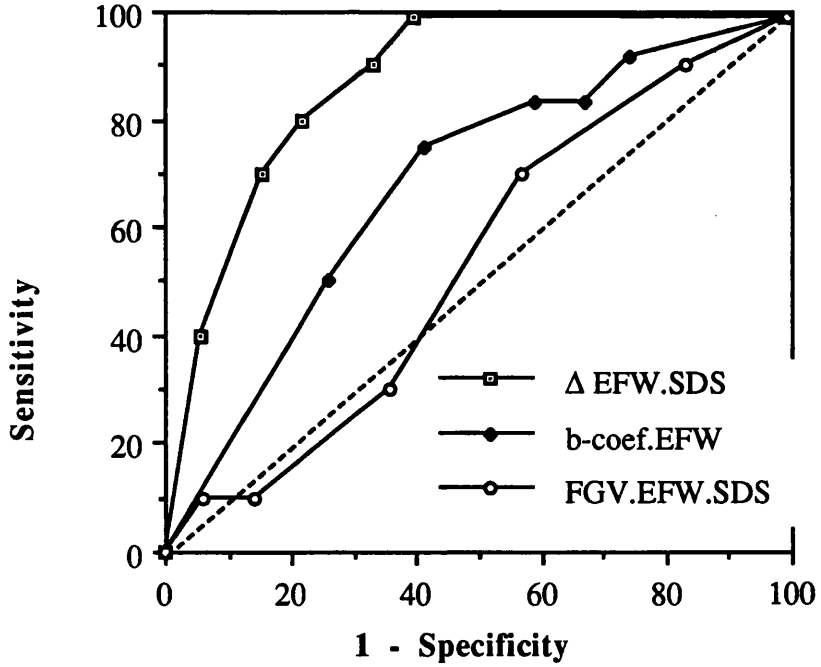
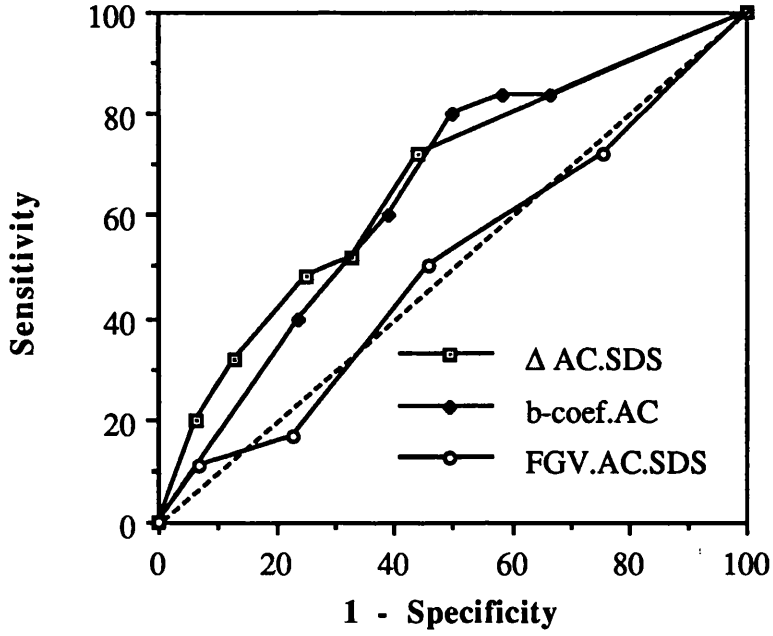


Figure 7.7 Receiver operating characteristic curves in the prediction of subscapular skinfold thickness



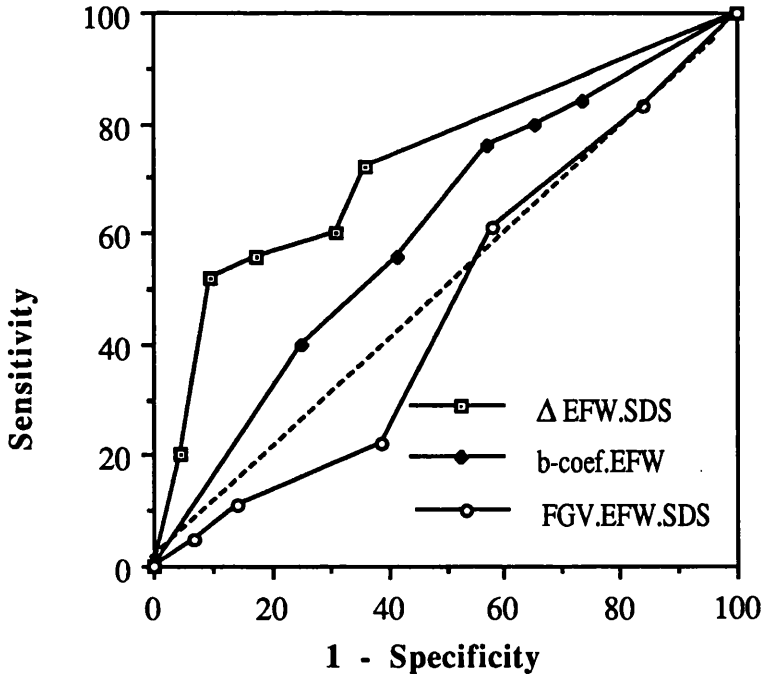
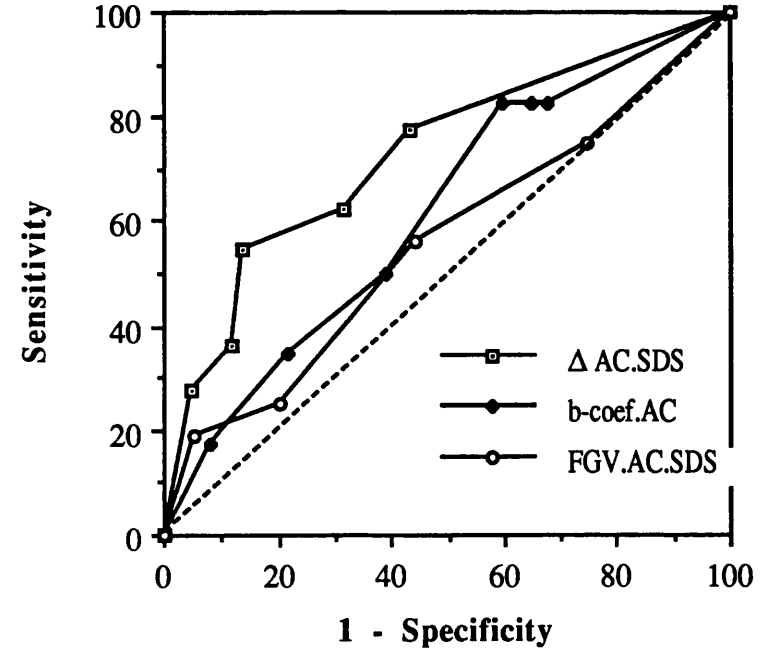


Figure 7.8 Receiver operating characteristic curves in the prediction of total neonatal morphometric score



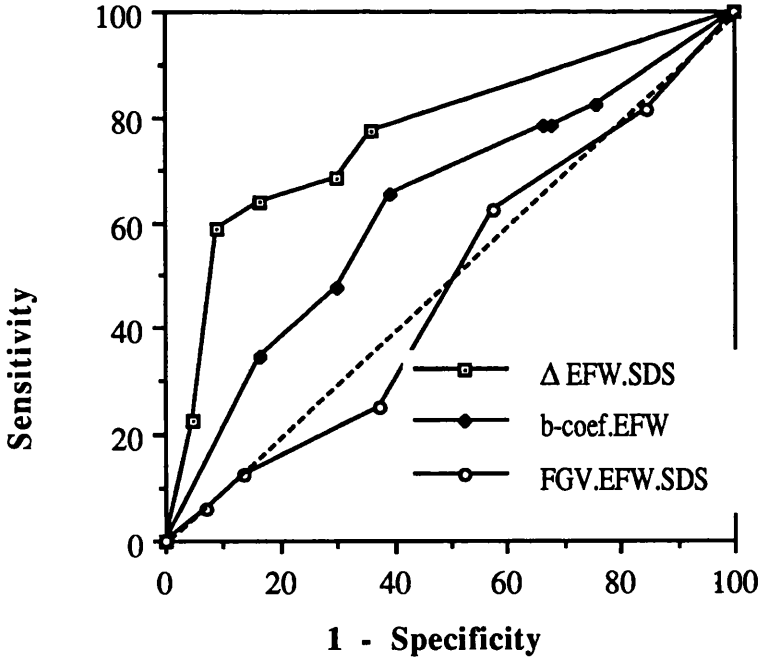
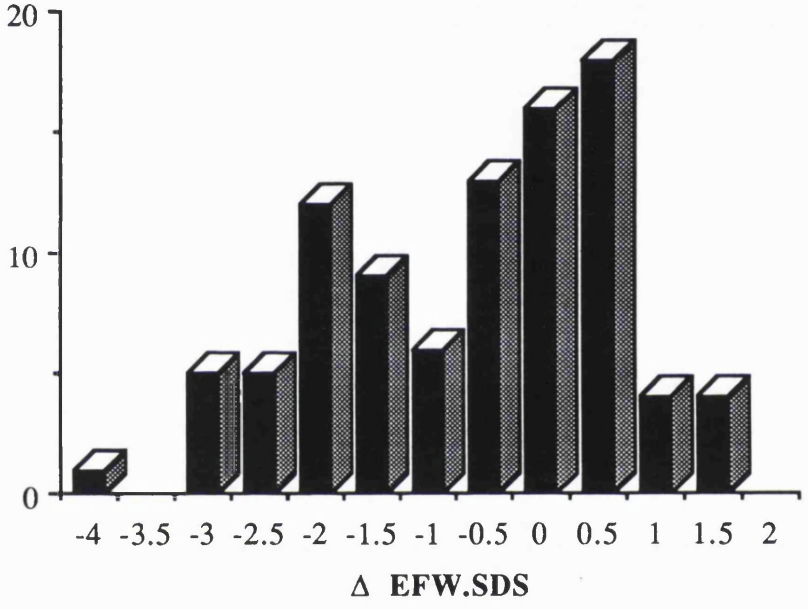


Figure 7.9 Distribution of Δ EFW.SDS



7.4.2. Comparison with other ultrasound parameters

Having determined Δ EFW.SDS to be the optimal method of quantifying ~~serial~~ serial ultrasound data, this ultrasound measure was then compared with other standard ultrasound parameters currently used to evaluate fetal well-being, namely umbilical artery PI and last EFW.SDS, a single estimate of fetal size. Correlations between these ultrasound parameters and neonatal morphometric indices were determined using linear regression analysis. Although Δ EFW.SDS and last EFW.SDS both demonstrated significant correlation with each of the four neonatal indices, umbilical artery PI showed a significant correlation with subscapular and triceps skinfolds, but not with ponderal index or MAC / HC ratio (Table 7.8).

ROC curves were derived for these ultrasound parameters in the prediction of abnormal neonatal morphometry (total score ≥ 2) (Figure 7.10) and adverse perinatal outcome (Figure 7.11). Comparisons of the areas under the ROC curves showed that Δ EFW.SDS was superior to umbilical artery PI and last EFW.SDS in the prediction of both abnormal neonatal morphometry and adverse perinatal outcome (Table 7.9)

These ultrasound parameters were then compared using both the optimal ultrasound cut-off derived from the ROC curve and a standard ultrasound cut-off of ≥ 2 SD. In the prediction of abnormal neonatal morphometry, the optimal cut-off for Δ EFW.SDS (≥ -1.5) resulted in a sensitivity (63.6%), specificity (83.6%), Kappa (0.33) and odds ratio (6.1), higher than comparable data for any other ultrasound parameter (Table 7.10). A similar pattern was evident for the prediction of significant perinatal morbidity, irrespective of whether the optimal or standard cut-off criteria were used (Table 7.10).

These results suggest that serial values of EFW (as quantified by Δ SDS) were superior to umbilical artery PI and the last EFW in the prediction of neonatal morphometry and adverse perinatal outcome.

Table 7.8 Comparison with other ultrasound parameters: Linear regression analysis of relationship with neonatal morphometry.

Ultrasound measure (y)	Morphometric index (x)	Equation	R-value	p
Δ EFW.SDS	Ponderal index	$y = -4.71 + 1.67x$	0.37	0.000
	MAC / HC	$y = -7.14 + 23.5x$	0.40	0.000
	Subscapular	$y = -3.27 + 0.882x$	0.38	0.000
	Triceps	$y = -2.73 + 23.5x$	0.30	0.002
Umbilical artery PI	Ponderal index	-	-	NS
	MAC / HC	-	-	NS
	Subscapular	$y = 3.08 - 0.59x$	0.27	0.008
	Triceps	$y = 2.71 - 0.466x$	0.21	0.032
Last EFW.SDS	Ponderal index	$y = -4.03 + 0.685x$	0.18	0.04
	MAC / HC	$y = -5.17 + 10.2x$	0.23	0.013
	Subscapular	$y = -3.9 + 0.518x$	0.29	0.002
	Triceps	$y = -3.59 + 0.417x$	0.23	0.014

Table 7.9 Comparison with other ultrasound parameters: area under the ROC curves in the prediction of abnormal neonatal morphometry and adverse perinatal outcome.

Ultrasound criteria	Outcome criteria	AUC	SE
Δ EFW.SDS	**Total score ≥ 2**	0.66	0.02
Umbilical artery PI		0.43	0.02 *
Last EFW.SDS		0.58	0.02 *
Δ EFW.SDS	**Significant perinatal morbidity**	0.62	0.03
Umbilical artery PI		0.48	0.02 *
Last EFW.SDS		0.55	0.02

* $p < 0.05$ vs. Δ EFW.SDS

Table 7.10 Comparison with other ultrasound parameters: Prediction of abnormal neonatal morphometry and adverse perinatal outcome

Ultrasound parameter	Outcome	Se	Sp	Kappa	OR	(95% CI)
Optimal antenatal cut-off						
Umbilical artery PI.SDS > 1.5	Total score ≥ 2	61.9	63.8	0.22	2.9	(1.1, 8.3)
Last EFW.SDS ≥ -2.5		59.1	70.7	0.25	3.5	(1.3, 9.4)
Δ EFW.SDS ≥ -1.5		63.6	83.6	0.33	6.1	(2.0, 18.8)
Standard antenatal cut-off (2 SD)						
Umbilical artery PI	Total score ≥ 2	42.9	79.0	0.21	2.8	(1.0, 8.2)
Last EFW.SDS		81.8	45.3	0.17	3.7	(1.1, 12.2)
Δ EFW.SDS		59.1	91.0	0.52	14.7	(4.4, 49.1)
Optimal antenatal cut-off						
Umbilical artery PI.SDS ≥ 1.5	Adverse perinatal outcome	48.4	58.9	0.07	1.3	(0.5, 3.3)
Last EFW.SDS ≥ -2.0		62.5	45.8	0.06	1.4	(0.5, 3.8)
Δ EFW.SDS ≥ -1.5		53.5	80.4	0.3	3.6	(1.3, 9.5)
Standard antenatal cut-off (2 SD)						
Umbilical artery PI	Adverse perinatal outcome	29.0	75.0	0.04	1.2	(0.5, 3.3)
Last EFW.SDS		62.5	45.8	0.06	1.4	(0.5, 3.8)
Δ EFW.SDS		39.3	84.3	0.25	3.5	(1.2, 10.2)

Figure 7.10 Comparison with other ultrasound parameters: Receiver operating characteristic curves in the prediction of total neonatal morphometric score

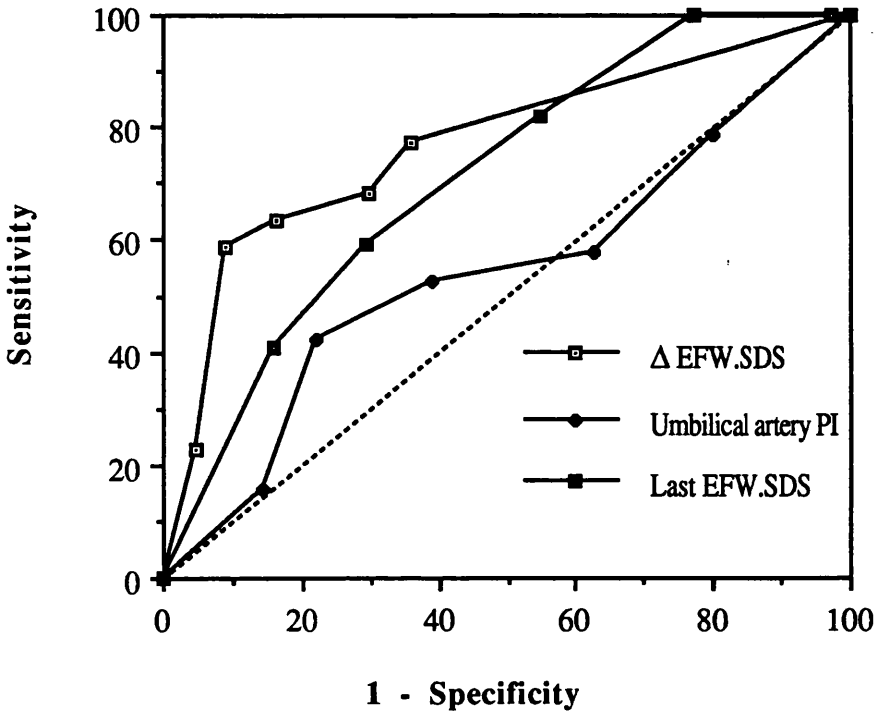
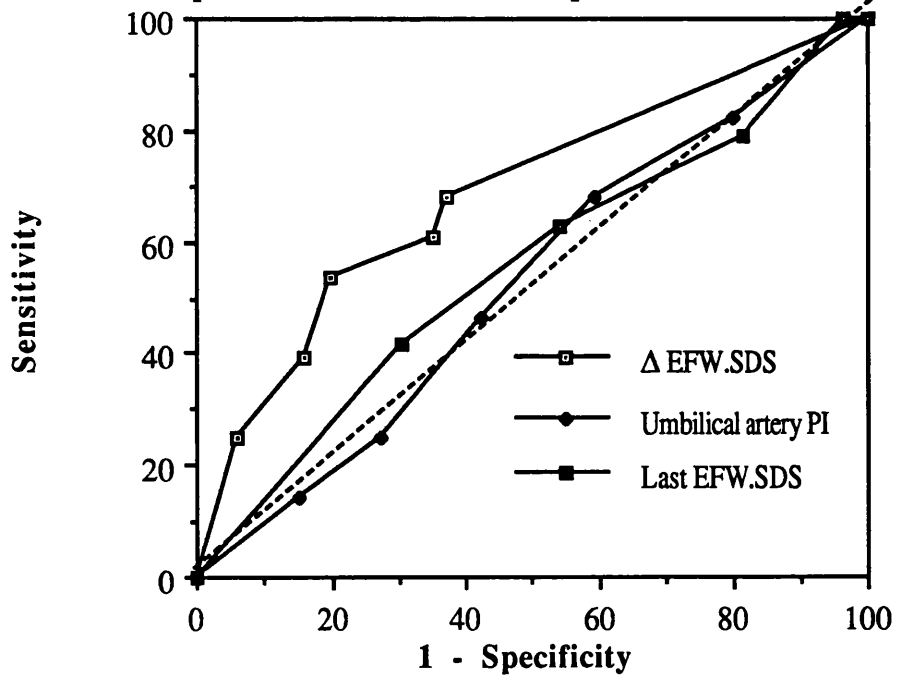


Figure 7.11 Comparison with other ultrasound parameters: Receiver operating characeteristic curves in the prediction of adverse perinatal outcome



7.4.3 Morbidity and biochemistry

Having determined Δ EFW.SDS ≥ -1.5 SD to be the optimal cut-off for the prediction of abnormal neonatal morphometry, this antenatal criteria was used to separate the fetuses into two groups, those with ultrasound evidence of normal growth and those with evidence of IUGR. Twenty nine (27.9%) of the 104 fetuses were categorized as growth retarded by this criteria. The median (range) intervals between first and last scans in fetuses with ultrasound evidence of normal and subnormal growth were 36 (14 - 63) and 59 (14 - 77) days respectively; these intervals were not significantly different between the two groups.

Perinatal morbidity data and cord biochemical indices of IUGR in the two groups are summarized in Table 7.11. Median birthweight was lower in those with ultrasound evidence of IUGR compared with those who had normal ultrasonic growth; however, this difference was not statistically significant. No significant differences were noted in the arterial or venous pH and BE, or in the incidence of Caesarean section for fetal distress in labour or Apgar score < 7 at 5 minutes between the two groups. There was no significant relationship between Δ EFW.SDS and umbilical arterial or venous pH (n = 95). Neonates were more likely to be admitted to NICU with complications related to IUGR if they had ultrasonic evidence of IUGR ($p < 0.05$).

Sixty seven umbilical venous samples were collected for the assessment of the metabolic and endocrine status of the newborn. Cord samples were obtained in 48 and 19 fetuses classified antenatally to have IUGR and non-IUGR respectively (48 / 75 vs. 19 / 29, $\chi^2 = 0.058$, NS). There were no significant differences in the cord venous glucose, insulin or triglyceride levels in the two groups. However, levels of IGF-1 were significantly lower in the group with IUGR [median 0.05 (range 0.00, 0.24) U / ml] compared with corresponding levels in the group with normal growth [median 0.13, (range 0.00, 0.94) U / ml].

Table 7.11 Perinatal and biochemical outcome in neonates classified by ultrasound criteria.

Perinatal outcome	**Non-IUGR (n = 75)**	**IUGR (n = 29)**	
Birthweight	2530 (1900 - 3200)	2390 (1360 - 2760)	NS
Gestational age at delivery	273 (252 - 292)	274 (252 - 290)	NS
Females	40 (53.3%)	18 (62.1%)	NS
Apgar at 5 minutes ≤ 7	2 (2.7%)	3 (10.3%)	NS
Caesarean section for fetal distress	7 (9.3%)	6 (20.7%)	NS
Umbilical arterial pH	7.32 (7.20 - 7.39)	7.26 (7.15 - 7.38)	NS
Umbilical arterial base excess (mmol / l)	4.5 (0.6 - 10.7)	4.9 (0.4 - 11.8)	NS
Umbilical venous pH	7.29 (7.07 - 7.48)	7.25 (7.08 - 7.40)	NS
Umbilical venous base excess (mmol / l)	4.5 (0.6 - 10.7)	4.9 (0.4 - 11.8)	NS
Admission to NICU	3 (4.6%)	5 (17.2%)	$p < 0.05$
Biochemical data (umbilical venous samples)	**Non-IUGR (n = 49)**	**IUGR (n = 19)**	
Glucose (mmol / l)	2.9 (1.6 - 5.0)	2.5 (1.2 - 4.8)	NS
Insulin (pmol / l)	27.3 (2.4 - 138.0)	31.2 (5.9 - 272.0)	NS
IGF-1 (U / ml)	0.13 (0.00 - 0.94)	0.05 (0.00 - 0.24)	$p < 0.05$
Triglycerides (mmol / l)	0.47(0.12 - 1.43)	0.44 (0.22 - 1.63)	NS

Results are median (range) or n (%).

7.5 Discussion

The serial ultrasound assessment of AC and EFW has not previously been validated against "gold standards" of IUGR. The study shows that Δ SDS.EFW (and to a lesser extent Δ AC.SDS) in the third trimester predicted subsequent neonatal morphometry accurately. This ultrasound measure was also superior to other standard ultrasonic parameters, such as umbilical artery PI and single estimates of fetal size, currently used to evaluate fetal well-being. The use of Δ EFW.SDS to divide small fetuses into those with or without IUGR resulted in two groups of neonates with some differences in perinatal morbidity and cord biochemistry suggestive of IUGR.

This study was restricted to third trimester fetuses with an AC < 10th centile (Deter et al. 1982a) at entry into the study. Whilst IUGR is by no means restricted to small fetuses, it was felt that the antenatal separation of small fetuses into those with normal or subnormal growth was important in the subsequent clinical management of such fetuses. The intrauterine growth of AGA fetuses can also be assessed by serial data to detect those with subnormal growth; such an analysis was not within the realms of this study. As the collection of ultrasound data was prospective, the 10 fetuses (9.6%) who subsequently had a birthweight > 10th centile were not excluded. The entry criteria with regards to gestational age assessment was necessarily strict as the interpretation of subsequent ultrasound measurements depended on knowledge of precise gestational ages. Data from fetuses who delivered before 36 weeks gestation were excluded (n = 4). This was to remove the confounding effects of prematurity on perinatal morbidity. However, it is acknowledged that this criteria resulted in the exclusion of fetuses with severe IUGR secondary to pre-eclampsia who were delivered early (Lin et al. 1991). The consequences of restricting this study to fetuses who delivered at term were partly reflected in the relatively low morbidity (only 8 of the 104 neonates were admitted to NICU with problems related to IUGR) and the absence of perinatal mortality in the study group. These small fetuses were therefore very different from those investigated by other workers; for example the cordocentesis data reported by Economides et al. (1989a) were based on fetuses with an AC < 2.5th centile, the majority of whom were delivered prematurely due to obstetric complications.

Neonatal morphometry was used as the main outcome measure of IUGR against which these ultrasound measures were to be evaluated. Numerous studies have assessed these morphometric indices (ponderal index, MAC / HC ratio, subscapular and triceps skinfold thickness) in the prediction of perinatal morbidity, but no one measure has been shown to be consistently superior to the others (Roord et al. 1978, Excler et al. 1985, Georgieff et al. 1988,Wolfe et al. 1990). The results in the present study were therefore reported using all these morphometric measures together with a summary total

morphometric score. Previous studies have reported the incidence of ponderal index < 10th centile in a SGA population to be 17.5% (432 / 2462) (Villar et al. 1990), 30.1 % (84 / 279) (Fay et al. 1991a), 38.6% (46 / 119) (Walther and Raemaker 1982) and 55.0% (11 / 20) (Patterson and Pouliot 1987a). The lower incidence in the present study of 14.3% (15 / 104) may be related to the use of a cut-off of ponderal index < 2 SD.

An analysis to determine the optimal method of quantifying serial values of AC and EFW was first undertaken. The non-linearity of fetal growth meant that changes in absolute measurements of AC or EFW could not be used to describe growth, unless standards for normal growth were also available for each interval between scan. The three statistical measures chosen to quantify serial data circumvented this problem. By expressing measurements of AC and EFW as SDS, changes in SDS (Δ SDS) could be used to define the growth trajectory independent of gestational age. The SDS at the first and last third trimester scans were used to calculate Δ SDS so that the growth trajectory could be measured over the largest possible scan interval. The disadvantage with such a method of quantifying serial data was that the SDS for all other measurements between the the first and last scan were not taken into consideration, thus placing great dependence on the accuracy of these two scans. The use of the quadratic *b*-coefficient theoretically overcame this disadvantage as the growth curve for each fetus was derived using all available scan data. The third method of quantifying serial data involved the calculation of the SDS of fetal growth velocity so that increments in absolute measurements per week could be evaluated against normal standards for the same interval. The use of the latter was crucial as increments in AC or EFW per week decreased with increasing gestational age (Table 5.7). As 88 (85%) fetuses had more than one FGV.SDS, the worst SDS (ie. representing the least growth) for each fetus was used to quantify growth.

Receiver operating characteristic curves were used to evaluate the diagnostic performance of each ultrasound measure (Figures 7.5 to 7.8). The area under the ROC curve reflects the overall performance of a diagnostic test (Hanley and MacNeil 1982). Furthermore, ROC curves allow standard statistics of sensitivity, specificity and odds ratio to be calculated using the optimal cut-off value for each test. Thus, different tests can be compared using their respective optimal cut-off values. This latter method of analysis has previously been used to identify the optimal cut-off point for ultrasound parameters in the prediction of a specific outcome (Miller and Gabert 1992). The results of the areas under the ROC curves clearly demonstrated that Δ SDS.EFW and Δ SDS.AC were the best predictors of subsequent neonatal morphometry. Furthermore, the highest ORs and Kappa values were achieved with these measures, irrespective of whether the optimal or standard ultrasound cut-off was used. However, the sensitivities achieved using Δ AC.SDS and Δ EFW.SDS were rather disappointing, despite the high specificities. The use of the optimal cut-offs for Δ AC.SDS and Δ SDS.EFW resulted in sensitivities ranging from 59 to 72%,

and 63 to 76% respectively.

By contrast, the other ultrasound measures, *b* -coefficient and FGV.SDS, were poor predictors of subsequent neonatal morphometry. This was particularly evident when evaluating the odds ratios and Cohen's Kappa, the overwhelming majority of which were not significantly different from zero and < 0.20 respectively. This was irrespective of whether the optimal or standardised antenatal criteria were used. The poor predictiveness of the *b* -coefficient may be due to the small SD values for *b* -coef.AC and *b* -coef.EFW as derived from the 67 reference fetuses. Values of the b -coefficient were about 8 SD below the respective means in order to achieve a sensible trade-off between sensitivity and specificity (Table 7.4). The poor results reported for FGV.SDS may be due to the use of a shorter interval between scans of two weeks compared with that used for calculating Δ SDS (median 4 weeks, range 2-13 weeks). Furthermore, the availability of more than one value of FGV.SDS for each fetus in 85% of cases led to the arbitrary use of the lowest value of FGV.SDS (ie. representing least growth) to define abnormality.

Several studies have previously investigated the ability of ultrasound to predict an abnormal neonatal ponderal index or MAC / HC ratio (Tables 2.7 and 2.8). The sensitivities and ORs reported here for the last EFW.SDS, a single estimate of fetal size, are comparable to those reported previously (Sarmandal and Grant 1990, Weiner and Robinson 1989). In the present study, the sensitivity and OR for Δ EFW.SDS were higher than those obtained using a single EFW (Tables 7.9 and 7.10). These results confirm the superiority of serial over single estimates of fetal size in the diagnosis of IUGR. This is not surprising as growth is a dynamic process involving change in fetal size (Altman and Hytten 1989), and is best assessed by serial measurements of fetal anthropometry (Deter et al. 1990, Deter and Harrist 1992). This measure of serial EFW was also superior to umbilical artery PI in the prediction of an abnormal total morphometric score (Tables 7.9 and 7.10). Two studies have previously reported data on the use of umbilical artery Doppler waveform indices to predict abnormal neonatal morphometry in low risk subjects (Sijmons et al. 1989, Beattie and Dornan 1989). The sensitivity obtained using umbilical artery PI in this study (61.9%) was higher than those previously reported by Sijmons et al. (1989) [≤ 27.3%] and Beattie and Dornan (1989) [≤ 28.0%]. These differences may be related to the different prevalences of IUGR in the different studies. Odds ratios could not be calculated for the study of Beattie and Dornan (1989). The OR for umbilical artery PI in the present study was comparable to that reported by Sijmons et al. (1989) (see Table 2.9), re-emphasising the advantage of using the OR in view of its prevalence-independence (Kahn and Sempos 1989).

The best measure of serial data, Δ EFW.SDS, was also compared with umbilical artery PI and estimates of fetal size in the prediction of adverse perinatal

outcome. Adverse perinatal outcome was defined using specific outcome criteria chosen to reflect morbidity associated with IUGR (Section 7.3). By excluding fetuses who delivered before 36 weeks gestation, the compounding effects of prematurity were excluded. Assisted vaginal delivery was not included as the majority of such babies show no objective evidence of compromise after birth (Sykes et al. 1983). Acidaemia was defined as a value < 10th centile rather than the arbitrary definition of an umbilical artery pH < 7.20 used in many previous studies. Neonates with hypoglycaemia were only included if hypoglycaemia was severe enough to warrant intravenous glucose treatment, thereby excluding otherwise asymptomatic infants with mild hypoglycaemia who respond to early feeding (Hawdon et al. 1992a). Finally, infants with cerebral dysfunction suggestive of ischaemic injury (Hull and Dodd 1991) were included in the outcome criteria. This study showed that Δ SDS.EFW was superior to umbilical artery PI in the prediction of adverse perinatal outcome. No previous study has addressed the ability of serial ultrasound measurements to predict adverse perinatal outcome. Five other studies (Laurin et al. 1987, Berkowitz et al. 1988b, Chambers et al. 1989, Burke et al. 1990, Gudmundsson and Marsal 1991b) have specifically addressed the ability of umbilical artery Doppler waveform indices to predict perinatal morbidity in small fetuses (Table 2.12). The outcome criteria reported varied from study from study. The sensitivities and ORs obtained using umbilical artery PI in this study were similar to those reported by Berkowitz et al. (1988b). Comparisons with other studies were difficult in view of the different outcome criteria used. The incidence of adverse perinatal outcome varied from study to study. The incidence of Caesarean section for fetal distress in small fetuses had previously been reported to be between 7.2% and 14.5% (Chambers et al. 1989, Burke et al. 1990), compared with 12.5% in the present study. Burke et al. (1990) reported a rate of admission to NICU of 29.6% compared with 7.7% in the present study. However, Burke and co-workers (1990) included data from premature fetuses as well as those with congenital anomalies. The poor results obtained using umbilical artery PI in the present study may be due to the low incidence of morbidity in a group of fetuses who delivered beyond 36 weeks gestation.

The optimal cut-off for Δ SDS.EFW (≥ -1.5 SD) was used to separate the study group into fetuses with IUGR and those with normal growth. The resultant two groups did not differ in the proportion who were delivered by Caesarean section for fetal distress, or who had a low Apgar score (≤ 7) at 5 minutes (Table 7.11). The lack of differences in these measures of asphyxia between the two groups may be readily explained by the poor specificity of these outcome measures to IUGR. Umbilical arterial and venous pH and BE were obtained from cord samples at 95 of the 104 deliveries; no significant differences in pH or BE were found between these two groups of fetuses. It is known from both data obtained at cordocentesis (Soothill et al. 1987a, Nicolaides et al. 1989) and at delivery (Lin et al. 1980, Dijxhoorn et al. 1987, Steer 1989) that SGA per se confers an increased

risk of acidaemia. The incidence of acidaemia of 24% in the present study group was 24%. Even if small fetuses with ultrasound evidence of IUGR were more likely to be more acidaemic than those with normal growth, a background incidence of acidaemia of 24% would have necessitated a larger number of fetuses to be studied to demonstrate a significant difference between these two groups. The degree of hypoglycaemia in the group as a whole was not severe. Only two neonates, one in each group, actually required intravenous glucose infusion. Jones and Robertson (1986), who reported an equally low incidence of SGA infants requiring intravenous glucose treatment (1 in 165), have questioned the usefulness of hypoglycaemia as a measure of serious morbidity. However, Lucas et al. (1988) have reported data to show that moderate hypoglycaemia may nevertheless have serious neurodevelopmental consequences.

Significantly greater proportion of fetuses with ultrasound evidence of IUGR were admitted to NICU. One neonate in the IUGR group had clinical symptoms and signs suggestive of necrotising enterocolitis. However, this was not confirmed on radiological examination and the symptoms resolved on expectant management. The complications in this group included 4 infants with neurological disturbances and 1 with hypoglycaemia requiring intravenous glucose. Given the absence of serious morbidity associated with IUGR such as necrotising enterocolitis and pulmonary haemorrhage in the study group, neurological deficits may be the most important sequelae in such a cohort of fetuses born at term. Most studies on the neurodevelopmental assessment of such infants are valid only if carried out until 4 years of age. However, recent studies have shown that early objective measures of brain structure and function in the neonate can predict longer term outcome in considerable detail (Stewart 1989). Cranial ultrasound findings and neurological examination in the first few days of life predict subsequent neurodevelopmental outcome with an accuracy of 98% (Stewart et al. 1988). Therefore, it is all the more significant that differences were noted between the two groups despite the small numbers with evidence of cerebral dysfunction.

No significant differences in umbilical venous glucose, insulin or triglyceride levels were found between the two groups at delivery. It is possible that as biochemical indices at delivery were obtained in only about two-thirds of all fetuses in the study group, the resultant group in whom biochemical indices were available differed from the original study group. However, this is unlikely to be the case as χ^2 analysis showed no significant difference in the proportions of fetuses with umbilical cord blood data in the two groups. Previous studies have shown that small fetuses per se are more likely to be hypoglycaemic (Soothill et al. 1987a, Economides et al. 1989b, Hawdon et al. 1992), hypoinsulinaemic (Economides et al. 1989b), and hypertriglyceridaemic (Economides et al. 1988) than normal sized fetuses. Hawdon et al. (1992a) are the only workers to have reported data on the further sub-division of small fetuses by ultrasound with respect to umbilical venous

glucose levels at delivery (Section 2.3.4.1.). Interestingly, they also found no significant differences in glucose levels in small fetuses with absent end-diastolic flow in the umbilical artery compared with those where end-diastolic flow was present. The present study group (n = 67) was larger than that in the study of Hawdon et al. (1992) [n = 25]. No study has investigated the sub-division of small fetuses by ultrasound with respect to insulin or triglyceride levels at delivery. It is evident from Table 7.11 that there was considerable overlap in the values of these indices in the two groups. As no appropriately derived reference ranges exist for insulin or triglyceride levels at delivery (Section 2.2.4.4), it was not possible to determine whether values for these biochemical indices were low in both groups with respect to normal reference ranges. Significant differences in serum IGF-1 levels were noted between the two groups. Previous studies on umbilical cord data at delivery have shown IGF-1 levels to be lower in SGA neonates than AGA neonates (Foley et al. 1980, Bennett et al. 1983, Gluckman et al. 1983). These findings have also been confirmed in a study based on data obtained at cordocentesis (Lassarre et al. 1991). This is the first study to show that within a group of small fetuses, those with ultrasound evidence of IUGR had significantly reduced IGF-1 levels in umbilical venous blood at delivery compared with normally growing fetuses.

7.6 Summary

In a group of small fetuses, serial measurements of AC and EFW, as quantified by a change in standard deviation scores (Δ SDS) in the third trimester, predicted subsequent neonatal morphometry accurately. This ultrasound measure was also superior to other standard ultrasonic parameters, such as umbilical artery PI and single estimates of fetal size, in the prediction of abnormal neonatal morphometry and adverse perinatal outcome. The use of Δ EFW.SDS to divide small fetuses into those with or without IUGR resulted in two groups of neonates with some differences in perinatal morbidity and cord biochemistry indicative of IUGR. Serial ultrasound measurements of anthropometric size are therefore useful in separating small fetuses into those with antenatal evidence of growth failure and those with ultrasonic normal growth.

CHAPTER 8

DISCUSSION

8.1 Introduction

In a commentary in the British Journal of Obstetrics and Gynaecology, Altman and Hytten (1989) wrote regarding the antenatal diagnosis of IUGR; "The estimation of fetal growth rather than fetal size is important, because the object of antenatal surveillance is to anticipate or demonstrate clinical problems that can be helped by appropriate action. Establishing that a fetus is small does not establish a clinical problem; only the demonstration of impaired growth can do that. That basic point has been recognised by those studying child growth, but not usually by those studying fetal growth."

Altman and Hytten (1989) highlighted much of the confusion which surrounds the antenatal diagnosis of IUGR. Much of the obstetric literature has confused SGA with IUGR. A single estimate of fetal size in late pregnancy suggestive of a small fetus may lead the clinician to assume that some pathological interference has led to impaired growth; however, size alone does not prove it. The term IUGR should be restricted to those fetuses where there is definite evidence that growth has altered. Ultrasound provides the opportunity to assess change in fetal size with gestation in the individual fetus, but the robustness of ultrasound in diagnosing reduced growth in-utero has not hitherto been properly evaluated. The use of serial ultrasound is especially important in identifying small fetuses with evidence of growth failure, in view of the increased perinatal morbidity in this sub-group of fetuses (Villar et al. 1990). This would enable increased surveillance to be concentrated in these fetuses and reduce surveillance and intervention in those small fetuses with normal growth and who are constitutionally small.

Any ultrasound measure purported to be of use in the diagnosis of IUGR must be validated against a "gold standard". There is little agreement as to the outcome measure of choice for the purposes of confirming IUGR. Whilst some have advocated the use of Doppler and cordocentesis to define IUGR (Campbell 1989), others have used neonatal morphometry to define IUGR (Patterson and Pouliot 1987a, Villar et al. 1990, Fay et al. 1991a). In this study, neonatal morphometry was used as the definitive outcome measure for the evaluation of serial ultrasound data.

The originality of this work can be summarised as follows;
(1) Study 1 is the first to report appropriately derived reference ranges for AC and EFW based on the longitudinal data of a large cohort of fetuses.
(2) Study 2 is the only study which has assessed the reproducibility of EFW using the limits of agreement method of Bland and Altman (1986).
(3) Study 3 is the only study which has evaluated serial AC and EFW data in the diagnosis of IUGR within a group of small fetuses.

8.2 Pathology of IUGR: Rationale for neonatal morphometry indicative of malnutrition

Any ultrasound parameter purported to be of use in the diagnosis of IUGR has to be validated against a "gold standard". As was apparent from Sections 2.2.1 to 2.2.4, low birthweight, neonatal morphometric indices indicative of "wasting", measures of perinatal morbidity and biochemical indices of malnutrition have all been used to define IUGR. In this study, neonatal morphometric indices were used as the" gold standard", the reasons for which have been enumerated in Section 2.2.5.

There were two reasons for choosing a variety of morphometric indices to define IUGR. First, although all reflect soft tissue "wasting" in response to intrauterine starvation (Hill et al. 1984, Chard et al. 1992), skinfold thicknesses are direct measures of subcutaneous fat whilst the ratios quantify soft tissue wasting relative to normal skeletal growth. Implicit in the use of the ponderal index and MAC / HC ratio indices to quantify "wasting" is that in normal fetal growth, all parts of the fetus grow in parallel with no evidence of body disproportion. However, numerous studies have questioned the validity of the latter assumption as head growth has also been shown to decrease in IUGR (Crane and Kopta 1980, Kramer et al. 1989, Colley et al. 1991). Second, no one index has been shown to be consistently superior to the others in the prediction of perinatal morbidity associated with IUGR (Roord et al. 1978, Excler et al. 1985, Georgieff et al. 1988,Wolfe et al. 1990). As the studies in Section 6.2 showed all indices to be reproducible, the results of this study were therefore reported using all the above morphometric indices.

Although previous ultrasound studies used a cut-off of < 10th centile to define IUGR (Ott 1985, Vintzileos et al. 1986, Patterson et al. 1987b, Weiner and Robinson 1989, Sarmandal and Grant 1990), a lower cut-off (of -2 SD for all neonatal morphometric indices) was used in this study.This was due to the observation that the ponderal index and MAC / HC ratio varied with birthweight (Chard et al. 1992). As this study only addressed a population of small fetuses, it was imperative to use a low cut-off as such fetuses were more likely to have lower morphometric indices at birth by virtue of being small per se (Chard et al. 1992).

The results of this study showed that the same ultrasound cut-off (Δ EFW.SDS > -1.5 SD) was the best predictor of all neonatal morphometric indices evaluated. The cumbersome use of so many different outcome criteria led to the derivation of a summary outcome measure, the total neonatal morphometric score. Whilst the same ultrasound cut-off also predicted a low total morphometric score, future studies are needed to develop a more refined score to quantify neonatal wasting in the newborn.

8.3 Pathology of IUGR: rationale for serial assessment of fetal size

8.3.1 Cellular hyperplasia and hypertrophy

In studies on fetal and newborn rats, which subsequently extended to studies on the human fetus, Winick (1971) and Rosso and Winick (1974) described three phases of fetal cellular growth. The first phase is cellular hyperplasia, which includes an increase in cell number occurring during the first 16 weeks of embryonic and cellular life. Fetal insults during this phase lead to SGA with the effect being reduction in cell numbers. These infants are usually small in size with a reduction in all external measurements like weight, length and head circumference, thus resulting in a normal neonatal ponderal index. This type of SGA may be associated with congenital malformations and viral infections.

The second phase of fetal growth is concomitant cellular hyperplasia and hypertrophy. This lasts from 16 to 32 weeks during which there is a progressive decrease in the rate of hyperplasia and an increase in the rate of cellular hypertrophy. Fetal insult during this phase usually produces a mixed pattern of wasting and proportional smallness. The third phase of fetal growth extends from 32 weeks to term (Brar and Rutherford 1988). During this period, cell size increases rapidly together with increased glycogen and fat deposition. An insult during this phase leads to clinical evidence of wasting (Lin et al. 1991).

During the period of cellular hypertrophy, there is a concomitant marked increase in fetal fat accretion. Total fetal fat increases from 4% of body weight at 28 weeks gestation to 14% of body weight at 40 weeks gestation (Widdowson et al. 1979). By 28 weeks gestation, a 1.5 kg fetus has 50 g of fat; this increases to 500 g of fat in a term 3.0 kg fetus (Milner 1989). Any insult which reduces the rate of cellular hypertrophy and fat accumulation would therefore have its greatest effect in the third trimester of pregnancy.

This simplistic model of growth describes cellular growth in general but does not take into account the different rates at which different tissues develop and mature. For example, proliferation of cerebral neurones is maximal up to 18 weeks with predominantly hypertrophy occurring thereafter (Dobbing and Sands 1970) whilst cerebellar neurones proliferate throughout the first year of life (Dobbing and Sands 1973). Nevertheless, this model helps to explain the clinical observation that an insult early in pregnancy has an effect on growth potential whilst an insult later in pregnancy interferes with cellular hypertrophy (Lin et al. 1991). This model also concurs with studies on fat accretion where fetal fat, visualised using nuclear magnetic resonance imaging, increases markedly in the

third trimester of pregnancy (Garden et al. 1991).

Animal experiments suggest that cellular growth and differentiation are mediated by macromolecules which are expressed at critical periods of development and coordinate the process (Underwood et al. 1986, Gluckman et al. 1987, Han 1989). Altering the balance of expression of growth promoting factors and growth inhibitory factors may lead to reduced cellular growth and fat accumulation, the net effect of which is expressed in a reduction in fetal growth. As serial measurements of fetal size quantify the anthropometric changes of fetal growth, they reflect any reduction in the rate of cellular hypertrophy or fat accumulation particularly in the third trimester of pregnancy.

8.3.2 Uteroplacental blood flow and substrate supply

The commonest cause of IUGR in the Western world is believed to be uteroplacental insufficiency secondary to failure of the second wave of trophoblastic invasion of spiral arteries in the placental bed (Robertson et al. 1981). This underlying pathological process helps to explain the rationale of using serial measurements of fetal size to describe fetal growth.

Uteroplacental blood flow increases from 50 ml / minute at 10 weeks gestation to 600 ml / minute at term (Martin 1968). In order to achieve this vast increase, the spiral arteries of the non-pregnant uterus are modified into the uteroplacental vessels of pregnancy. At term, maternal blood enters the intervillous space through 100 to 200 spiral arteries (Brosens and Dixon 1966). Trophoblasts can be found in maternal spiral arteries from the time these arteries communicate with the intervillous space. The trophoblast disrupts the wall of the spiral artery, destroys the muscular and elastic tissues and replaces them with fibrinoid tissue. Invasion of the spiral arteries occurs in two stages (Robertson et al. 1975). The first wave occurs at the time of implantation and lasts until 10 weeks; it is limited in depth to the decidual parts of the spiral arteries. The second wave starts at 14 to 16 weeks gestation, lasts 4 to 6 weeks and invades as far as the radial artery. This second wave of invasion allows progressive distension of the arteries so that they can accommodate the increased uteroplacental blood flow.

The animal model, in particular the sheep model, allows the experimentation of IUGR and measurement of the adaptive responses of the fetus (Clapp 1989a, Clapp 1991). Different methods of inducing IUGR in different species have been used and are summarised by Cassady and Strange (1987). Chronic reduction in uteroplacental blood flow in sheep can be induced by vascular ligation or constriction that does not damage the placental tissue (Jones et al. 1985), or by embolic occlusion of the uteroplacental

vasculature that is associated with progressive placental damage (Clapp et al. 1980, Charlton and Johengen 1987). The animal model allows studies to be performed which shed some light on the chronology of IUGR. Reduced uteroplacental blood flow results in a decrease in placental size and fetal growth restriction. The degree of fetal growth restriction is linearly related to the reduction in uteroplacental blood flow (Clapp 1989b, Clapp 1991). This restriction in fetal size occurs well before metabolic changes in the fetus. Fetal substrate levels of glucose and lactate, and blood gases remain normal until late in the course, well after anthropometric evidence of IUGR is established (Robinson et al. 1985). This early reduction in fetal growth rate reduces energy requirements to match the new level of substrate availability. The precise factor that initiates the change in fetal growth rate is unknown. Continuous intravenous nutritional supplementation of the fetus during embolic placental damage completely prevents the restriction of fetal growth (Charlton and Johengen 1987), suggesting that changes in substrate availability to the placenta or fetus may be an important factor. Later on, when IUGR is well advanced, the hormonal profile in the animal model begins to change. These include a decrease in oxygen concentration, glucose and insulin levels and an increase in lactate, alanine and haematocrit (Clapp 1991). Therefore, the hypothesis is that reduced uteroplacental blood flow results in a reduction of substrate availability, thereby leading to anthropometric evidence of growth restriction well before any evidence of biochemical compensation.

The possibility that fetal growth is substrate-dependent is supported by experiments which have employed other methods of inducing IUGR. In the sheep, IUGR can also be induced by excision of the endometrial carúncles (Jones et al. 1985, Owens et al. 1987a, Owens et al. 1987b) or by restricting maternal caloric intake (Mellor 1983). By removing most of the implantation area prior to pregnancy, excision of the endometrial caruncles in the sheep results in a reduced placental size and growth restriction (Harding et al. 1985). Similarly, a reduction in maternal caloric intake reduces maternal substrate levels and produces a loss of placental and fetal weight (Mellor 1983). Although the model of IUGR induced by reduced uteroplacental blood flow may have a different pathophysiological basis to that by carunculectomy or maternal malnutrition, all result in reduced substrate uptake to match the reduced growth rate.

The precise stimulus that initiates the change in fetal growth rate in response to reduced substrate availability is unknown. Insulin and insulin-like growth factors (and associated binding proteins) have been implicated in the modulation of fetal growth. Insulin levels are reduced in response to reduced substrate supply and the effect of hypoinsulinism on fetal growth has been demonstrated in animal studies (Fowden et al. 1984, Fowden 1989, Hill 1989, Stevens et al. 1990, Fowden 1992). There is also evidence from animal data to suggest that IGF-1 may be involved in the regulation of fetal growth (Harding et al. 1985, Gluckman 1989). IGF-1 levels are decreased in experimental

models of IUGR induced by decreasing substrate supply (Robinson et al. 1985, Jones et al. 1988). Recent animal data suggest that the balance between substrate availability and fetal growth may be acutely regulated by IGF-1, IGF-2 and binding proteins (Jones 1985, Jones et al. 1988). Whether IGF-1 is directly regulated by glucose or mediated by insulin is unknown.

Whilst some of the findings in animal models may be similar to those observed in the growth retarded human fetus, not all can be extrapolated to the human fetus. In human pregnancies complicated by SGA, about 50% of women demonstrate inadequate invasion of the spiral arteries based on placental bed biopsy results (Sheppard and Bonnar 1976, Robertson et al. 1981). Therefore, only the decidual portion of each spiral artery is converted into a uteroplacental vessel, resulting in decreased perfusion of the intervillous space. This observation in only 50% of cases may be due to the fact that not all small fetuses are growth retarded. Inadequate blood volume expansion (Gibson 1973) and prolonged exercise (Clapp et al. 1991) have also been known to result in reduced fetal size, possibly exerting their restrictive effect on fetal growth through their effects on uteroplacental blood flow. The increased incidence of IUGR in patients with sickle cell disease, auto-immune diseases and pre-eclampsia may also be due to a reduction in uteroplacental blood flow.

The chronology of IUGR in the human fetus appears to show a similar pattern to that in the animal model. Data from cordocentesis studies provide indirect evidence that significant changes in oxygen, glucose and lactate concentrations do not occur until the fetus is restricted in its growth ie. size is severely reduced (Soothill et al. 1986, Economides et al. 1989b). Fetal Doppler studies based on small fetuses also provide indirect evidence that the balance between substrate availability and energy demands is maintained initially by a reduction in growth rate, but eventually results in redistribution of cardiac output that selectively protects the brain (Vyas et al. 1990, Gudmundsson and Marsal 1991b). However, all the above cordocentesis and fetal Doppler studies are based on fetuses with an AC < 2.5th centile, the inherent assumption being that reduced fetal size is the end-result of growth restriction. There have been no longitudinal studies which have directly assessed the chronology of IUGR in the human fetus with respect to serial changes in fetal size and substrate uptake.

These results from human studies are consistent with animal data which suggest that reduced substrate availability secondary to uteroplacental insufficiency results in growth restriction, measured indirectly as a severe reduction in fetal size. It is postulated that this morphometric growth restriction and reduced substrate uptake by fetal tissues commensurate with reduced energy demands is an early adaptation to the pathology of IUGR. Only when IUGR is well advanced do the mechanisms of tissue catabolism (Cetin

et al. 1988) and redistribution of cardiac output (Peeters et al. 1979) become evident. Therefore, this chronology of events provide the pathophysiological basis for the serial assessment of fetal size in the early diagnosis of IUGR.

8.3.3 Summary

From the evidence discussed above, the following factors influence fetal growth and provide the pathophysiological basis for serial measurements of fetal size in the diagnosis of IUGR;

Cellular hypertrophy and fat accretion exert their main effect in the third trimester of pregnancy. Animal experiments suggest that cellular growth and differentiation are mediated by macromolecules which are expressed at critical periods of development and coordinate the process (Gluckman et al. 1987, Han 1989).

Reduced uteroplacental blood flow may exert their influence on cellular control of fetal growth by reducing substrate supply. The first consequence of reduced uteroplacental blood flow is growth restriction and a reduction in substrate uptake by fetal tissues. Only when IUGR is well advanced do the mechanisms of tissue catabolism and Doppler abnormalities in umbilical and fetal blood flow become evident. The chronology of events suggests that anthropometric growth restriction is one of the first adaptations to IUGR and therefore provide the pathophysiological basis for quantifying serial measurements of fetal size.

8.4 Robustness of ultrasound in the diagnosis of growth failure in small fetuses

The concept of fetal growth is one which involves a change in fetal size (Altman and Hytten 1989). The robustness of ultrasound in assessing change in size in a group of small fetuses had not previously been tested against standard neonatal morphometry.

Although SGA per se already confers increased risk of perinatal morbidity and mortality (Section 2.2.3.2), the growth retarded small fetus is at further additional risk of subsequent perinatal morbidity (Walther and Raemaker 1982, Haas et al. 1987, Patterson and Pouliot 1987a, Villar et al. 1990, Fay et al. 1991a). For this reason, the present study was confined to a cohort of fetuses defined to be small by antenatal criteria. The ability of serial ultrasound measurements of fetal size to separate these small fetuses into those with and without ultrasound evidence of IUGR would be tested against neonatal morphometry.

Abdominal circumference and EFW are the best ultrasound predictors of a SGA infant at birth (Section 2.3.1). Therefore, serial ultrasound measurements of AC and EFW were evaluated in their ability to diagnose IUGR within a group of small fetuses. Crucial to the diagnosis of abnormal growth was the availability of appropriate reference standards of normal growth. Reference standards derived from longitudinal data are preferable to cross-sectionally collected data especially when the express purpose of such standards is to define growth rather than size (Evans et al. 1990, Sparks and Cetin 1991, Bland and Altman 1992). The limitations with all previous AC and EFW reference standards derived from longitudinal data have been enumerated in Section 3.1.2. It was apparent that all such reference ranges were derived using inappropriate methods of statistical analysis. Furthermore, conflicting data on whether growth is linear have been reported. The results reported in Section 5.3 clearly demonstrated that the linear model was inappropriate for the purposes of describing growth. Use of the linear model would have resulted in a systematic over-estimation in AC and EFW values towards term (Figure 5.2). Therefore, use of a linear model to evaluate growth, as suggested by Deter et al. (1982a), would lead to a disproportionately high number of normally growing fetuses being diagnosed as growth retarded towards term. By contrast, the quadratic model did not systematically under- or over-estimate values of AC and EFW towards term (Table 5.3). Therefore, the reference standards reported in Tables 5.4 and 5.5 provided appropriate standards for subsequent ultrasound assessment of fetal growth.

The results of reproducibility of ultrasound measurements (Section 6.1) had significant implications on the interpretation of serial data obtained in the subsequent study. In clinical practice, ultrasound scans are usually performed at two-weekly intervals

for assessment of change in fetal size. However, the magnitude of the error of ultrasound measurements had not previously been compared with the change in fetal size over two weeks. Assessment of the inter-observer variability of ultrasound measurements by the limits of agreement method (Bland and Altman 1986) enabled such a comparison to be made. The 95% inter-observer prediction intervals for AC and EFW were (-16.8 g, 13.0 mm) and (-160, 124 g) respectively (Table 6.2). The limits of agreement for AC were greater than the ±2 SD limits of increments in AC per week at any gestation (Figure 6.2). Sarmandal et al. (1989) reported limits of agreement for AC which were much wider than those in this study. These results of reproducibility of ultrasound measurements seriously question the validity of performing ultrasound scans at less than two-weekly intervals. For this very reason, the minimum time interval over which Δ SDS was calculated was two weeks. Similarly, FGV was quantified using increments in AC and EFW over two weeks (Section 4.5.3).

The quantification of fetal growth posed a particular problem as growth was non-linear. The use of absolute increments in AC and EFW per two weeks to quantify growth would require normal standards for each gestational age interval. Two groups have previously reported reference standards for increments in AC (Fescina et al. 1982, Deter and Harrist 1992). Reference standards for increments in AC and EFW derived from two-weekly data were also reported in this study (Tables 5.6). However, these reference standards are of limited value in clinical practice when ultrasound scans may have been performed at intervals greater than two weeks gestation. Reference standards for a whole permutation and combination of intervals other than two weeks do not exist and would be very cumbersome to use even if available.

The use of Δ SDS allowed change in size to be measured relative to standards which took into account the mean (and SD) measurements at the relevant gestational ages. Conceptually, this was an appealing method of quantifying changes in measurements of fetal size in view of the non-linearity of growth. The Δ SDS for AC and EFW in this study were calculated using only the first and last SDS for each fetus. Whilst this represented the change in fetal size over the entire gestational age interval over which the fetus was assessed, data collected between these two gestations were not used. The limitations with this are obvious; it placed an over-dependence on the accuracy of these two scans and did not take into account SDS values in the intervening period. Nevertheless, this mirrored clinical practice to some extent as decisions are usually made in the light of results of the last ultrasound scan. Whilst future mathematical models will have to be developed to take into account all SDS values, the methodology described here represented the first step in the quantification of serial ultrasound data.

The disadvantage with Δ SDS was the assumption that growth was normal up to

the first ultrasound scan and that IUGR could be accurately diagnosed by assessing the degree of deviation in SDS thereafter. If the first scan was performed at or before 28 weeks gestation, this assumption may be valid as cellular hypertrophy (Section 8.2.2) and fat accretion (Section 8.2.2.) have all been shown to exert their main effect in the third trimester. However, only 37 (35.6%) of all the fetuses studied had their first ultrasound scan performed by the author at or before 28 weeks gestation. The gestation at the first assessment would not have mattered if the fetus was SGA but growing normally as the SDS would be approximately the same at any given gestation. However, a growth retarded fetus could have had a reduction in growth prior to the first ultrasound assessment. By quantifying the change in SDS thereafter using Δ SDS, the degree of IUGR in that particular fetus may be under-estimated. Nevertheless, the advantage with this method of quantifying growth was that it could be used irrespective of the gestational age at which the fetus was first assessed.

Other workers have suggested an additional routine ultrasound scan at 26 weeks gestation to establish the growth potential for individual fetuses, so that any growth deviation thereafter can be diagnosed relative to the individual growth trajectory of that fetus (Deter and Rossavik 1987, Simon et al. 1989, Deter Deter et al. 1990). The Rossavik Growth Model requires the data from two ultrasound scans performed prior to 27 weeks gestation (Deter et al. 1989a, Deter et al. 1989b) (see Section 3.1.3). Whilst a level II scan at 18 to 20 gestation is standard practice in normal obstetric care, the introduction of a second scan at 26 weeks for all obstetric patients would incur great expense and double the routine scanning workload in an ultrasound department. The benefits of such a scanning practice would need to be evaluated in a randomised controlled trial, before its introduction into routine clinical practice.

In the present study, the antenatal diagnosis of IUGR was made using a cut-off of Δ EFW.SDS > -1.5 SD. This same value was found to be the optimal cut-off for predicting all neonatal morphometric indices and adverse perinatal outcome. Receiver operating characteristic curves obviated the use of an arbitrary cut-off to divide small fetuses into those with or without ultrasound evidence of IUGR. It is of note that Δ EFW.SDS was superior to Δ AC.SDS, suggesting the additional benefit of including EFW in the ultrasound assessment of small fetuses. This same ultrasound measure proved to be superior to single estimates of fetal size in the prediction of neonatal morphometry and adverse perinatal outcome. This was not surprising as AC.SDS and EFW.SDS at the last scan were single static measures of size, and conveyed little information on preceding growth. Nevertheless, the sensitivities achieved with these single measures of size were comparable to those using Δ SDS. This suggests that fetuses with severe IUGR ultimately ended up with low values of AC.SDS and EFW.SDS, especially as all fetuses in this study already had, by definition, an AC < 10th centile for gestational age at inclusion into

the study. However, the odds ratios for Δ SDS were overwhelmingly superior to those for measures of size, rendering the the latter of less value in the management of the individual small fetus.

The superiority of Δ EFW.SDS over umbilical artery Doppler waveform indices may be related to the chronology of IUGR. Animal studies have previously shown that reduction in substrate availability lead to a reduction in morphometric growth before any biochemical or umbilical Doppler evidence of IUGR (Clapp 1991) (Section 8.2.2). The Doppler assessment of the umbilical circulation provides information about downstream vascular resistance. Giles et al. (1985) examined the placentas from three groups of women; those with normal pregnancies and a normal systolic / diastolic ratio from the umbilical artery, those with an at-risk pregnancy but a normal systolic / diastolic ratio, and those with an at-risk pregnancy and abnormal systolic / diastolic ratio. They demonstrated a reduction in the number of tertiary stem arterioles in women with abnormal systolic / diastolic ratios. Tertiary stem arterioles, the fetal placental resistance vessels, are the level at which there is a maximum fall in blood pressure. This work has been confirmed by McCowan et al. (1989) who also demonstrated a linear inverse relationship between the systolic / diastolic ratio and the number of small arterioles.

If the pathophysiological basis for IUGR is reduced uteroplacental perfusion leading to hypoxic ischaemia of the intervillous space and vasoconstriction of the fetal placental vessels (Rankin and McLoughlin 1979), such an increase in peripheral vascular resistance will be reflected in abnormal umbilical waveforms. Experiments on fetal lambs have shown that abnormal umbilical artery Doppler waveforms were produced using the gradual embolisation model (Morrow et al. 1989) but not by mild to moderate hypoxia (Morrow et al. 1990). In the embolisation model, initial changes in umbilical blood flow act to improve placental perfusion; this helps to maximise the remaining placental transfer capacity (Clapp 1989a). With increasing embolisation, a gradual change in the Doppler waveform is observed with progressive reduction in the diastolic component (Morrow et al. 1989). It is therefore possible that the poor results obtained using umbilical artery PI in this study may be related to the mild degree of IUGR in these small fetuses at term where significant biochemical or Doppler abnormalities had not yet occurred.

The results of the present study show that serial ultrasound measurements of AC and EFW, as quantified by by Δ SDS, are predictive of subsequent neonatal morphometry in a group of small fetuses. The superiority of this ultrasound measure of growth over single estimates of fetal size and umbilical artery PI confirm its usefulness in the description of fetal growth and antenatal surveillance. This method of quantifying growth relies on the assumption that growth prior to ultrasound assessment was normal. This may be true in IUGR where the aetiology is one of reduced uteroplacental perfusion leading to

reduced cellular growth and fat deposition, this process becoming progressively more pronounced with advancing gestational age in the third trimester (Sections 8.3.1 to 8.3.2). However, an insult such as viraemia occurring earlier in pregnancy would have altered growth in-utero which then "resets" at a lower level (Clapp 1991). Though serial ultrasound may be useful in evaluating fetal growth, its limitations are highlighted by the tests of reproducibility. A reliable diagnosis of IUGR using serial ultrasound can only be made if serial scans are performed at intervals not less than two weeks.

8.5 Clinical implications of a diagnosis of IUGR using serial ultrasound assessment of fetal size

The clinical significance of separating small fetuses into those with ultrasound evidence of IUGR and those with normal growth was assessed by comparing the perinatal outcome and biochemical indices of IUGR in the two groups. As premature delivery contributes significantly towards perinatal morbidity, only fetuses delivered beyond 36 weeks gestation were included in the study. This contrasts with most previous ultrasound studies where predictions of adverse perinatal outcome were evaluated in fetuses of all gestational ages (Berkowitz et al. 1988b, Chambers et al. 1989, Burke et al. 1990, Gudmundsson and Marsal 1991b). The disadvantage with these studies was that perinatal outcome was not solely related to the effects of IUGR, in view of the compounding effects of prematurity.

The resultant two groups of fetuses were significantly different in the proportion of neonates admitted to NICU for morbidity related to IUGR (Table 7.11). However, no significant differences in the incidence of fetal distress in labour, pH at delivery or incidence of low Apgar score at 5 minutes were evident between the two groups. This may have been related to the low morbidity in the group as a whole. However, it is also possible that the classification of small fetuses by serial ultrasound was of little clinical relevance when these measures were used to define asphyxia.

Neonates classified using ultrasound criteria had significant differences in umbilical venous IGF-1 levels at delivery but no significant differences in the biochemical indices of malnutrition investigated. In the light of animal studies (Section 8.3.2), these findings are consistent with the hypothesis that one of the earliest manifestations of reduced substrate availability is the a reduction in fetal growth rate prior to any biochemical evidence of decompensation. Human cordocentesis studies have shown that some fetuses with an AC < 2.5th centile are hypertriglyceridaemic (Economides et al. 1988), hypoinsulinaemic (Economides et al. 1989b), hypoxic, acidaemic and hypoglycaemic (Soothill et al. 1987a). However, the fetuses evaluated in the present study are not comparable to those investigated by cordocentesis. The degree of IUGR in the present study may therefore be commensurate with ultrasound evidence of change in size without significant changes in the metabolic status of the fetus. The significant differences in IGF-1 levels may be related to its role in the modulation of fetal growth. Animal studies suggest that the balance between substrate availability and fetal growth may be acutely regulated by IGF-1 and IGF-2 and binding proteins (Jones 1985). In response to decreased uteroplacental flow, the IGFs and binding proteins act to reduce fetal protein synthesis and ultimately result in morphometric evidence of IUGR (Jones et al. 1988). Such findings

would explain the significant differences in umbilical venous IGF-1 levels in the absence of metabolic changes related to IUGR.

In this study, the clinical significance of separating small fetuses by their antenatal growth profile was only evaluated with respect to measures of perinatal morbidity and biochemical indices of malnutrition at delivery. However, it is evident that these represent only two of the many clinical consequences of IUGR. Numerous studies have reported an association between IUGR and subsequent postnatal growth, neurodevelopmental outcome and predisposition to certain diseases in later adult life. As most of these associations have been reported using indirect or inappropriate measures of fetal growth, serial ultrasound provides the opportunity to investigate these purported associations further.

Previous studies have reported the phenomenon of "catch-up growth" in the postnatal period following IUGR (Davies et al. 1979, Villar et al. 1982, Walther and Raemaker 1982, Villar et al. 1984). This phenomenon describes the accelerated somatic growth which follows a period of growth restriction, the purpose of which is to return the growing infant to its normal growth trajectory (Davies 1981). All previous studies which have evaluated the postnatal growth profile have defined antenatal growth using the ponderal index at birth. Small for gestational age neonates with a low ponderal index have been shown to grow at a faster rate than SGA neonates with a normal ponderal index (Walther and Raemaker 1982, Villar et al. 1984, van Vugt et al. 1991). As yet, no study has evaluated the relationship between fetal growth as assessed by serial measurements of fetal size and postnatal growth. Such longitudinal follow-up studies would further define the relationship between fetal and postnatal growth.

There is much controversy about the mental, behavioural and neurodevelopment of infants with IUGR. Many of these follow-up studies have inappropriately defined IUGR as a low birthweight centile (Ounsted et al. 1983, Rantakillio 1985). Most studies on term infants with a birthweight less than the 3rd or 10th centile have shown that the vast majority of such infants have a normal IQ and show no evidence of major handicap at follow-up (Allen 1984). However, these studies have in reality only addressed the outcome of SGA infants rather than infants with evidence of IUGR. Four studies have previously reported data on neurodevelopmental outcome in a group of SGA neonates who had undergone serial ultrasound measurements of BPD in-utero (Fancourt et al. 1976, Harvey et al. 1976, Parkinson et al 1981, Harvey et al. 1982). Those SGA neonates whose rate of growth of BPD had begun to slow in-utero before 34 weeks gestation were more likely to have a height and weight below the 10th centile in the first years of life (Fancourt 1976) and have a lower developmental quotient and cognitive ability (Harvey et al. 1982). No differences were noted when such infants were divided

into groups using birthweight centiles. In a study assessing school achievement and behaviour of children, those with a slowing of head growth before 26 weeks were poorer achievers, were less able to concentrate and had more problems at school (Parkinson et al. 1981). However, the major limitation with these studies was the use of BPD, one of the last indices to be altered in IUGR as a result of the "brain-sparing" effect (Campbell and Thoms 1977, Kramer et al. 1989), to quantify fetal size. Two other studies (Villar et al. 1984, Blair and Stanley 1992) have suggested that SGA neonates with a low ponderal index or morphometric evidence of malnutrition at birth were at increased risk of cerebral palsy in later childhood life. However, no study has compared fetal growth (quantified using changes in AC or EFW) with subsequent neurological outcome.

Finally, Barker and co-workers (1992) have suggested that fetal growth may be inextricably linked with development of cardiovascular disease and other important disorders in later adult life. In geographical studies based on men born between 1920 and 1930 in Hertfordshire, the incidence of hypertension in adult of life was significantly related to the birthweight / placental weight ratio, with a low ratio predictive of subsequent hypertension (Barker et al. 1990). The hypothesis suggested by Barker et al. (1990) was that placental weight is increased relative to birthweight in IUGR, with important implications in later life as a result of this " programming" in fetal life. Other studies performed by the same group (Barker et al. 1989, Hales et al. 1991) related the development of ischaemic heart disease and maturity-onset diabetes mellitus in later adult life to birthweight and weight at one year of age. Quite how these indices reflected fetal growth remain unclear. If the placental weight / birthweight ratio is indeed indicative of IUGR, serial ultrasound measurements of fetal size would provide an invaluable starting point with which to define fetal growth and confirm or refute this hypothesis.

In summary, the robustness of serial ultrasound measurements of anthropometric size allow fetal growth to be quantified antenatally. Some differences in perinatal morbidity and biochemical indices of malnutrition were evident between the two groups. This classification of small fetuses on the basis of serial ultrasound data provide the basis for future follow-up studies investigating the other consequences of IUGR in childhood and later adult life.

8. 6 Epilogue

The use of serial ultrasound measurements of anthropometric size to quantify fetal growth represents a significant improvement in the diagnosis of altered growth in-utero. The replacement of birthweight by neonatal morphometric indices of "wasting" as the "gold standard" further redefines the outcome criteria to be used in subsequent studies of IUGR. This study was confined to the small fetuses; the usefulness of serial AC or EFW measurements in the separation of AGA fetuses into those with or without IUGR was not investigated. Further studies will have to be performed to address this different group of fetuses.

The results of the reproducibility studies suggested that intervals between measurements of less than 2 weeks were not appropriate for accurate quantification of true fetal growth. Beyond that, the optimal interval between scans was not evaluated in this study. Furthermore, the usefulness of this method of quantifying fetal growth needs to compared with that suggested by Deter and co-workers who espouse the use of the Rossavik Growth Model.

The effect of quantifying serial ultrasound measurements of size on subsequent perinatal outcome will need to be evaluated in a randomised controlled study. Serial measurements of size are already used clinically but the quantification of fetal growth needs to be tested against the visual interpretation of graphical plots of serial measurements to ascertain if the former presents any advantage over the latter in relation to subsequent outcome.

The use of serial values of AC or EFW to quantify growth in-utero provides an opportunity to evaluate the morphological and neurodevelopmental outcome of such fetuses in the first years of life. Many such studies have inappropriately used birthweight to define IUGR. The infants from Studies 1 and 3 have been recruited into a follow-up study to investigate the relationship between fetal growth and growth in the first four years of life; this study is being coordinated by Professor Michael Preece and Dr. Les Cox at the Department of Growth and Development, Institute of Child Health, London. The neurodevelopmental outcome of infants from Studies 1 and 3 is also being evaluated by Dr. Simon Roth, Department of Paediatrics and Neonatology, University College Hospital, London at yearly intervals to ascertain any differences in the outcome of fetuses classified to be growth retarded based on serial ultrasound data. It is hoped that these studies will answer some of the questions concerning the long-term outcome of growth retarded fetuses.

SUMMARY

Introduction

1. Many studies have used SGA to define IUGR. However, this is inappropriate as many SGA neonates are constitutionally small with no morphometric evidence of IUGR. Intrauterine growth retardation is best defined using neonatal morphometric indices which reflect "wasting", such as the ponderal index, MAC / HC ratio, and subscapular and triceps skinfold thicknesses.
2. The antenatal separation of small fetuses into those with IUGR and those with normal growth is important in view of the increased perinatal morbidity and mortality in the former compared with the latter.
3. As AC and EFW are the best ultrasound parameters for predicting SGA, serial ultrasound assessment of AC and EFW may be useful in identifying those small fetuses with evidence of growth retardation.

Aim

The aim of the study was to evaluate the ability of serial ultrasound measurements of AC and estimates of fetal weight to diagnose IUGR in a group of small fetuses who delivered at term.

Factors influencing study design

1. All published reference standards for AC and EFW constructed from longitudinal data have been derived using inappropriate statistical methods of analysis. There was therefore a need to derive appropriate reference standards for these ultrasound parameters from longitudinal data using a statistical method which took into account the inter-dependence of repeated measurements.
2. The reproducibility of these ultrasound measurements had to be evaluated in order to determine the minimum interval between ultrasound assessments.
3. All previous methods of quantifying serial values of AC and EFW have been based on the Rossavik Growth Model. However, this model has never been validated against neonatal morphometric indices, the "gold standard" for defining IUGR. There was therefore a need to determine the best method of quantifying serial measurements.

Methods

Study 1: Reference standards for AC and EFW were derived from the serial ultrasound assessment of 67 Caucasian fetuses from 20 weeks until term. Estimated fetal weight was calculated using the four-parameter formula of Hadlock et al. (1985). Four classes of mathematical models (linear, quadratic, Gompertz and Rossavik) were fitted to the $\log_{10}$(AC) and $\log_{10}$(EFW) data from each fetus using least squares regression analysis.

The best model would result in the lowest standard deviation of the residual error, and would not systematically over- or under-estimate the final $\log_{10}$(AC) or $\log_{10}$(EFW) when these data are omitted from the fitting process.

Study 2: The reproducibility of AC and EFW was investigated in a study involving 40 fetuses. The reproducibility of neonatal morphometric measurements was also investigated in another group of 30 neonates. Intra-observer and inter-observer variability were assessed using one-way analysis of variance and the limits of agreement method of Bland and Altman (1986) respectively.

Study 3a: In a study of 104 small fetuses (defined as an AC < 10th centile in the third trimester of pregnancy), ROC curves were used to compare three different methods (Δ SDS, value of the *b* -coefficient and fetal growth velocity per two weeks) of quantifying serial values of AC and EFW in the prediction of abnormal neonatal morphometric indices of IUGR. The best method would result in the largest area under the ROC curves, and highest sensitivities, odds ratios and Cohen's Kappa values.

Study 3b: Receiver operating characteristic curves were also used to compare the best measure of serial AC / EFW data with umbilical artery PI and single estimates of fetal size in the prediction of abnormal neonatal morphometry and adverse perinatal outcome.

Study 3c: The optimal cut-off criterion on the ROC curve was used to separate these fetuses into two groups. Measures of perinatal morbidity and biochemical indices of IUGR were compared in the resultant two groups.

Results

1. Study 1: The linear model resulted in the largest standard deviation of the residual error. The linear model also systematically over-estimated the final $\log_{10}$(AC) while the linear, Rossavik and Gompertz models all over-estimated the final $\log_{10}$(EFW) when these data were omitted from the curve fitting process. The log quadratic model was therefore used to derive reference ranges for AC and EFW.

2. Study 2: The intra-observer standard deviations for AC and EFW were < 5 mm and < 70 g respectively. The corresponding 95% limits of agreement for inter-observer differences were (-16.8, 13.0 mm) and (-159.9, 124.3 g) respectively. The intra-observer variability of all neonatal morphometric measurements was low. The inter-observer differences, expressed as limits of agreement, were low with respect to the reference ranges for these morphometric measurements.

Study 3a: The statistical method, Δ SDS, resulted in the largest area under the ROC curve in the prediction of all neonatal morphometric indices. This measure of serial AC / EFW data also resulted in the highest sensitivities, odds ratios and Cohen's Kappa values.

Study 3b: Serial values of AC and EFW, as quantified by Δ SDS, were superior to umbilical artery PI and single estimates of fetal size in the prediction of abnormal neonatal morphometry and adverse perinatal outcome.

Study 3c: Separation of the 104 small fetuses into two groups according to their Δ SDS values resulted in two groups with some differences in perinatal morbidity and biochemical indices of IUGR. Significantly more neonates were admitted to NICU in the group with ultrasound evidence of IUGR. No significant differences in the incidence of Caesarean section for fetal distress, acidaemia or low Apgar scores at 5 minutes were found between the two groups. Neonates classified antenatally to be growth retarded had significantly lower umbilical venous IGF-1 levels than those with ultrasonic normal growth. No significant differences in umbilical venous levels of glucose, insulin or triglycerides were noted between the two groups.

Discussion

1. Reference standards for AC and EFW were derived by fitting a log quadratic model to the individual data of 67 fetuses. These standards proved to be useful for the subsequent assessment of serial ultrasound measurements.
2. The tests of reproducibility of the ultrasound measurements, AC and EFW, suggested that the minimum interval between assessments should be two weeks. The intra- and inter-observer variabilities of neonatal morphometric measurements were low, confirming their usefulness as the "gold standard" to define IUGR.
3. Serial values of AC and EFW, as quantified by Δ SDS, were accurate in the prediction of neonatal morphometry in a group of small fetuses. Division of small fetuses according to their serial measurements resulted in two groups with some differences in perinatal morbidity and biochemical indices indicative of IUGR.
4. The pathophysiology of IUGR is believed to be reduced uteroplacental perfusion, resulting in reduced substrate supply. One of the earliest fetal responses to this is restriction in fetal growth. This provides the pathophysiological basis for the diagnosis of IUGR using serial assessment of fetal size. The chronology of events in IUGR may explain the superiority of serial measurements of fetal size to umbilical artery Doppler waveform indices or single estimates of fetal size in the diagnosis of IUGR.
5. The accurate quantification of fetal growth by serial AC and EFW data provides the basis for future research into the postnatal morphological and neurodevelopmental consequences of altered fetal growth in-utero.

Publications

Four publications have resulted from this work:

1) Chang TC, Robson SC, Boys RJ, Spencer JAD. Prediction of the small-for-gestational age infant: Which ultrasound measurement is best? Obstet Gynecol 1992; 80: 1030-1038.

2) Gallivan S, Robson SC, Chang TC, Vaughan J, Spencer JAD. An investigation of fetal growth using serial ultrasound data. Ultrasound Obstet Gynecol 1993; in press.

3) Chang TC, Robson SC, Spencer JAD, Gallivan S. Prediction of fetal growth retardation in small babies: Comparison of Doppler waveform indices and serila ultrasound measurements of abdominal circumference and fetal weight. Obstet Gynecol 1993; in press.

4) Chang TC, Robson SC, Spencer JAD, Gallivan S. Ultrasonic evaluation of fetal weight: Analysis of intra- and inter-observer variability. J Clin Ultrasound 1993; in press.

In addition, parts of the material have been published as abstracts and presented at research meetings, both national and international.

BIBLIOGRAPHY

Abdulla U, Campbell S, Dewhurst CJ, Talbert D, Lucas M, Mullarkey M. Effect of diagnostic ultrasound on maternal and fetal chromosomes. Lancet 1971; ii: 829-831.

Abdulla U, Talbert D, Lucas M, Mullarkey M. Effect of ultrasound on chromosomes of lymphocyte cultures. Br Med J 1972; 3: 797-799.

Abramowicz JS, Warsof SL, Sherer DM, Levy DL, Woods JR. Value of a random single Doppler study of the umbilical artery for predicting perinatal outcome. J Ultrasound Med 1991; 10: 337-339.

Adelstein P, Fedrick J. Antenatal identification of women at increased risk of being delivered of a low birth weight infant at term. Br J Obstet Gynaecol 1978; 85: 8-12.

Albano JDM, Ekins RP, Maritz G. A selective precise radioimmunoassay of serum insulin relying on charcoal separation of bound and free hormone moieties. Acta Endocrinol 1972; 70: 487-509.

Al-Ghazali W, Chapman MG, Allan LD. Doppler assessment of the cardiac and uteroplacental circulations in normal and complicated pregnancies. Br J Obstet Gynaecol 1988; 95: 575-80.

Allen MC. Developmental outcome and follow-up of the small-for-gestational age infant. Semin Perinatol 1984; 8: 123-156.

Alstrom H, Axelsson O, Cnattingius S, Ekman G, Maesel A, Ulmsten U, Arstrom K, Marsal K. Comparison of umbilical-artery velocimetry and cardiotocography for surveillance of small-for-gestational age fetuses. Lancet 1992; 340: 936-940.

Altman DG, Coles ES. Nomograms for precise determination of birthweight for dates. Br J Obstet Gynaecol 1980; 87: 81-86.

Altman DG, Hytten FE. Intrauterine growth retardation: let's be clear about it. Br J Obstet Gynaecol 1989; 96: 1127-1128.

American Institute of Ultrasound in Medicine. Bio-effects considerations for the safety of diagnostic ultrasound. J Ultrasound Med 1988; 7 (suppl).

American Institute of Ultrasound in Medicine. Standards for real-time display of thermal and mechanical acoustic output indices on diagnostic ultrasound equipment. Rockville, Maryland. 1992.

Amiel-Tison C, Barrier G, Shnider SM, Levinson G, Hughes SC, Stefani SJ. A new neurologic and adaptive capacity scoring system for evaluating obstetrc medications in full-term newborns. Anethesiology 1982; 56: 340-350.

Arabin B, Mohnkaupt A, Becker R, Weitzel HK. Compariosn of the prognostic value of pulsed Doppler blood flow parameters to predict SGA and fetal distress. Ultrasound Obstet Gynecol 1992; 2: 272-278.

Arduini D, Rizzo G, Romanini C, Mancuso S. Fetal blood flow velocity waveforms as predictors of growth retardation. Obstet Gynecol 1987; 70: 7-11.

Arduini D, Rizzo G. Normal values of pulsatility index from fetal vessels: A cross-sectional study on 1556 healthy fetuses. J Perinat Med 1990; 18: 165-172.

Bailey SM, Sarmandal P, Grant JM. A comparison of three methods of assessing inter-observer variation applied to measurement of the symphysis-fundal height. Br J Obstet Gynaecol 1989; 96: 1266-1271.

Baker DW. Pulsed ultrasonic Doppler blood-flow sensing. IEEE Trans on S and US 1970; SU-17: 1-185.

Bakketeig LS, Jacobsen G, Brodtkorb C, Eriksen BC, Eik-Nes SH, Ulstein M, Balstad P, Jorgansen NP. Randomised controlled trial of ultrasonographic screening in pregnancy. Lancet 1984; ii: 207-210.

Barker DJP, Osmond C, Winter PD, Margetts B, Simmonds S. Weight in infancy and death from ischaemic heart disease. Lancet 1989; i: 577-580.

Barker DJP, Bull AR, Osmond C, Simmonds SJ. Fetal and placental size and risk of hypertension in later adult life. Br Med J 1990; 301: 259-262.

Barker DJP. Fetal growth and adult disease. Br J Obstet Gynaecol 1992; 99: 275-282.

Bause GS, Niebyl JR, Sanders RC. Doppler ultrasound and maternal erythrocyte fragility. Obstet Gynecol 1983; 62: 7-10.

Beattie RB, Dornan JC. Antenatal screening for intrauterine growth retardation with umbilical artery Doppler ultrasonography. Br Med J 1989; 298: 631-635.

Bennett A, Wilson DM, Liu F, Nagashima R, Rosenfield RG, Hintz RL. Levels of insulin-like growth factors I and II in human cord blood. J Clin Endocrinol Metab 1983; 57: 609-612.

Benson CB, Doubilet PM, Saltzman DH. Intrauterine growth retardation: Predictive value of ultrasound criteria for antenatal diagnosis. Radiology 1986; 160: 415-417.

Berkowitz GS, Chitkara U, Rosenberg J, Cogswell C, Berkowitz RL. Sonographic estimation of fetal weight and Doppler analysis of umbilical artery velocimetry in the prediction of intrauterine growth retardation: A prospective study. Am J Obstet Gynecol 1988a; 158: 1149-1153.

Berkowitz GS, Mehalek KE, Chitkara U, Rosenberg J, Cogswell C, Berkowitz RL. Doppler umbilical velocimetry in the prediction of adverse perinatal outcome in pregnancies at risk of intrauterine growth retardation. Obstet Gynecol 1988b; 74: 742-725.

Bernadini I, Evans MI, Nicolaides KH, Economides DL, Gahl WA. The fetal concentrating index as a gestational age-independent measure of placental dysfunction in intrauterine growth retardation. Am J Obstet Gynecol 1991; 164: 1481-1490.

Bernstein IM, Meyer MC, Simmons GM, Capeless EL. Fetal growth charts: comparison of cross-sectional ultrasound examinations with birthweight. Am J Obstet Gynecol 1992; 166: 329 (abstract).

Bewley S, Cooper D, Campbell S. Doppler investigations of uteroplacental blood flow resistance in the second trimester: a screening study for pre-eclampsia and intrauterine growth retardation. Br J Obstet Gynaecol 1991; 98: 871-879.

Bilardo CM, Nicolaides KH, Campbell S. Doppler measurements of fetal and uteroplacental circulation: relationship with umbilical venous blood gases measuremd at cordocentesis. Am J Obstet Gynecol 1990; 162: 115-120.

Birnholz L. Ultrasound measurements in Obstetrics. In: Deter RL ed. Ouantitative obstetrical ultrasonography. John Wiley and Sons, Chichester. 1986; 23-60.

Blair E, Stanley F. Intrauterine growth and spastic cerebral palsy. II. The association with morphology at birth. Early Hum Dev 1992; 28: 91-103.

Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; i: 307-310.

Bland JM, Altman DG. Statistics as applied to Doppler ultrasound studies. In: Pearce JM. ed. Doppler Ultrasound in Perinatal Medicine. Oxford University Press, Oxford. 1992: 17-62.

Bottoms SF, Zador IE, Chan KL. Prematurity and fetal growth: "Normal" weights based on abnormal pregnancies. Am J Obstet Gynecol 1992; 166: 419 (abstract).

Bozzeti P, Ferrari MM, Marconi AM, Ferrazzi E, Pardi G, Makowski E, Battaglia FC. The relationship of maternal and fetal glucose concentrations in the human from mid-gestation until term. Metabolism 1988; 37: 358-363.

Brans YW, Sumners JE, Dweck HS. A non-invasive approach to body composition in the newborn: Dynamic skinfold measurements. Pediatr Res 1974; 8: 215-222.

Branson RS, Vaucher YE, Harrison GG et al. Inter- and intra-observer reliability of skinfold thickness measurements in newborn infants. Hum Biol 1982; 54: 137-143.

Brar HS, Rutherford SE. Classification of intrauterine growth retardation. Semin Perinatol 1988; 12: 2-10.

Brennan P, Silman A.Statistical methods for assessing observer variability in clinical measures. Br Med J 1992; 304: 1491-1494.

Brenner WE, Edelman DA, Hendricks CH. A standard of fetal growth for the United States of America. Am J Obstet Gynecol 1976; 126: 555-564.

Breslow NE, Day NE. The analysis of case-control studies. IARC Scientific Publications No. 32, Lyon, France. 1980.

Brosens I, Dixon HG. Anatomy of the maternal side of the placenta. J Obstet Gynaecol Br Commonw 1966; 73: 357-363.

Brown HL, Miller JM, Gabert HA, Kissling G. Ultrasonic recognition of the small-for-gestational age fetus. Obstet Gynecol 1987; 69: 631-635.

Buckton KE, Baker NV. An investigation into possible chromosome damaging effects of

ultrasound on human blood cells. Br J Radiol 1972; 45: 340-342.

Burke G, Stuart B, Crowley P, Scanaill SN, Drumm J. Is intrauterine growth retardation with normal umbilical artery blood flow a benign condition? Br Med J 1990; 300:1044-1045.

Burkinshaw L, Jones PRM, Krupowicz DW. Observer error in skinfold thickness measurements. Human Biol 1973; 45: 273-279.

Burns PN. Measuring volume flow with Doppler ultrasound - an old nut. Ultrasound Obstet Gynecol 1992; 2: 238-241.

Campbell S. The prediction of fetal maturity by ultrasound measurement of the biparietal diameter. J Obstet Gynaecol Br Commonw 1969; 76: 603-607.

Campbell S, Dewhurst CJ. Diagnosis of small-for-dates dates fetus by serial ultrasonic cephalometry. Lancet 1971; ii: 1002-1006.

Campbell S, Wilkin D. Ultrasound measurement of fetal abdominal circumference in the estimation of fetal weight. Br J Obstet Gynaecol 1975; 82: 689-697.

Campbell S. Fetal growth. In: Beard RW, Nathanielsy PR eds. Fetal Physiology and Medicine. WB Saunders and Co. Ltd, London. 1976: 271-300.

Campbell S, Thoms A. Ultrasound measurement of the fetal head to abdomen ratio in the assessment of growth retardation. Br J Obstet Gynaecol 1977; 84: 165-170.

Campbell S, Warsof SL, Little D, Cooper D. Routine ultrasound screening for the prediction of gestational age. Obstet Gynecol 1985; 65: 613-620.

Campbell S. The detection of intrauterine growth retardation. In: Sharp F, Fraser RB, Milner RDG eds. Fetal Growth. Proceedings of the 20th Study Group of the Royal College of Obstetricians and Gynaecologists. Royal College of Obstetricians and Gynaecologists, London. 1989: 251-261.

Cassady G, Strange M. The small for gestational age infant. In: Avery GB ed. Neonatology. Pathophysiology and management of the newborn. JB Lippincott Company, Philadelphia. Third edition. 1987: 299-331.

Catalano PM, Tyzbir ED, Allen SR, McBean JH, McAuliffe TL. Evaluation of fetal

growth by estimation of neonatal body composition. Obstet Gynecol 1992; 79: 46-50.

Cetin I, Marconi AM, Bozzetti P, Sereni LP, Corbetta C, Pardi G, Battaglia FC. Umbilical amino acid concentrations inappropriate and small for gestational age infants: a biochemical difference present in-utero. Am J Obstet Gynecol 1988; 158: 120-126.

Chamberlain PF, Manning FA, Morrisin I, Harman CR, Lange IR. Ultrasound evaluation of amniotic fluid volume. I. The relationship of decreased amniotic fluid volume to perinatal outcome. Am J Obstet Gynecol 1984; 150: 245-250.

Chambers SE, Hoskins PR, Haddad NG, Johnstone FD, McDicken WN, Muir BB. A comparison of fetal abdominal circumference measurements and Doppler ultrasound in the prediction of small-for-dates babies and fetal compromise. Br J Obstet Gynaecol 1989; 96: 803-808.

Chang TC, Robson SC, Boys RJ, Spencer JAD. Prediction of the small-for-gestational age infant: Which ultrasound measurement is best? Obstet Gynecol 1992; 80: 1030-1038.

Chard T. Hormonal control of growth in the human fetus. J Endocrinol 1989; 123: 3-9.

Chard T, Costeloe K, Leaf A. Evidence of growth retardation in neonates of apparently normal birthweight. Eur J Obstet Gynecol Reprod Biol 1992; 45: 59-62.

Charlton V, Johengen M. Fetal intravenous nutritional supplementation ameliorates the development of embolization-induced growth retardation in sheep. Pediatr Res 1987; 22: 55-61.

Chinn DH, Filly RA, Callen PW. Prediction of intrauterine growth retardation by sonographic estimation of total intrauterine volume. J Clin Ultrasound 1981; 9: 175-179.

Clapp JF, Szeto HH, Larrow R, Mann LI. Umbilical blood flow response to embolization of the uterine circulation. Am J Obstet Gynecol 1980; 138: 60-67.

Clapp JF, Szeto HH, Larrow R, Hewitt J, Mann LI. Fetal metabolic response to embolisation of the uterine circulation. Am J Obstet Gynecol 1981; 140: 446-451.

Clapp JF. Uteroplacental blood flow and fetal growth. In: Sharp F, Fraser RB, Milner RDG eds. Fetal Growth. Proceedings of the 20th Study Group of the Royal College of Obstetricians and Gynaecologists. Royal College of Obstetricians and Gynaecologists, London. 1989a: 251-261.

Clapp JF. Physiological adaptations in fetal growth retardation. In: Spencer JAD ed. Fetal Monitoring. Physiology and techniques of antenatal and intrapartum assessment.Castle House Publications, Tunbridge Wells. 1989b: 15-19.

Clapp JF, Capeless EL. Neonatal morphometrics after endurance exercise during pregnancy. Am J Obstet Gynecol 1990; 163: 1805-1811.

Clapp JF. Etiology and pathophysiology of intrauterine growth retardation. In: Divon MJ ed. Fetal Growth. Elsevier Science Publishing Co, New York. 1991: 83-97.

Clement D, Silverman R, Scott D, Hobbins JC. Comparisons of abdominal circumference measurements by real-time and B-scan techniques. J Clin Ultrasound 1981; 9: 1-3.

Cochrane WJ, Klish WJ, Wong WW, Klein PD. Total body electrical conductivity used to determine body composition in infants. Pediatr Res 1986; 20: 561-564.

Colley NV, Tremble JM, Henson G, Cole TJ. Head circumference / abdominal circumference ratio, ponderal index and fetal malnutrition. Should head circumference / abdominal circumference ratio be abandoned? Br J Obstet Gynaecol 1991; 98: 524-527.

Crane JP, Kapta MM. Comparative newborn anthropometric data in symmetric versus asymmetric intrauterine growth retardation. Am J Obstet Gynecol 1980; 138: 518-522.

Danielian PJ, Allman ACJ, Steer PJ. Is obstetric and neonatal outcome worse in fetuses who fail to reach their own growth potential? Br J Obstet Gynaecol 1992; 99: 452-454.

Dauncey MJ, Gandy G, Gairdner D. Assessment of total body fat in infancy from skinfold thickness measurements.Arch Dis Child 1977; 52: 223-227.

Davies DP. The infants self-regulation of food intake and weight gain (letter). Lancet 1975; ii: 366-367.

Davies DP, Platts P, Pritchard JM, Wilkinson PW. Nutritional status of light-for-date infants at birth and its influence on early postnatal growth. Arch Dis Child 1979; 54: 703-706.

Davies DP. Growth of small-for dates babies. Early Hum Dev 1981; 5: 95-105.

Davies JA, Lee A, Spencer JAD. Variability of continuous-wave doppler flow velocity

waveform indices from the umbilical artery. Obstet Gynecol 1990; 76: 366-369.

Davies JA, Gallivan S, Spencer JAD. Randomised controlled trial of Doppler ultrasound screening of placental perfusion during pregnancy. Lancet 1992; 340: 1299-1303.

Davies PSW, Lucas A. The prediction of total body fatness in early infancy. Early Hum Dev 1990; 24: 193-198.

Day NE. Fitting curves to longitudinal data. Biometrics 1966; 22: 276-291

Degani S, Paltiely Y, Lewinsky R, Shapiro I, Sharif M. Fetal blood flow velocity waveforms in pregnancies complicated by intrauterine growth retardation. Isr J Med Sci 1990; 26: 250-254.

Dempster J, Mires GJ, Patel N, Taylor DJ. Umbilical artery velocity waveforms: poor association with small-for-gestational-age babies. Br J Obstet Gynecol 1989; 96: 692-696.

Deter RL, Harrist RB, Hadlock FP, Poindexter AN. Longitudinal studies of fetal growth with use of dynamic image ultrasonography. Am J Obstet Gynecol 1982a; 143: 545-554.

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 1982b; 10: 365-372.

Deter RL, Harrist RB, Hadlock FP, Cortissoz CM, Batten GW. Longitudinal studies of fetal growth using volume parameters determined with ultrasound. J Clin Ultrasound 1984; 12: 313-324.

Deter RL. Detection of fetal growth abnormalities. In: Deter RL ed. Ouantitative obstetrical ultrasonography. John Wiley and Sons, Chichester. 1986: 123-140.

Deter RL, Rossavik IK. A simplified method for determining individual growth curve standards. Obstet Gynecol 1987; 70: 801-806.

Deter RL, Rossavik IK, Harrist RB. Development of individual growth curve standards for estimated fetal weight: I. Weight estimation procedure. J Clin Ultrasound 1988; 16: 215-225.

Deter RL, Hill RM, Tennyson LM. Predicting the birth characteristics of normal fetuses 14 weeks before delivery. J Clin Ultrasound 1989a; 17: 89-93.

Deter RL, Rossavik IK, Carpenter RJ. Development of individual growth standards for estimated fetal weight: II. Weight prediction during the third trimester and at birth. J Clin Ultrasound 1989b; 17: 83-88.

Deter RL, Harrist RB, Hill RM. Neonatal growth assessment score: A new approach to the detection of intrauterine growth retardation in the newborn. Am J Obstet Gynecol 1990; 162: 1030-1036.

Deter RL, Harrist RB. Growth standards for anatomic measurements and growth rates derived from longitudinal studies of normal fetal growth. J Clin Ultrasound 1992; 20: 381-388.

Devoe LD, Gardner P, Dear C, Castillo RA. The diagnostic values of concurrent nonstress testing, amniotic fluid measurement, and Doppler velocimetry in screening a general high-risk population. Am J Obstet Gynecol 1990; 163: 1040-1048.

Dhanireddy R, Hamosh M, Siva KN. Post-heparin lipolytic activity and intralipid clearance in very low birthweight infants. J Pediatr 1981; 98: 617-622.

Dijxhoorn MJ, Visser GHA, Fidler VJ, Touwen BCL, Huisjes HJ. Apgar score, meconium and acidaemia at birth in relation to neonatal neurological morbidity in term infants. Br J Obstet Gynaecol 1986; 93: 217-222.

Dijxhoorn MJ, Visser GHA, Touwen BCL,Huisjens HS. Apgar score, meconium and acidaemia at birth in small-for-gestational age infants born at term, and their relation to neonatal neurological morbidity. Br J Obstet Gynecol 1987; 94: 873-879.

Divon MY, Chamberlain PF, Sipos L, Manning FA, Platt LD. Identification of the small for gestational age fetus with the use of gestational age-independent indices of fetal growth. Am J Obstet Gynecol 1986; 155: 1197-1201.

Divon MY, Guidetti DA, Braverman JJ, Oberlander E, Langer O, Merkatz IR. Intrauterine growth retardation - a prospective study of the diagnostic value of real-time sonography combined with umbilical artery flow velocimetry. Obstet Gynecol 1988; 72: 611-614.

Divon MY, Hsu HW. Maternal and fetal blood flow velocity waveforms in intrauterine growth retardation. Clin Obstet Gynecol 1992; 35: 156-171.

Dobbing J, Sands J. Timing of neuroblast multiplication in developing human brain.

Nature 1970; 226: 639-640.

Dobbing J, Sands J. Quantitative growth and development of the human brain. Arch Dis Child 1973; 48: 757-767.

Dobson P, Abell A, Beischer N. Antenatal pregnancy complications and fetal growth complications. Aust N Z J Obstet Gynaecol 1982; 22: 203-205.

Donald I, Brown C. Demonstration of tissue interfaces within the body by ultrasonic echo sounding. Br J Radiol 1961; 34: 539-546.

Drillien CM. The small-for-dated infant: Aetiology and prognosis. Pediatr Clin North Am 1970; 17: 9-24.

Duff GB. The realities of screening for the small for dates fetus using ultrasound measurement. Aust NZ J Obstet Gynaecol 1986; 26: 102-105.

Dyson M. A review of recent experimental evidence of the effects of diagnostic ultrasound on tissue. In: Evans JA ed. Physics in medical ultrasound. Institute of Physical Sciences in Medicine. 1986: 1-11.

Economides DL, Crook D, Nicolaides KH. Investigation of hypertriglyceridaemia in small for gestational age fetuses. Fetal Ther 1988; 3: 165-172.

Economides DL, Nicolaides KH, Gahl WA, Bernadini I, Evans MI. Plasma amino acids in appropriate- and small-for-gestational age fetuses. Am J Obstet Gynecol 1989a; 161: 1219-1227.

Economides DL, Proudler A, Nicolaides KH. Plasma insulin in appropriate- and small-for-gestational age fetuses. Am J Obstet Gynecol 1989b; 160: 1091-1094.

Eik-Nes SH, Brubakk AO, Ulstein M. Measurement of human fetal blood flow. Br Med J 1980; 280: 283-284.

Eik-Nes SH, Okland O, Aure JC, Ulstein M. Ultrasound screening in pregnacy: A randomised controlled trial. Lancet 1984a; i: 1347.

Eik-Nes SH, Marsal K, Kristofferson K.Methodology and basic problems related to blood flow studies in the human fetus. Ultrasound Med Biol 1984b; 10: 329-337.

Eldridge MW, Alverson DC, Howard EA, Berman W. In: Berman W ed. Pulsed Doppler ultrasound: Principles and Instrumentation. Futura, New York. 1983: 5-40.

Eskes TKAB, Jongsma HW, Houx PCW. Percentiles for gas values in human umbilical cord blood. Europ J Obstet Gynec Reprod Biol 1983; 14: 341-346.

European Federation of Societies for Ultrasound in Medicine and Biology. Clinical Safety Statement. Newsletter. June, 1992.

Evans T, Farrant P, Gowland M, McNay M. Clinical application of ultrasonic fetal measurement. Report of the Fetal Measurement Working Party. British Medical Ultrasound Society, London. 1990: 1-13.

Evans T. The Physics of Doppler Ultrasound. In Pearce JM ed. Doppler Ultrasound in Perinatal Medicine. Oxford University Press, Oxford. 1992: 3-16.

Excler JL, Sann L, Lasne Y. Anthropometric assessment of nutritional status in newborn infants: discriminative value of mid-arm circumference and skinfold thickness. Early Hum Dev 1985; 11: 169-172.

Fancourt R, Campbell S, Harvey D, Norman AP. Follow-up study of small-for-dates babies. Br Med J 1976; 1: 1435-1437.

Fay RA, Dey PL. Saadie CM, Buhl JA, Gebski VJ. Ponderal index: A better definition of the 'at risk' group with intrauterine problems than birth-weight for gestational age in term infants. Aust NZ J Obstet Gynaecol 1991a; 31: 17-19.

Fay TN, Patodi M, Crocker SG. Antenatal prediction of 'small-for-dates' babies. What proportion are growth retarded? J Obstet Gynaecol 1991b; 11: 237-240.

Fee SC, Malee K, Deddish R, Minogue JP, Socol ML. Severe acidosis and subsequent neurological status. Am J Obstet Gynecol 1990; 162: 802-806.

Ferrazi E, Nicoloni U, Kustermann A, Pardi G. Routine obstetric ultrasound: Effectiveness of cross-sectional screening for fetal growth retardation. J Clin Ultrasound 1986; 14: 17-22.

Ferrazzi E, Vegni C, Bellotti M, Borboni A, Peruta SD, Barbera A. Role of umbilical Doppler velocimetry in the biophysical assessment of the growth retarded fetus. J Ultrasound Med 1991; 10: 309-315.

Fescina RH, Ucieda FJ. Reliability of fetal anthropometry by ultrasound. J Perinatal Med 1980; 8: 93-99.

Fescina RH, Ucieda FJ, Cordano MC, Nieto F, Fenzer SM, Lopez R. Ultrasonic patterns of intrauterine fetal growth in a Latin American country. Early Hum Dev 1982; 6: 239-242.

Fescina RH, Martell M, Martinez G, Lastra L, Scharcz R. Small for dates: Evaluation of different diagnostic methods. Acta Obstet Gynecol Scand 1987; 66: 221-226.

Fitzgerald DE, Drumm JE. Non-invasive measurements of human fetal circulation using ultrasound: a new method. Br Med J 1977; 2: 1450-1451.

Fleischer A, Schulman H, Farmakides G, Bracero L, Blattner P, Randolph G. Umbilical artery velocity waveforms and intrauterine growth retardation. Am J Obstet Gynecol 1985; 151: 502-505.

Foley TP, de Philip R, Perricelli A, Miller A. Low Somatomedin activity in cord serum from infants with intrauterine growth retardation. J Pediatr 1980; 96: 605-610.

Forbes GB. Report of the Committee on Nutrition: Methods for determining composition of the human body, with a note on the effect of diet on body composition. Pediatrics 1962; 29: 477-494.

Fowden AL, Comline RS. The effects of pancreatectomy on the sheep fetus in utero. Quart J Exp Physiol 1984; 69: 319-330.

Fowden AL. The role of insulin in prenatal growth. J Develop Physiol 1989; 12: 173-182.

Fowden AL. The role of insulin in fetal growth. Early Hum Dev 1992; 29: 177-181.

Freeman JM, Nelson K. Intrapartum asphxia and cerebral palsy. Pediatrics 1988; 82: 240-249.

Frisancho AR, Compton A, Matos J. Ineffectiveness of body mass indices for the evaluation of neonate nutritional status. J Pediatr 1986; 108: 993-995.

Gairdner D, Pearson J. Revised Gairdner-Pearson growth charts. Arch Dis Child 1985; 60: 1202.

Galbraith RS, Karchmar EJ, Piercy WN, Low JA. The clinical prediction of intrauterine growth retardation. Am J Obstet Gynecol 1979; 133: 281-286.

Gampel B. The relation of skinfold thickness in the neonate to sex, length of gestation,size at birth and maternal skinfold. Hum Biol 1965; 37: 29-37.

Garden AS, Weindling AM, Griffiths RD, Martin PA. Assessment of fetal growth using fast-scan magnetic resonance imaging. J Maternal Fetal Invest 1991; 1: 7-13.

Gardosi J, Chang A, Kalyan B, Sahota D, Symonds EM. Customised antenatal growth charts. Lancet 1992; 339: 283-287.

Gart J. Point and interval estimation of the common odds ratio in the combination of 2 X 2 tables with fixed marginals. Biometrika 1970; 57: 471-475.

Gaziano E, Knox E, Wager GP, Bendel RP, Boyce DJ, Olson J. The predictability of the small-for-gestational age infant by real-time ultrasound-derived measurements combined with pulsed Doppler umbilical artery velocimetry. Am J Obstet Gynecol 1988; 158: 1431-1439.

Geirsson RT, Patel NB, Christie AD. Efficacy of intrauterine volume, fetal abdominal area and biparietal diameter measurements with ultrasound in screening for small-for-dates babies. Br J Obstet Gynaecol 1985a; 92: 929-935.

Geirsson RT, Patel NB, Christie AD. Intrauterine volume, fetal abdominal area and biparietal diameter measurements with ultrasound in the prediction of small-for-dates babies in a high-risk population. Br J Obstet Gynaecol 1985b; 92: 936-940.

Geirsson RT. Ultrasound instead of the last menstrual period as the basis of gestational age assignment. Ultrasound Obstet Gynecol 1991; 1: 212-219.

Georgieff MK, Sasanow SR, Mammel MC, Pereira GR. Mid-arm circumference / head circumference ratios for identification of symptomatic LGA, AGA and SGA newborn infants. J Pediatr 1986; 109: 316-321.

Georgieff MK, Sasanow SR, Chockalingam UM, Pereira GR. A comparison of the mid-arm circumference / head circumference ratio and ponderal index for the evaluation of newborn infants after abnormal intrauterine growth. Acta Paediatr Scand 1988; 77: 214-219.

Gerhard I, Vollmar B, Runnebaum B, Klinga K, Haller U, Kubli F. Weight percentile at birth: Prediction by endocrinological and sonographic measurements. Eur J Obstet Gynecol Reprod Biol 1987; 26: 313-328.

Gibson HM. Plasma volume and glomerular filtration rate in pregnancy and their relation to differences in fetal growth. Br J Obstet Gynaecol 1973; 80: 1067-1074.

Giles WB, Trudinger BJ, Baird PJ. Fetal umbilical artery flow velocity waveforms and placental resistance: a pathological correlation. Br J Obstet Gynaecol 1985; 92: 31-38.

Gill RW, Kossoff G. Pulsed Doppler ultrasound combined with B-mode imaging for blood flow measurements. Contrib Obstet Gynecol 1979; 6: 139-141.

Gilstrap III LC, Hauth JC, Hankins GDV, Beck AW. Second stage fetal heart rate abnormalities and type of neonatal acidaemia. Obstet Gynecol 1987; 70: 191-195.

Gluckman PD, Johnson-Barrett JJ, Butler JH, Edgar BW, Gunn TR. Studies of insulin-like growth factor I and II by specific radioligand assays in umbilical cord blood. Clin Endocrinol 1983; 19: 405-413.

Gluckman PD, Butler JH, Comline R, Fowden A. The effects of pancreatectomy on the plasma concentrations of insulin-like growth factors 1 and 2 in the sheep fetus. J Dev Physiol 1987; 9: 79-88.

Gluckman PD. Fetal growth : An endocrine perspective. Acta Paediatr Scand (Suppl) 1989; 349: 21-25.

Gohari P, Berkowitz RL, Hobbins JC. Prediction of intrauterine growth retardation by determination of total intrauterine volume. Am J Obstet Gynecol 1977; 127: 255-260.

Goldberg SJ, Allen HD, Marx GR, Donnerstein RL. In: Doppler echocardiography. Lea and Fabiger, Philadelphia. 1988: 1-186.

Goldenberg RL, Cutter GR, Hoffman HJ, Foster JM et al. Intrauterine growth retardation: Standards for diagnosis. Am J Obstet Gynecol 1989; 161: 271-277.

Goldenberg RL, Hoffman HJ, Cliver SP, Cutter GR, Nelson KG, Cooper RL. The influence of previous low birth weight on birth weight, gestational age, and anthropometric measurements in the current pregnancy. Obstet Gynecol 1992; 79: 276-

280.

Gosling RG, King DH. Ultrasound angiology. In: Harcus A, Adamson L eds. Ultrasonic angiology in arteries and veins. Churchill Livingstone, Edinburgh. 1975: 61-98.

Gramellini D, Folli MC, Raboni S, Vadora E, Merialdi A. Cerebral-umbilical Doppler ratio as a predictor of adverse perinatal outcome. Obstet Gynecol 1992; 79: 416-420.

Grant A, Mohide P. Screening and diagnostic tests in antenatal care. In: Enkin M, Chalmers I eds. Clinics in Developmental Medicine Nos. 81 / 82. Effectiveness and Satisfaction in Antenatal Care. W. Heinemann, London. 1982; 22-59.

Gruenwald P. Chronic fetal distress and placental insufficiency. Biol Neonat 1963; 5: 215-265.

Gudmundsson S, Marsal K. Receiver operating characteristic curves of fetal, umbilical and uteroplacental blood velocity waveforms as predictors of fetal outcome. Zent bl Gynakol 1991a; 113: 601-607.

Gudmundsson S, Marsal K. Blood velocity waveforms in the fetal aorta and umbilical artery as predictors of fetal outcome: A comparison. Am J Perinatol 1991b; 8: 1-6.

Gustafson A, Kjellmer I, Olegard R. Nutrition in low birthweight infants. Acta Pediatr Scand 1972; 61: 149-158.

Haas J, Balcazor H, Caulfield L. Variation in early neonatal mortality for different types of fetal growth retardation. Am J Phys Anthropol 1987; 73: 467-473.

Hackett GA, Campbell S, Gamsu H, Cohen-Overbeek T, Pearce JMF. Doppler studies in the growth-retarded fetus and prediction of neonatal necrotising enterocolitis, haemorrhage and neonatal mortality. Br Med J 1987; 294: 13-16.

Hadlock FP, Deter RL, Harrist RB, Park SK. Fetal abdominal circumference as a predictor of menstrual age. Am J Roentgen 1982a; 139: 367-370.

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

Hadlock FP, Harrist RB, Carpenter RJ, Deter RL, Park SK. Sonographic estimation of

fetal weight. Radiology 1984; 150: 535-540.

Hadlock FP, Harrist RB, Sharman RS, Deter RL, Park SK. Estimation of fetal weight with the use of head, body and femur measurements - a prospective study. Am J Obstet Gynecol 1985; 151: 333-337.

Hales CN, Barker DJP, Clark PMS, Cox LJ, Fall C, Osmond C, Winter PD. Fetal and infant growth and impaired glucose tolerance at age 64. Br Med J 1991; 303: 1019-1022.

Hamilton PA, Costello AM deL, Stewart A, Baudin J, Bradford B, Reynolds EOR. Ultrasound brain scanning in very preterm infants and outcome at 18 months and 4 years of age. Pediatr Res 1986; 20: 1040.

Han VKM. Genetic mechanisms of regulation of fetal growth. In: Sharp F, Fraser RB, Milner RDG eds. Fetal Growth. Proceedings of the 20th Study Group of the Royal College of Obstetricians and Gynaecologists. Royal College of Obstetricians and Gynaecologists, London. 1989: 77-81.

Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982; 143: 29-36.

Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983; 148: 839-843.

Hanley JA. Receiver operating characteristic (ROC) methodology: The state of the art. Crit Rev Diagn Imaging 1989; 29: 307-335.

Hanretty KP, Primrose MH, Neilson JP, Whittle MJ. Pregnancy screening by Doppler uteroplacental and umbilical artery waveforms. Br J Obstet Gynaecol 1989; 96: 1163-1167.

Harding JE, Jones CT, Robinson JS. Studies on experimental growth retardation in sheep. The effect of a small placenta in restricting transport to and growth of the fetus. J Dev Physiol 1985; 7: 427-442.

Hardy BJ, Drage JS, Jackson EC. The first year of life. Johns Hopkins University Press, Baltimore. 1979: 38-70.

Harvey D, Prince J, Bunton J, Parkinson C, Campbell S. Abilities of children who were small-for-dates at birth and whose growth in-utero was measured by ultrasonic

cephalometry. Pediatr Res 1976; 10: 891.

Harvey D, Prince J, Bunton J, Parkinson C, Campbell S. Abilities of children who were small-for-gestational age babies. Pediatrics 1982; 69: 296-300.

Hassan MM, Bottoms SF, Mariona FG, Syner FN, Simkowski KM, Sokol RJ. The use of clinical, biochemical and ultrasound parameters for the diagnosis of intrauterine growth retardation. Am J Perinatol 1987; 4: 191-194.

Hawdon JM, Ward Platt MP, Aynsley-Green A. Patterns of metabolic adaptation for preterm and term infants in the first neonatal week. Arch Dis Child 1992a; 67: 357-365.

Hawdon JM, Ward Platt MP, McPhail S, Cameron H, Walkinshaw SA. Prediction of impaired metabolic adaptation by antenatal Doppler studies in small for gestational age fetuses. Arch Dis Child 1992b; 67: 789-792.

Hayashi S, Sareda K, Sagawa N, Yamoda W, Kido K. Umbilical vein-artery differences of plasma amino acids in the last trimester of human pregnancy. Biol Neonate 1978; 34: 11-18.

Haymond MW, Karl IE, Pagliara AS. Increased gluconeogenic substrate in the small for gestational age infant. N Eng J Med 1974; 291: 322-328.

Hays D, Patterson RM. A comparison of fetal biometric ratios to neonatal morphometrics. J Ultrasound Med 1987; 6: 71-73.

Healy MJR. Measuring measuring errors. Statistics in Medicine 1989; 8: 893-906.

Hecher K, Spernol R, Stettner H, Szalay S. Potential for diagnosing imminent risk to appropriate- and small-for-gestational age fetuses by Doppler sonographic examination of umbilical and cerebral arterial blood flow. Ultrasound Obstet Gynecol 1992; 2: 266-271.

Heinonen K, Matilainen R, Koski H, Launila K. Intrauterine growth retardation (IUGR) in preterm infants. J Perinat Med 1985; 13: 171-178.

Hellman LM, Duffus GM, Donald I, Sunden B. Safety of diagnostic ultrasound in obstetrics. Lancet 1970; i: 1133-1135.

Hill DJ, Milner RDG. Insulin as a growth factor. Pediatr Res 1985; 19: 879-886.

Hill DJ, Han VKM. Control of cellular multiplication and differentiation. In: Sharp F, Fraser RB, Milner RDG eds. Fetal Growth. Proceedings of the 20th Study Group of the Royal College of Obstetricians and Gynaecologists. Royal College of Obstetricians and Gynaecologists, London. 1989: 83-100.

Hill LM, Guzick D, Belfar HL, Peterson C, Rivello D, Hixson J. A combined historic and sonographic score for the detection of intrauterine growth retardation. Obstet Gynecol 1989; 73: 291-296.

Hill RM, Verniaud WM, Deter RL, Tennyson LM, Rettig GM, Zion TE, Vordeman AL, Holms PG, Mc Culley LB, Hill LL. The effect of intra-uterine malnutrition on the term infant: a 14 year progressive study. Acta Paediatr Scand 1984; 73: 482-487.

Himes JH, Roche AF, Siervogel RM. Compressibility of skinfolds and the measurement of subcutaneous fatness. Am J Clin Nutr 1979; 32: 1734-1740.

Hoffbauer H, Pachaly J, Arabin B, Baumann ML. Control of fetal development with multiple ultrasonic body measures. Contrib Gynecol Obstet 1979; 6: 147-156.

Hoffman AL, Hjortdal JO, Secher NJ, Weile B. The relationship between Apgar score, umbilical artery pH and operative delivery for fetal distress in 2778 infants born at term. Eur J Obstet Gynecol Reprod Biol 1991; 38: 97-101.

Hughey MJ. Routine ultrasound for detection and management of small-for-gestational age fetus. Obstet Gynecol 1984; 64: 101-107.

Hull J, Dodd KL. What is birth asphyxia? Br J Obstet Gynaecol 1991; 98: 953-955.

Jacobson S-L, Imhof R, Manning N, Mannion V, Little D, Rey E, Redman C. The value of Doppler assessment of the uteroplacental circulation in predicting preeclampsia or intrauterine growth retardation. Am J Obstet Gynecol 1990; 162: 110-114.

Jarai I, Mestyan J, Schultz K, Lazar A, Halasz M, Kraussy I. Body size and neonatal hypoglycaemia in intrauterine growth retardation. Early Hum Dev 1977; 1: 25-32.

Jeanty P, Cousaert E, Cantraine F. Normal growth of the abdominal perimeter. Am J Perinatol 1984a; 1: 129-135.

Jeanty P, Cantraine F, Romero R, Cousaert E, Hobbins JC. A longitudinal study of fetal weight growth. J Ultrasound Med 1984b; 3: 321-328.

Jones CT. Reprogramming of metabolic development by restriction of fetal growth. Biochem Soc Trans 1985; 13: 89-91.

Jones CT, Gu W, Harding JE et al. Studies on the growth of the fetal sheep. Effects of surgical reduction in placental size, or experimental manipulation of uterine blood flow on plasma sulphation activity and on the concentration of insulin-like growth factors I and II. J Dev Physiol 1988; 10: 179-189.

Jones RAG, Robertson NRC. Problems of the small-for-dates baby. Clin Obstet Gynecol 1984; 11: 499-524.

Jones RAG, Robertson NRC. Small-for-dates babies: are they really a problem? Arch Dis Child 1986; 61: 877-880.

Jurgens- van der Zee AD, Bierman-van Eendenburg MEC, Fidler VJ,Olinga AA, Visch JH, Touwen BCL, Huisjes HJ. Preterm birth, growth retardation and acidaemia in relation to neurological abnormality of the newborn. Early Hum Dev 1979; 3: 141-154.

Kahn HA, Sempos CT. Relative risk and odds ratio. In: Statistical methods in Epidemiology. Monographs in Epidemiology and Biostatistics No.12. Oxford University Press. 1989: 45-71.

Kaminski M, Papiernik E. Multifactorial study of the risk of prematurity at 32 weeks of gestation. II. A comparison between an empirical prediction and a discriminant analysis. J Perinat Med 1974; 2: 37-44.

Kay HH, Carol BB, Dahmus M, Killam AP. Sonographic measurements with umbilical artery Doppler analysis in suspected intrauterine growth retardation. J Reprod Med 1991: 36: 65-68.

Kazzi GM, Gross TL, Sokol RJ. Fetal biparietal diameter and placental grade: predictors of intrauterine growth retardation. Obstet Gynecol 1983a; 62: 755-759.

Kazzi GM, Gross TL, Sokol RJ, Kazzi NJ.Detection of intrauterine growth retardation: a new use for sonographic placental grading. Am J Obstet Gynecol 1983b; 145: 733-737.

Kiely M, Kiely JL. Intrauterine growth retardation: 1988 US data compared to previous standards. Am J Obstet Gynecol 1992; 166: 420 (abstract).

Koh THHG, Aynsley-Green A, Tarbit M, Eyre JA. Neuronal dysfunction during hypoglycaemia. Arch Dis Child 1988; 63: 1353-1358.

Koops BL, Morgan LJ, Battaglia FC. Neonatal mortality risk in relation to birth weight and gestational age. J Pediatr 1982; 101: 969.

Kramer MS, McLean FH, Olivier M, Willis DM, Usher RH. Body proportionality and head and length "sparing" in growth-retarded neonates: A critical reappraisal. Pediatrics 1989; 84: 717-723.

Kremkau FW. Biological effects and possible hazards. Clin Obstet Gynecol 1983; 10: 395-405.

Laird NM, Ware JH. Random-effect models for longitudinal data. Biometrics 1982; 38: 963-974.

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

Larsen T, Larsen JF, Petersen S, Greisen G. Detection of small-for-gestational age fetuses by ultrasound screening in a high risk population: a randomized controlled study. Br J Obstet Gynaecol 1992; 99: 469-474.

Lassarre C, Hardouin S, Daffos F, Forestier F, Frankenne F, Binoux M. Serum insulin-like growth factors and insulin-like growth factor binding proteins in the human fetus: Relationship with growth in normal subjects and in subjects with intrauterine growth retardation. Pediatr Res 1991; 29: 219- 225.

Laurin J, Marsal K, Persson P-H,Lingman G. UItrasound measurement of fetal blood flow in predicting fetal outcome. Br J Obstet Gynaecol 1987; 94: 940-948.

Levene MI, Kornberg J, Wiulliams THC. The incidence and severity of post-asphyxial encephalopathy in full-term infants. Early Hum Dev 1985; 11: 21-26.

Lin CC, Moawad AH, Rosenow PJ, River P. Acid-base characteristics of fetuses with intrauterine growth retardation during labour and delivery. Am J Obstet Gynecol 1980; 137: 553-559.

Lin CC, Moawad AH, River PH, Blix P, Abraham M, Rubenstein AH. Amniotic fluid C-peptide as an index for intrauterine fetal growth. Am J Obstet Gynecol 1981; 139: 390-

396.

Lin CC, Su SJ, River P. Comparison of associated high-risk factors and perinatal outcome between symmetric and assymetric fetal intrauterine growth retardation.
Am J Obstet Gynecol 1991; 164: 1535-1542.

Low JA. The current status of maternal and fetal blood flow velocimetry. Am J Obstet Gynecol 1991; 164: 1049-1063.

Lowery CL, Henson BV, Wan J, Brumfield CG. A comparison between umbilical artery velocimetry and standard antepartum surveillance in hospitalised high-risk patients. Am J Obstet Gynecol 1990; 162: 710-714.

Lubchenco L, Hansman C, Dressler M, Boyd E. Intrauterine growth as estimated from live born weight data. Pediatrics 1963; 32: 793-800.

Lubchenco LO, Bard H. Incidence of hypoglycaemia in newborn infants classified by birthweight and gestational age. Pediatrics 1971; 47:831-838.

Lucas A, Morley R, Cole T. Adverse neurodevelopmental outcome of moderate neonatal hypoglycaemia. Br Med J 1988; 297: 1304-1308.

Lucas M, Mullarkey M, Abdulla U. Study of chromosomes in the newborn after ultrasound fetal heart monitoring in labour. Br Med J 1972; 3: 795-796.

MacDonald D, Grant A, Sheridan-Pereira M, Boylan P, Chalmers I. The Dublin randomized trial of intrapartum fetal heart rate monitoring. Am J Obstet Gynecol 1985; 152: 524-539.

MacIntosh IJC, Davey DA. Chromosome aberrations induced by an ultrasonic fetal pulse detector. Br Med J 1970; 4: 92-93.

MacIntosh IJC, Davey DA. Relationships between intensity of ultrasound and induction of chromosome aberrations. Br J Radiol 1972; 45: 340-342.

Makanson DO, Oh W. Hyperviscosity in small-for-gestational age infants. Biol Neonate 1980; 37: 109-112.

Malcolm G, Ellwood D, Devonald K, Beilby R, Henderson-Smart D. Absent or reversed end diastolic flow velocity in the umbilical artery and necrotising enterocolitis. Arch Dis

Child 1991; 66: 805-807.

Manning FA, Hill LM, Platt LD. Qualitative amniotic fluid volume determination by ultrasound: antepartum detection of intrauterine growth retardation. Am J Obstet Gynecol 1981; 139: 254-258.

Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst 1959; 22: 719-748.

Marsal K. Fetal breathing movements in man - Characteristics and clinical significance. Obstet Gynecol 1978; 52: 394-401.

Marsal K. Fetal and placental blood flow. In: Sharp F, Fraser RB, Milner RDG eds. Fetal Growth. Proceedings of the 20th Study Group of the Royal College of Obstetricians and Gynaecologists. Royal College of Obstetricians and Gynaecologists, London. 1989: 297-308.

Martin CB. The anatomy and circulation of the placenta. In Barns AC ed. Intra-uterine development. Lea and Fabiger, Philadelphia. 1968: 35-67.

Maulik D, Yarlagadda P, Youngblood JP, Ciston P. The diagnostic efficacy of the umbilical arterial systolic / diastolic ratio as a screening tool: a prospective blinded study. Am J Obstet Gynecol 1990; 162: 1518-1525.

McCallum WD, Williams CS, Napel S, Diagle RE. Fetal blood velocity waveforms. Am J Obstet Gynecol 1978; 132: 425-429.

McCowan LM, Mullen BM, Ritchie JWK. Umbilical artery flow velocity waveforms and the placental vascular bed. Am J Obstet Gynecol 1989; 157: 900-902.

McGowan A, Jordan M, MacGregor J. Skinfold thickness in neonates. Biol Neonate 1975; 25: 66-84.

Meadows NJ, Till J, Leaf A, Hughes E, Jani B, Larcher V. Screening for intrauterine growth retardation using ratio of mid-arm circumference to occipitofrontal circumference. Br Med J 1986; 292: 1039-1040.

Meire HB, Farrant P. Ultrasound demonstration of an unusual fetal growth pattern in Indians. Br J Obstet Gynaecol 1981; 88: 260-263.

Mellor DJ. Nutritional and placental determinants of fetal growth rate in sheep and consequences for the newborn lamb. Br Vet J 1983; 139: 307-324.

Mestyan J, Soltesz G, Schultz K, Horvath M. Hyperaminoacidaemia due to the accumulation of gluconeogenic amino acid precursors in hypoglycaemic small for gestational age infants. J Pediatr 1975; 87: 409-414.

Miller HC, Hassanein K. Diagnosis of impaired fetal growth in newborn infants. Pediatrics 1971; 48: 511-522.

Miller HC, Merritt TA. Fetal growth in humans. Chicago: Year book. 1979; 31-57.

Miller HC. Intrauterine growth retardation; an unmet challenge. Am J Dis Child 1981; 135: 944.

Miller JM, Kissling GA, Brown HL, Gabert HA. The relationship of placental grade to fetal size and growth at term. Am J Perinatol 1988; 5: 19-21.

Miller JM, Gabert HA. Comparison of dynamic and pulsed Doppler ultrasonography for the diagnosis of the small-for-gestational age fetus. Am J Obstet Gynecol 1992; 166: 1820-1826.

Milner RDG. Mechanisms of overgrowth. In: Sharp F, Fraser RB, Milner RDG eds. Fetal Growth. Proceedings of the 20th Study Group of the Royal College of Obstetricians and Gynaecologists. Royal College of Obstetricians and Gynaecologists, London. 1989: 139-148.

Morrow RJ, Adamson SL, Bull SB, Ritchie JWK. The effect of placental embolization on the umbilical artery waveform in sheep. Am J Obstet Gynecol 1989; 161: 1055-1060.

Morrow RJ, Adamson SL, Bull SB, Ritchie JWK. Acute hypoxaemia does not affect umbilical artery waveforms in sheep. Obstet Gynecol 1990; 75: 590-593.

Mulders LGM, Wijn PFF, Jongsma HW, Hein PR. A comparative study of three indices of umbilical blood flow in relation to prediction of growth retardation. J Perinat Med 1987; 15: 3-12.

Neilson JP, Whitfield CR, Artchison T. Screening for the small for-dates fetus: A two-stage ultrasonic examination schedule. Br Med J 1980; 208: 1203-1206.

Neilson JP, Munjaja SP, Whitfield CR. Screening for small-for-dates fetuses: a controlled trial. Br Med J 1984; 289: 1179-1182.

Neligan GA, Prudham D, Seiner H. Variations in intrauterine growth. The formative years. Oxford University Press, London. 1974.

Nelson KB, Ellenberg JH. Apgar scores as predictors of chronic neurologic disability. Pediatrics 1981; 68: 36-44.

Nelson KB, Ellenberg JH. Obstetric complications as risk factors for cerebral or seizure disorders. JAMA 1984; 251: 1843-1848.

Newnham JP, Patterson LL, James IR, Diepeveen DA, Reid SE. An evaluation of the efficiency of Doppler flow velocity waveform analysis as a screening test in pregnancy. Am J Obstet Gynecol 1990; 162: 403-410.

Newnham JP, O'Dea MRA, Reid KP, Diepeveen DA. Doppler flow velocity waveform analysis in high-risk pregnancies: a randomized controlled trial. Br J Obstet Gynaecol 1991; 98: 956-963.

Nicolaides KH, Bilardo CM, Soothill PW, Campbell S. Absence of end-diastolic frequencies in the umbilical artery: a sign of fetal hypoxia and acidosis. Br Med J 1988; 297: 1026-1027.

Nicolaides KH, Economides DL, Soothill PW. Blood gases, pH and lactate in appropriate- and small-for-gestational age fetuses. Am J Obstet Gynecol 1989; 161: 996-1001.

Nicolini U, Nicolaidis P, Fisk NM, Vaughan J, Fusi L, Gleeson R, Rodeck R. Limited role of fetal blood sampling in the prediction of outcome in intrauterine growth retardation. Lancet 1990; 336: 768-772.

Nyborg WL, Steele RB. Temperature elevation in a beam of ultrasound. Ultrsound Med Biol 1983; 9: 611-620.

Oakley JR, Parsons RJ. Skinfold thickness as an indicator of neonatal hypoglycaemia in infants with birthweights over 2500 grams. Dev Med Child Neurol 1977a; 19: 585-588.

Oakley JR, Parsons RJ, Whitelaw AGL. Standards for skinfold thickness in British newborn infants. Arch Dis Child 1977b; 52: 287-290.

Ott WJ, Doyle S. Ultrasonic diagnosis of altered fetal growth by use of a normal ultrasonic fetal weight curve.Obstet Gynecol 1984; 63: 201-204.

Ott WJ. Fetal femur length, neonatal crown-heel length, and screening for intrauterine growth retardation.Obstet Gynecol 1985; 65: 460-464.

Ott WJ. The diagnosis of altered fetal growth. Obstet Gynecol Clin North Am 1988; 15: 237-263.

Ott WJ. Defining altered fetal growth by second-trimester sonography. Obstet Gynecol 1990a; 75: 1053-1059.

Ott WJ. Comparison of dynamic image and pulsed doppler ultrasonography for the diagnosis of intrauterine growth retardation. J Clin Ultrasound 1990b; 18: 3-7.

Ounsted MK, Moar VA, Scott A. Small-for-dates babies at the age of four years: health, handicap and developmental status. Early Hum Dev 1983; 8: 243-258.

Owens JA, Falconer J, Robinson JS. Effect of restriction of placental growth on umbilical and uterine blood flows. Am J Physiol 1986; 250: 427-434.

Owens JA, Falconer J, Robinson JS. Effect of restriction of placental growth on oxygen delivery to and consumption by the pregnant uterus and fetus. J Dev Physiol 1987a; 9: 137-150.

Owens JA, Falconer J, Robinson JS. Effect of restriction of placental growth on fetal and uteroplacental metabolism. J Dev Physiol 1987b; 9: 225-238.

Palo P, Erkkola R, Piiroinen O, Ruotsalainen P. Accuracy of ultrasonic fetal weight estimation and detection of small for gestational age fetuses. Am J Perinatol 1989; 6: 400-404.

Pardi G, Buscaglia M, Ferrazi E et al. Cord sampling for the evaluation of oxygenation and acid-base balance in growth-retarded human fetuses. Am J Obstet Gynecol 1987; 157: 1221-1228.

Parer JT, Livingston EG. What is fetal distress? Am J Obstet Gynecol 1990; 1662: 1421-1427.

Parizkova J. Total body fat and skinfold thickness in children. Metabolism 1961; 10: 794-807.

Parkinson CE, Wallis S, Harvey D. School achievement and behaviour of children who were small-for-dates a birth. Develop Med Child Neurol 1981; 23: 41-50.

Patterson RM, Prihoda TJ, Gibbs CE, Wood RC. Analysis of birth weight percentile as a predictor of perinatal outcome. Obstet Gynecol 1986; 68: 459-463.

Patterson RM, Pouliot RN. Neonatal morphometrics and perinatal outcome: Who is growth retarded? Am J Obstet Gynecol 1987a; 157: 691-693.

Patterson RM, Prihoda TJ, Pouliot MR. Sonographic amniotic fluid measurement and fetal growth retardation: a reappraisal. Am J Obstet Gynecol 1987b; 157: 406-410.

Pearce JM, Campbell S, Cohen-Overbeek T, Hackett G, Hernadez J, Royston JP. Reference ranges and sources of variation for indices of pulsed Doppler flow velocity waveforms from the uteroplacental and fetal circulation. Br J Obstet Gynaecol 1988; 95: 248-256.

Peeters LL, Sheldon RE, Jones MD Jr, Makowski EL, Meschia G. Blood flow to fetal organs as a function of arterial oxygen content. Am J Obstet Gynecol 1979; 135: 637-642.

Persson P-H, Weldner B-M. Intrauterine weight curves obtained by ultrasound. Acta Obstet Gynecol Scand 1986; 65: 169-173.

Persson P-H. Fetal growth curves. In: Sharp F, Fraser RB, Milner RDG eds. Fetal Growth. Proceedings of the 20th Study Group of the Royal College of Obstetricians and Gynaecologists. Royal College of Obstetricians and Gynaecologists, London. 1989: 13-25.

Petersen S, Gotfredsen A, Knudsen FU. Lean body mass in small-for-gestational age and appropriate-for-gestational age infants. J Pediatr 1988; 113: 886-889.

Philipson EH, Sokol RJ, Williams T. Oligohydramnios: clinical associations and predictive value for intrauterine growth retardation. Am J Obstet Gynecol 1983; 146: 271-276.

Pielet PW, Sabaggha RE, Scott N, MacGregor DO, Tamura RK, Feigenbaum SL. Ultrasonic prediction of birth weight in preterm fetuses: Which formula is best? Am J

Obstet Gynecol 1987; 157: 1411-1414.

Pildes R, Forkes AE, O' Connor SM, Cornblath M. The incidence of neonatal hypoglycaemia - A completed survey. J Pediatr 1967; 70: 76-82.

Portman RJ, Carter BS, Gaylord MS, Murphy MG, Thieme RE, Merenstein GB. Predicting neonatal morbidity after perinatal asphyxia: a scoring system. Am J Obstet Gynecol 1990; 162: 174-182.

Quaranta P. Currell R, Redman CWG, Robinson JS. Prediction of small-for-dates infants by measurements of symphysial-fundal height. Br J Obstet Gynaecol 1981; 88: 115-119.

Rankin JHG, McLaughlin MK. The regulation of placental blood flow. J Dev Physiol 1979; 1: 3-30.

Rantakillio P. A 14 year follow-up of children with normal and abnormal birth weight for their gestational age. A population study. Acta Pediatr Scand 1985; 74: 62-69.

Reuwer PJHM, Sijmons EA, Rietman GW, Van Tiel MVM, Bruinse HW. Intrauterine growth retardation: prediction of perinatal distress by Doppler ultrasound. Lancet 1987; 2: 415-418.

Richardson DK, Schwartz JS, Weinbaum PJ, Gabbe SG. Diagnostic tests in obstetrics: A method for improved evaluation. Am J Obstet Gynecol 1985; 152: 613-618.

Robertson WB, Brosens I, Dixon HG. Uteroplacental vascular pathology. Eur J Obstet Gynecol Reprod Biol 1975; 5: 47-65.

Robertson WB, Brosens I, Dixon HG. Maternal blood supply in growth retardation. In: van Assche FA and Robertson WA eds. Fetal growth retardation. Churchill Livingstone, Edinburgh. 1981: 126-138.

Robinson JS, Falconer J, Owens JA. Intrauterine growth retardation: Clinical and experimental. Acta Pediatr Scand Suppl 1985; 319: 135-142.

Rochelson BL, Shulman H, Fleischer A et al. The clinical significance of Doppler umbilical artery velocimetry in the small-for-gestational age fetus. Am J Obstet Gynecol 1987; 156: 1223-1226.

Roord JJ, Raemakers LH. Quantifications of intrauterine malnutrition. Biol Neonate 1979;

33: 273-277.

Rosen KG, Murphy KW. How to assess fetal metabolic acidosis from cord samples. J Perinat Med 1991; 19: 221-226.

Rosendahl H, Kivinen S. Routine ultrasound screening for early detection of small for gestational fetuses. Obstet Gynecol 1988; 71: 518-521.

Rosendahl H, Kiniven S. Detection of small for gestational age fetuses by the combination of clinical risk factors and ultrasonography. Eur J Obstet Gynecol Reprod Biol 1991; 39: 7-11.

Rossavik IK, Deter RL. Mathematical modeling of fetal growth: I. Basic principles. J Clin Ultrasound 1984; 12: 529-533.

Rosso P, Winick M. Intrauterine growth retardation: A new approach based on the clinical and biochemical characteristics of this condition. J Perinatal Med 1974; 2: 147-160.

Ruth VJ, Raivio KO. Perinatal brain damage: predictive value of metabolic acidosis and the Apgar score. Br Med J 1988; 297: 24.

Sabbagha RE, Barton FB, Barton BA. Sonar biparietal diameter: I. Analysis of percentile of growth differences in two populations using the same methodology. Am J Obstet Gynecol 1976; 126: 479-484.

Sabbagha RE, Minogue J, Tamura RK, Hungerford SA. Estimation of birth weight by use of ultrasonographic formulas targeted to large-, appropriate- and small-for-gestational-age fetuses. Am J Obstet Gynecol 1989; 160: 854-862.

Salvesen KA, Vatten L, Jacobsen G, Eik-Nes S, Okland O, Molne K, Bakketeig L. Routine ultrasonography in utero and subsequent vision and hearing at primary school age. Ultrasound Obstet Gynecol 1992; 2: 243-247.

Sands J, Dobbing J, Gratrix CA. Cell number and cell size: organ growth and development and the control of catch-up growth in rats. Lancet 1979; ii: 503-505.

Sarmandal P, Bailey SM, Grant JM. A comparison of three methods of assessing inter-observer variation applied to ultrasonic fetal measurement in the third trimester. Br J Obstet Gynaecol 1989; 96: 1261-1265.

Sarmandal P, Grant JM. Effectiveness of ultrasound determination of fetal abdominal circumference and fetal ponderal index in the diagnosis of asymmetrical growth retardation. Br J Obstet Gynaecol 1990; 97: 118-123.

Sasonow SR, Georgieff MK, Pereira GR. Mid-arm and mid-arm/head circumference ratios: Standard curves for anthropometric assessment of neonatal nutritional status. J Pediatr 1986; 109: 311-315.

Satomura S. Ultrasonic Doppler method for the inspection of cardiac functions. J Acoust Soc Am 1957; 29: 1181-1185.

Scheidt PC, Stanley F, Bryla DA. One year follow-up of infants exposed to ultrasound in-utero. Am J Obstet Gynecol 1978; 131: 743-748.

Scott KE, Usher R. Fetal malnutrition: its incidence, causes and effects. Am J Obstet Gynecol 1966; 94: 951-963.

Secher NJ, Hansen PK, Lenstrup C, Eriksen SP. Controlled trial of ultrasound screening for light for gestational age infants in late pregnancy. Eur J Obstet Gynecol Reprod Biol 1986a; 23: 307-313.

Secher NJ, Hansen PK, Lenstrup C, Pedersen-Bjergaard L, Eriksen SP, Thomsen BL, Kielding N. Birthweight-for-gestational age charts based on early ultrasound estimation of gestational age. Br J Obstet Gynaecol 1986b; 93: 128-134.

Secher NJ, Hansen PK, Lenstrup C, Eriksen PS, Morsing G. A randomized study of fetal abdominal diameter and fetal weight estimation for the detection of light-for-gestational age infants in low-risk pregnancies. Br J Obstet Gynaecol 1987a; 94: 105-109.

Secher NJ, Hansen PK, Lenstrup C, Eriksen SP, Thomsen BL, Keiding N. On the evaluation of routine ultrasound screening in the third trimester for the detection of light for gestational age infants. Acta Obstet Gynecol Scand 1987b; 66: 463-471.

Secher NJ, Hansen PK, Thomsen BL, Kielding N. Growth retardation in preterm infants. Br J Obstet Gynaecol 1987c; 94: 115-120.

Secher NJ, Lundbye-Christensen S, Qvist I, Bagger P. An evaluation of clinical estimation of fetal weight and symphysis fundal distance for detection of SGA infants. Eur J Obstet Gynecol Reprod Biol 1990; 38: 91-96.

Selbing A, Wichman K, Ryden G. Screening for detection of intrauterine growth retardation by means of ultrasound. Acta Obstet Gynecol Scand 1984; 63: 543-548.

Shepard M. A standardized plane for biparietal diameter measurements. J Ultrasound Med 1982a; 1: 145-150.

Shepherd MJ, Richards VA, Berkowitz RL, Warsof SL, Hobbins JC. An evaluation of two equations for predicting fetal weight by ultrasound. Am J Obstet Gynecol 1982b; 142: 47-54.

Sheppard BL, Bonnar J. The ultrastructure of the arterial supply of the human placenta in pregnancy complicated by fetal growth retardation. Br J Obstet Gynaecol 1976; 83: 948-959.

Sijmons EA, Reuwer PJHM, van Beek E, Bruinse HW. The validity of screening for small-for-gestational-age and low-weight-for-length infants by Doppler ultrasound. Br J Obstet Gynaecol 1989; 96: 557-581.

Simon NV, Levisky JS, Shearer DM, O'Lear MS, Flood JT. Influence of fetal growth patterns on sonographic estimation of fetal weight. J Clin Ultrasound 1987; 15: 376-383.

Simon NV, Levisky JS, Shearer DM, Morris KC, Hansberry PA. Predictiveness of sonographic fetal weight estimation of prior probability of intrauterine growth retardation. J Clin Ultrasound 1988; 16: 285-294.

Simon NV, Deter RL, Shearer DM,Levisky JS. Prediction of normal fetal growth by the Rossavik Growth Model using two scans before 27 weeks menstrual age. J Clin Ultrasound 1989; 17: 237-243.

Simon NV, O'Connor TJ, Shearer DM. Detection of intrauterine fetal growth retardation with abdominal circumference and estimated fetal weight using cross-sectional growth curves. J Clin Ultrasound 1990a; 18: 685-690.

Simon NV, Surosky BA, Shearer DM, Levisky JS. Effectiveness of the pretest probability of intrauterine growth retardation on the predictiveness of sonographic estimated fetal weight in detecting IUGR: A clinical application of Bayes' Theorem. J Clin Ultrasound 1990b; 18: 145-153.

Sly PD, Drew JH. Massive pulmonary haemorrhage: a cause of sudden unexpected deaths in severely growth-retarded infants. Aust Paediatr 1981; 17: 32-34.

Soothill PW, Nicolaides KH, Bilardo CM, Campbell S. Relationship of fetal hypoxia in growth retardation to mean blood velocity in the fetal aorta. Lancet 1986; ii: 1118-1120.

Soothill PW, Nicolaides KH, Campbell S. Prenatal asphyxia, hyperlacticaemia, hypoglycaemia, and erythroblastosis in growth-retarded fetuses. Br Med J 1987a; 294: 1051-1053.

Soothill PW, Nicolaides KH, Rodeck CH, Campbell S. Amniotic fluid and fetal tissues are not heated by obstetric ultrasound scanning. Br J Obstet Gynaecol 1987b; 94: 675-677.

Sparks JW, Cetin I. Fetal growth: Energy and substrate requirements. In: Divon MJ ed. Fetal Growth. Elsevier Science Publishing Co, New York. 1991: 1-27.

Spellacy WN, Buhi WC, Bradley B, Holsinger KK. Maternal, fetal, and amniotic fluid levels of glucose, insulin and growth hormone. Diabetologia 1987; 30: 394-396.

Stanley KP, Fraser RB, Milner M, Bruce C. Cord insulin and C-peptide - distribution in an unselected population. Br J Obstet Gynaecol 1992; 99: 512-515.

Starfield B, Shapiro S, McCormick M, Bross D. Mortality and morbidity in infants with intrauterine growth retardation. J Pediatr 1982; 101: 978-983.

Stark CR, Orleans M, Haverkamp AD, Murphy J. Short- and long-term risks after exposure to diagnostic ultrasound in-utero. Obstet Gynecol 1984; 63: 194-200.

Starritt HC, Duck FA. A comparison of ultrasound exposure in therapy and pulsed Doppler fields. Br J Radiol 1992; 65: 557-563.

Steele SA, Pearce JM, Chamberlain GV. Doppler ultrasound of the uteroplacental circulation as a screening test for severe pre-eclampsia with intrauterine growth retardation. Europ J Obstet Gynecol Reprod Biol 1988; 18: 45-47.

Steer PJ. Intrapartum monitoring in IUGR. In: Sharp F, Fraser RB, Milner RDG eds. Fetal Growth. Proceedings of the 20th Study Group of the Royal College of Obstetricians and Gynaecologists. Royal College of Obstetricians and Gynaecologists, London. 1989: 381-387.

Steer PJ, Eigbe F, Lissauer TJ, Beard RW. Inter-relationships among cardiotocogram in

labour, meconium staining of amniotic fluid, arterial cord blood pH and Apgar scores. Obstet Gynecol 1989; 74: 715-720.

Stefos T, Deter RL, Simon NV. Effect of timing of initial scan and interval between scans on Rossavik Growth Model specification. J Clin Ultrasound 1989; 17: 319-325.

Stempel L. Eenie, meenie, minie, mo... What do the data really show? Am J Obstet Gynecol 1982; 144: 745-752.

Stewart A, Hope PL, Hamilton P, Costello AM deL, Baudin J, Bradford B, Amiel-Tison C, Reynolds EOR. Prediction in very preterm infants of satisfactory neurodevelopmental progress at 12 months. Dev Med Child Neurol 1988; 30: 53-63.

Stewart A. Fetal growth: mortality and morbidity. In: Sharp F, Fraser RB, Milner RDG eds. Fetal Growth. Proceedings of the 20th Study Group of the Royal College of Obstetricians and Gynaecologists. Royal College of Obstetricians and Gynaecologists, London. 1989: 403-412.

Sumners JE, Findley GM, Ferguson KA. Evaluation of methods for intrauterine growth using neonatal fat scores instead of birth weight as outcome measures: Fetal and neonatal measurements correlated with neonatal skinfold thickness. J Clin Ultrasound 1990; 18: 9-14.

Sykes GS, Molloy PM, Johnson P, Stirrat GM, Turnbull AC. Fetal distress and the condition of the newborn infants. Br Med J 1983; 287: 943-945.

Tamura RK, Sabbagha RE. Percentile ranks of sonar fetal abdominal circumference measurements. Am J Obstet Gynecol 1980; 138: 475-479

Tamura RK, Sabbagha RE, Depp R, Vaisrub N, Dooley SL, Socol ML. Diminished growth in fetuses born preterm after spontaneous labour or rupture of membranes. Am J Obstet Gynecol 1984; 148: 1105-1110.

Tanner JM, Whitehouse RH. Revised standards for triceps and subscapular skinfold thickness in British children. Arch Dis Child 1975; 50: 142-145.

Taylor AM, Dunger DB, Grant DB, Preece MA. Somatomedin-C / IGF-1 measured by radioimmunoassay and somatomedin bioactivity in adolescents with insulin dependent diabetes compared with puberty matched controls. Diabetic Res 1988; 9: 177-181.

Teberg AJ, Walther FJ, Pena IC. Mortality, morbidity, and outcome of the small-for-gestational age infant. Semin Perinatol 1982; 12: 84-94.

Tenovuo AH, Kero PO, Korvenranta HJ.Risk factors associated with severely small for gestational age neonates. Am J Perinatol 1988; 5: 267-271.

Ter Haar GR. Safety of routine ultrasound. Ultrasound Obstet Gynecol 1992; 2: 237-238.

Thomson AM, Billewicz WZ, Hytten FE.The assessment of fetal growth. J Obstet Gynaecol Br Commonw 1968; 75: 903-916.

Thompson SG, Pocock SJ. Can meta-analyses be trusted? Lancet 1991; 338: 1127-1130.

Thurneau GR, Tamura RK, Sabbagha RE. A simple estimated fetal weight equation based on real-time ultrasound measurements of fetuses less than 34 weeks gestation. Am J Obstet Gynecol 1983; 145: 557-561.

Trudinger BJ, Giles WB, Cook CM, Bombardieri J, Collins L. Fetal umbilical artery flow velocity waveforms and placental resistance: clinical significance. Br J Obstet Gynaecol 1985; 92: 23-30.

Trudinger BJ, Cook CM, Jones L, GIles WB. A comparison of fetal heart rate monitoring and umbilical artery waveforms in the recognition of fetal compromise. Br J Obstet Gynaecol 1986; 93: 171-175.

Trudinger BJ, Giles WB, Cook CM, Connelly A, Thompson RS. Umbilical artery flow velocity waveforms in high-risk pregnancy: randomised controlled trial. Lancet 1987; 1: 188-190.

Trudinger BJ, Cook CM, Jones L, Giles WB. A comparison of fetal heart rate monitoring and umbilical artery waveforms in the recognition of fetal compromise. Br J Obstet Gynaecol 1991; 98: 378-384.

Tyrell SN, Lilford RJ, MacDonald HN, Nelson EJ, Porter J, Gupta JK. Randomized comparison of routine vs highly selective use of Doppler ultrasound and biophysical scoring to investigate high risk pregnancies. Br J Obstet Gynaecol 1990; 97: 909-916.

Udall JN, Harrison GG, Vaucher Y, Walson PD, Morrow G. Interaction of maternal and neonatal obesity. Pediatrics 1978; 68: 17-21.

Ulrich M. Fetal growth patterns in a population of Danish newborn infants. Acta Paediatr Scand (Suppl) 1982; 292: 27-45.

Underwood LE, Clemmons DR, Maes M, D'Ercole AJ, Ketelsleger J-M. Regulation of somatomedin-C / insulin-like growth factor 1 by nutrients. Hormone Res 1986; 24: 166-176.

Usher R, McLean F, Scott KE. Clinical and therapeutic aspects of malnutrition. Pediatr Clin North Am 1970; 17: 169-177.

Van Assche JA, de Prins F, Aerts L, Verjans M. The endocrine pancreas in small for dates infants. Br J Obstet Gynaecol 1977; 84: 751-753.

Villar J, Belizan JM, Spalding J, Klein RE. Postnatal growth of intrauterine growth retarded infants. Early Hum Dev 1982; 6: 262-271.

Villar J, Smeriglio V, Martorell R, Brown CH, Klein RE. Heterogeneous growth and mental development of intra-uterine growth-retarded infants during the first 3 years of life. Pediatrics 1984; 74: 783-791.

Villar J, Belizan JM. The evaluation of the methods used in the diagnosis of intrauterine growth retardation. Obstet Gynecol Survey 1986; 41: 187-199.

Villar J, de Onis M, Kestler E, Bolanos F, Cerezo R, Bernedes H. The differential neonatal morbidity of the intrauterine growth retardation syndrome. Am J Obstet Gynecol 1990; 163: 151-157.

Vintzileos AM, Lodeiro JG, Feinstein SJ, Campbell WA, Weinbaum PJ, Nochimson DJ. Value of fetal ponderal index in predicting growth retardation. Obstet Gynecol 1986; 67: 584-588.

Vintzileos AM, Campbell WA, Rodis JF, McLean DA, Fleming AD, Scorza WE. The relationship between fetal biophysical assessment, umbilical artery velocimetry, and fetal acidosis. Obstet Gynecol 1991; 77: 622-626.

Vosmer MBJG, Jongsma HW, van Dongen PWJ. The value of ultrasonic placental grading: No correlation with intrauterine growth retardation or with maternal smoking. J Perinat Med 1989; 17: 137-143.

van Vugt JMG. Validity of umbilical artery blood velocimetry in the prediction of

intrauterine growth retardation and fetal compromise. J Perinat Med 1991; 19: 15-20.

Vyas S, Nicolaides KH, Campbell S. Renal artery blood flow velocity waveforms in normal and hypoxaemic fetuses. Am J Obstet Gynecol 1989; 161: 168-175.

Vyas S, Nicolaides KH, Bower S, Campbell S. Middle cerebral artery flow velocity waveforms in fetal hypoxaemia. Br J Obstet Gynaecol 1990; 97: 797-803.

Wald NJ, Cuckle HS. Reporting the assessment of screening and diagnostic tests. Br J Obstet Gynaecol 1989; 96: 389-396.

Walther FJ, Raemaker LHJ. The ponderal index as a measure of the nutritional status at birth and its relation to some aspects of neonatal morbidity. J Perinatal Med 1982; 10: 42-47.

Warda AH, Deter RL, Rossavik IK. Fetal femur length: a critical re-evaluation of the relationship to menstrual age. Obstet Gynecol 1985; 66: 69-75.

Warsof SL, Cooper DJ, Little D, Campbell S. Routine ultrasound screening for antenatal detection of intrauterine growth retardation. Obstet Gynecol 1986; 67: 33-39.

Warsof SL, Gohari P, Berkowitz RL, Hobbins JC. The estimation of fetal weight by computer-assisted analysis. Am J Obstet Gynecol 1978; 128: 881-892.

Weile B, Bach-Mortensen N, Peitersen B. Caliper skinfold measurements in newborns: Analysis of a method. Biol Neonate 1986; 50: 192-199.

Weiner CP, Sabbagha RE, Tamura RK, Dalcombo S. Sonographic abdominal circumference: dynamic vs static imaging. Am J Obstet Gynecol 1981; 139: 953-955.

Weiner CP, Sabbagha RE, Vaisrub N, Depp R. A hypothetical model suggesting suboptimal intrauterine growth in infants delivered preterm. Obstet Gynecol 1985a; 65: 323-326.

Weiner CP, Sabbagha RE, Vaisrub N, Socol ML. Ultrasonic fetal weight prediction: role of head circumference and femur length. Obstet Gynecol 1985b; 65: 812-816.

Weiner CP, Robinson D. Sonographic diagnosis of intrauterine growth retardation using the postnatal ponderal index and the crown-heel length as standards of diagnosis. Am J Perinatol 1989; 6: 380-383.

Weinraub Z, Scheider D, Langer R, Brown M, Caspi E. Ultrasonographic measurement of fetal growth parameters for estimation of gestational age and fetal weight. Isr J Med Sci 1979; 15: 829-832.

Weiss PAM, Hofmann H, Purstner P, Winter R, Lichtenegger W. Fetal insulin balance, gestational diabetes and post-partal screening. Obstet Gynecol 1984; 64: 65-68.

Wells PNT. The safety of diagnostic ultrasound. Report of the British Institute of Radiology Working Group. Br J Radiol (suppl) 1987; 20: 1-43.

Wennergren M, Karlsson K, Olsson T. A scoring system for antenatal identification of fetal growth retardation. Br J Obstet Gynaecol 1982; 89: 520-524.

Whitelaw A. Subcutaneous fat measurement as an indication of nutrition of the fetus and the newborn. In: Visser HKA ed. Nutrition and metabolism of the fetus and the infant. Martinus Nijhoff Publishers, The Hague. 1979: 131-143.

Widdowson EM. Intrauterine growth retardation in the pig. I. Organ size and cellular development at birth and after growth to maturity. Biol Neonate 1971; 19: 329-340.

Widdowson EM, Southgate DAT, Hey EN. Body composition of the fetus and infant. In: Visser HKA ed. Nutrition and metabolism of the fetus and the infant. Martinus Nijhoff Publishers, The Hague. 1979: 169-177.

Winick M. Cellular changes during placental and fetal growth. Am J Obstet Gynecol 1971; 109: 166-176.

Wolfe HM, Brans YW, Gross TL, Bhatia RK, Sokol RJ. Correlation of commonly used measures of intrauterine growth with estimated neonatal body fat. Biol Neonate 1990; 57: 167-171.

Wolfe HM, Dombrowski MP, Sokol RJ, Brans YW. Obstetric prediction of the growth-retarded neonate. Am J Obstet Gynecol 1992; 166: 419 (abstract).

Woo JSK, Liang ST, Wan C, Ghosh A, Cho KM, Wong V. Abdominal circumference vs. abdominal area - Which is better? J Ultrasound Med 1984; 3: 101-106.

Yoon BH, Syn HC, Kim SK. The efficacy of Doppler umbilical artery velocimetry in identifying fetal acidosis. J Ultrasound Med 1992; 11: 1-6.

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.

Zelen M. The analysis of several 2 X 2 contingency tables. Biometrika 1971; 58: 129-133.

ACKNOWLEDGEMENTS

Many people have helped and encouraged me during the course of the work described in this thesis. I would like to thank my supervisor, Mr. John Spencer, Senior Lecturer / Consultant in Obstetrics and Gynaecology, University College Hospital, London for his encouragement and sound advice, and Mr. Stephen Robson, Senior Lecturer / Consultant in Obstetrics and Gynaecology, University of Newcastle-upon-Tyne for his help in study construction and participation in the study whilst he was in London. I would also like to thank the Rockefeller Foundation and Appeals Trust, University College London for funding the work for the entire two years.

I must pay tribute to the many women who agreed to participate in these studies. This work was carried out on the Fetal Medicine / Day Assessment Unit at University College Hospital and I am indebted to the midwives on the Unit for their cooperation and help. I am also grateful to the radiographers in the Ultrasound Department for their help in recruiting patients for the study. I would like to thank all the consultant staff at University College Hospital for permission to study their patients. I am grateful to Dr. Stephen Gallivan (Clinical Operations Research Unit, Department of Statistics, University College London) and Dr. Richard Boys (Department of Statistics, University of Newcastle-upon-Tyne) for their help with the statistical analyses. I would like to thank Professor Charles Brook, Department of Child Growth and Development, Middlesex Hospital, London for the use of his equipment for morphometric measurements and in providing training for such purposes. I acknowledge the contribution of Mr. Stephen Robson in assisting in the reproducibility studies and in undertaking the majority of the neonatal morphometric measurements. Dr. Simon Roth, Lecturer in Neonatology at University College Hospital performed neurological examination on all babies born in this study; to him we are grateful for his assistance and helpful advice. I would like to thank Dr. David Crook and Dr. Anthony Proudler, Cavendish Laboratories, Wynn Institute, London for their help in the analysis of triglyceride and insulin levels, and Professor Michael Preece, Dr. David Morrell and Dr. Jennifer Jones, Department of Growth and Development, Institute of Child Health, London for their help in the analysis of insulin-like growth factor levels.

I would also like to thank Professor Charles Rodeck and Mr. Humphry Ward for their constant encouragement during the two years of the study. I am indebted to Miss Gillian Lachelin who read through this thesis and offered helpful advice in the final stages of its preparation.

Finally, I must thank my wife, Jacquie, who supported and encouraged me throughout this work. She has cheerfully endured the preparation of this thesis and has been prepared to sacrifice much to ensure its completion. This thesis is dedicated to her.

ERRATA (MD THESIS - TC CHANG)

p 24 para 2 lines 12-13	should read "... was therefore reflected in the diminution in skinfold thickness."
p 29 para 4 line 9	should read "... heterogeneous group and the majority of such infants are healthy ..."
p 35 para 1 line 1	should read "The plasma non-essential / essential amino acid ratio was increased ..."
p 60 para 2 line 1	should read "The study of Danielian et al. (1992) is the only one to date ..."
p 65 para 5 line 1	should read "The use of data obtained at cordocentesis to confirm IUGR is unlikely to be applicable in ..."
p 76 para 3 line 4-9	should read "The amount of sound that is reflected depends on the degree of acoustic mismatch between the two tissues, the angle at which the beam strikes the interface and the relative thickness of the mismatched tissue compared with the wavelength."
p 79 para 5 line 7	should read "... of HC reduced errors of head measurement due to altered shape, such as dolichocephaly ..."
p 83 para 3 line 3	should read "... continually transmits a beam of ultrasound into the tissue ..."
p 83 para 3 line 6	should read "The advantage with CW equipment is that power levels are lower ..."
p 84 para 1 line 10	should read "... over the entire width of the cord in the same plane ..."
p 90 para 4 line 11	should read "... this was the method of choice."
p 92 para 5 line 8	should read "... defined as the occurrence of ..."
p 96 para 3 line 5	should read "... values were never extrapolated beyond the gestation of a fetus' final scan."
p 97 para 4 line 1	should read "Inter-observer variability was assessed, using the mean of each observer's three ..."
p 101 para 1 line 5	should read "... considered statistically significant (two-tailed)."

p 106 para 1 line 10	should read "... from the final analysis because of delivery at less than 37 weeks gestation ..."
p 119 para 2 line 1	should read "The subjects for the study were Caucasian women chosen sequentially from the two..."
p 130 para 2 line 11	should read "... coefficient of variation ..."
p 131 para 1 line 8	should read "... reported by Tamura et al. (1980) and ..."
p 135 para 1 line 2	should read "... depends to a large extent on their reproducibility."
p 141 para 1 line 8	should read "... were also presented as absolute values ..."
p 145 para 2 line 7	should read "... the findings were confirmed by the author."
p 145 para 4 line 8	should read "... infants had an adverse perinatal outcome ..."
p 146 table 7.1	should read "Umbilical arterial base excess -5.2 (3.3)"
p 163 para 1 lines 1-2	should read "... the optimal method of quantifying serial ultrasound data ..."
p 171 para 1 line 5	should read "... is likely to be related to the use of ..."
p 174 para 1 line 1	should read "The incidence of acidaemia in the present study group was 24%."
p 174 para 2 line 1-2	should read "A significantly greater proportion of fetuses with ultrasound evidence of IUGR was admitted to NICU."
p 198 para 3 line 2	should read "Comparison of Doppler waveform indices and serial ..."
p 229 bottom ref	should read "... inter-relationships among cardiotocograms in ..."

ERRATA (MD THESIS - TC CHANG) p 107

The following sentence is inserted at the end of paragraph 1.

"The number of subjects scanned after 36 weeks gestation were 62 at 37 weeks, 54 at 38 weeks, 44 at 39 weeks and 30 at 40 weeks gestation."

ERRATA (MD THESIS - TC CHANG) p 110

Table 5.4 Centiles for abdominal circumference (mm) [centiles were derived by calculation from an assumed normal distribution]

ERRATA (MD THESIS - TC CHANG) p 158-161

The x-axes for figures 7.5 to 7.8 should read "100 - specificity".

ERRATA (MD THESIS - TC CHANG) p 169

Footnote to **Table 7.11.**

The 5 neonates admitted to NICU from the IUGR group comprised of 1 with hypoglycaemia and 4 with neurological disturbances. The 3 neonates admitted to NICU from the non-IUGR group comprised of 2 with hypoglycaemia and 1 with neurological disturbances.