Accuracy of Prenatal Ultrasound in Detecting Growth Abnormalities in Triplets: A Retrospective Cohort Study

Cameron Sklar,[1] Maryna Yaskina,[2] Sue Ross,[1] and Kentia Naud[1]

[1]*Department of Obstetrics and Gynecology, Division of Maternal Fetal Medicine, University of Alberta, Lois Hole Hospital for Women, Edmonton, Alberta, Canada*
[2]*Women's and Children's Health Research Institute, University of Alberta, Edmonton, Alberta, Canada*

Significant management decisions in triplet pregnancies are made based mainly on ultrasound measurements of fetal growth, although there is a paucity of data examining the accuracy of fetal weight measurements in these gestations. To evaluate accuracy of prenatal ultrasound to diagnose growth abnormalities (intrauterine growth restriction, severe growth discordance) in triplet pregnancies, a retrospective cohort study of 78 triplet pregnancies (234 fetuses) delivered at a single tertiary hospital from January 2004 to May 2015 was performed. Growth percentiles from the last ultrasound were derived from estimated fetal weight using Hadlock's formula for each triplet. Growth discordance was calculated for each triplet set using the formula {(estimated fetal weight largest triplet - estimated fetal weight smallest)/estimated fetal weight largest}. These estimations were compared to birth weights. Sensitivity of ultrasound to predict ≥1 growth restricted fetus in a triplet set was 55.6% [95% CI 35.3, 74.5]; specificity was 100% [95% CI 93.0, 100]; positive predictive value (PPV) 100% [95% CI 74.7, 100]; negative predictive value (NPV) 81.0% [95% CI 73.2, 85.7%]. Sensitivity of ultrasound to detect fetal growth discordance >25% in a triplet set was 80.0% [95% CI 44.4, 97.5], specificity 94.1% [95% CI 85.6, 98.4]; PPV 66.7% [95% CI 42.4, 84.5]; NPV 97.0% [95% CI 90.2, 99.1]. Prenatal ultrasound currently remains the most reliable tool to screen for growth anomalies in triplet pregnancies; however, it appears to have less than ideal sensitivity, missing a number of cases of intra-uterine growth restriction and significant growth discordance.

■ **Keywords:** ultrasound, triplets, fetal growth restriction, growth discordance

Multiple gestation pregnancies have been associated with an increased risk of fetal growth restriction (FGR) and growth discordance (Alexander et al., 1998). Growth anomalies result in increased prenatal surveillance, interventions, preterm birth, and poor perinatal outcomes (Blickstein & Kalish, 2003, Blickstein & Keith, 2003, Jacobs et al., 2003). Low birth weight (LBW) and being born small for gestational age (SGA) are linked to increased mortality and morbidity, and contribute to the significant rate of neonatal mortality and morbidity in multiple pregnancies (Guilherme et al., 2009; Lynch et al., 1995, Weissman et al., 1990, Weissman et al., 1990). Similarly, birth-weight discordance among multiple pregnancies is common and associated with adverse neonatal outcome (Blickstein & Kalish, 2003; Branum & Schoendorf, 2003; Miller et al., 2012), and has been shown to be a strong (OR 10.88 for smallest triplet; Jacobs et al., 2003) independent risk factor to predict neonatal mortality in twins (D'Antonio et al,. 2013).

A paucity of evidence exists regarding the accuracy of prenatal ultrasound in predicting FGR and LBW among triplets, and studies that do attempt these measurements show suboptimal diagnostic accuracy (Guilherme et al., 2009). Past studies on twins also have shown suboptimal accuracy of prenatal ultrasound in detecting significant fetal weight discordance (Reberdao et al., 2010). Moreover, while it has been demonstrated that discordance in birth weight among multiples is common and often severe (Blickstein & Kalish, 2003, Branum & Schoendorf, 2003; D'Antonio et al., 2013), there is little evidence describing the sensitivity and

specificity of prenatal ultrasound in identifying and predicting growth discordance among triplet sets.

Ultrasound techniques for estimation of fetal weight have remained essentially the same for the last 2–3 decades, using the same fetal biometry and Hadlock's formula (Hadlock et al., 1985). Measuring fetal biometry is often challenging in multiple pregnancies, due to fetal crowding and presentation. The aim of this study was to ascertain the sensitivity, specificity, positive, and negative predictive values of modern tertiary level prenatal ultrasounds to predict growth abnormalities (FGR, SGA, growth discordance) in triplet pregnancies.

Materials and Methods

This was a retrospective study of all triplet pregnancies delivered at a single tertiary center (Royal Alexandra Hospital, Edmonton, Alberta, Canada), from January 1, 2004 to May 31, 2015. The main analyses compared the findings of the ultrasound scan closest to date of birth and actual fetal weight discordance observed at birth.

All triplet pregnancies were identified using medical coding in the Alberta Perinatal Health Program Database. Codes for 'triplet pregnancy', 'high-order multiple pregnancy', and 'triplets' were used.

Triplet pregnancies >18 weeks were included when documented on ultrasound; prenatal ultrasounds were performed at the Royal Alexandra Perinatal Clinic, and all triplets were delivered at the Royal Alexandra Hospital between the study dates. Cases were excluded if delivery occurred <23+0 weeks of gestational age (GA), if spontaneous reduction or multi-fetal reduction of a higher order multiple pregnancy into a triplet pregnancy occurred, if fetal reduction (spontaneous or not) of a triplet pregnancy into a twin/singleton pregnancy occurred, if the most recent ultrasound was performed more than 21 days before the delivery, or if the triplet pregnancy had no prenatal care or prenatal ultrasounds.

Ultrasound findings were recorded from the ultrasound reports attached to the maternal hospital records. Actual birth weights (ABWs) were collected from the birth records. Maternal characteristics, mode of conception, and complications of the index pregnancy were taken from the maternal hospital charts. GA was determined using the best estimation from LMP and early ultrasound dating. Chorionicity and amnionicity were assessed by early ultrasound and confirmed by histopathology of the placenta and membranes. GA at last ultrasound before birth, GA at delivery, and interval (days) from last ultrasound to birth were collected. Birth order for the triplets was taken from delivery records.

The following variables were calculated:

1. For each fetus, ultrasound estimated fetal weight (EFW) was calculated using Hadlock's formula, based on HC, abdominal circumference, and femur length (Hadlock et al., 1985).

2. For each fetus, FGR was defined as EFW < 10th percentile for GA using for reference the Canadian Perinatal Surveillance System singleton growth curves (Kramer et al., 2001). An FGR pregnancy was defined as a pregnancy where one or more fetuses were small for GA.

For each pregnancy, EFW discordance (%) was defined as (Largest triplet EFW - Smallest triplet EFW)/(Largest triplet EFW) × 100.

The following variables were obtained:

1. For each fetus, ABW was collected from the birth record.

2. For each fetus, SGA was defined as ABW < 10th percentile for GA using for reference the Canadian Perinatal Surveillance System singleton growth curves (Kramer et al., 2001). An SGA pregnancy was defined as a pregnancy where one or more babies were SGA.

3. For each set of newborn triplet, ABW discordance (%) was defined as (Largest triplet ABW – Smallest triplet ABW)/(Largest triplet ABW) × 100.

Data were entered into a secure database. Data management and analyses were carried out using SPSS 23 (IBM SPSS Version 23.0). For each pregnancy, FGR, SGA, EFW, and ABW discordances (%) were calculated according to the definitions above.

Cross-tabulation, using pregnancy as the unit of analysis, was used to estimate the sensitivity, specificity, positive, and negative predictive values of FGR pregnancy for SGA pregnancy. Cross-tabulation was similarly used to estimate the sensitivity, specificity, positive, and negative predictive values of EFW to predict strong (>25%) discordance versus ≤25% discordance in ABW. To estimate the agreement between EFW discordance and ABW discordance, two-way mixed intra-class correlation coefficient was computed with 95% confidence interval (CI). A Bland–Altman plot was constructed to estimate the bias (*SD*) in the relation between the average of EFW and ABW discordance versus the difference between the EFW and ABW discordance.

The local research ethics committee and health authority granted approvals for this study, and also provided waiver of consent as the study was retrospective in nature.

Results

Between January 1, 2004 and May 31, 2015, 85 triplet pregnancies were delivered at RAH. Of those, 78 pregnancies were included for this study (Figure 1). The median maternal age was 31 years, with an IQR of 7 years. All additional baseline characteristics of the pregnancies are presented in Table 1. The final ultrasound before delivery was performed at median 30.9 weeks of GA, (Table 2), with a median interval between last ultrasound and delivery of

Cameron Sklar, Maryna Yaskina, Sue Ross and Kentia Naud

TABLE 1

Characteristics of the Pregnancies

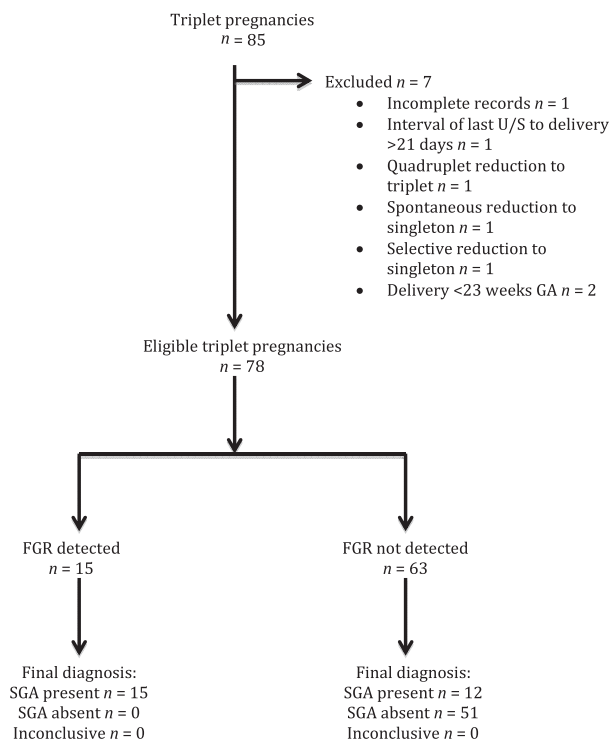|		Frequency	Percentage
Parity	Nulliparous	42	53.8%
	Multiparous	36	46.2%
	Total	78	100%
Mode of conception	Assisted	48	61.5%
	Spontaneous	30	38.5%
	Total	78	100%
Chorionicity	Monochorionic	6	7.7%
	Dichorionic	27	34.6%
	Trichorionic	45	57.7%
	Total	78	100%
Complications of index pregnancy	Smoker	6	7.7%
	Hypertensive disease of pregnancy	15	19.2%
	Gestational diabetes	12	15.4%

FIGURE 1

STARD flow diagram of participants.

8 days. Median GA at delivery was 31.71 weeks, and all included triplets were live births (Table 3). Out of all babies delivered, 49/234 (20.94%) were SGA (Table 4); while 26/234 fetuses (11.1%) were predicted by ultrasound to be growth restricted (Table 4).

Cross-tabulation was used to compare the ultrasound prediction of fetal growth anomalies to the ABWs at delivery (Table 5), as well as to compare prediction of severe growth discordance to ABW discordance (Table 6).

TABLE 2

Final Ultrasound Data

		Median	IQR
GA at last ultrasound (in weeks)		30.9	5.64
EFW at last ultrasound (in grams)	Total ($n = 234$ fetuses)	1,438	882
	A	1,518	940
	B	1,478	845
	C	1,417	901

TABLE 3

Birth Data of Triplet Pregnancies Included in Study

		Median	IQR
GA at delivery (in weeks)		31.71	5.29
ABW (in grams)	Total ($n = 234$ neonates)	1,518	840
	A	1,595	848
	B	1,430	780
	C	1,430	875

TABLE 4

Number of SGA Infants, and Number of Predicted FGR Fetuses in Triplet Pregnancies

		Frequency	Percentage
Pregnancy with SGA infants	0 infant SGA	51	65.4%
	One infant SGA	14	17.9%
	Two infants SGA	4	5.1%
	Three infants SGA	9	11.5%
	Total	78	100%
Pregnancy with ABW discordance ≥25%		10	12.8%
Pregnancy with FGR	0 fetus FGR	63	80.8%
	One fetus FGR	8	10.3%
	Two fetuses FGR	3	3.8%
	Three fetuses FGR	4	5.1%
	Total	78	100%
Pregnancy with EFW discordance ≥25%		12	15.4%

TABLE 5

Cross-Tabulation of Predicted FGR and Actual Birth Weight

FGR pregnancy (at least one fetus with EFW < 10th percentile)	SGA pregnancy (at least one baby with BW < 10th percentile)		Total
	Yes	No	
Yes	15	0	15
No	12	51	63
Total	27	51	78

TABLE 6

Cross-Tabulation of Predicted Severe Fetal Growth Discordance and Actual Birth-Weight Discordance

EFW discordance	ABW discordance		Total
	>25%	≤25%	
>25%	8	4	12
≤25%	2	64	66
Total	10	68	78

TABLE 7

Calculated Test Characteristics of Cross-Tabulation of FGR and Actual Birth Weight

	Estimate	95% confidence interval
Sensitivity	55.6%	[35.3%, 74.5%]
Specificity	100%	[93.0%, 100%]
Positive predictive value	100%	[74.7%, 100%]
Negative predictive value	81.0%	[73.2%, 85.7%]

TABLE 8

Calculated Test Characteristics of Cross-Tabulation of Predicted and Actual Birth Weight Discordance

	Estimate	95% confidence interval
Sensitivity	80.0%	[44.4%, 97.5%]
Specificity	94.1%	[85.6%, 98.4%]
Positive predictive value	66.7%	[42.4%, 84.5%]
Negative predictive value	97.0%	[90.2%, 99.1%]

Of all pregnancies evaluated, ultrasound correctly identified 55.6% of pregnancies as having at least one FGR triplet. Specificity and PPV were 100%, and NPV was 81.0% (Table 7). Cohen's kappa indicated moderate to good agreement between having at least one FGR triplet and at least one SGA triplet at birth ($\kappa = 0.620$ [(95% CI 0.348, 0.802], $p < .0005$).

For pregnancies with ABW discordance of >25%, the ultrasound identified 80.0% of these pregnancies as having severe EFW discordance. Specificity was 94.1%, NPV was 97.1% and PPV was 66.7% (Table 8). Cohen's kappa indicated moderate to good agreement between having EFW discordance of >25% and having ABW discordance of >25% at birth ($\kappa = 0.683$ [95% CI 0.446, 0.920], $p < .0005$).

Discussion

The potential for significant fetal morbidity and mortality in triplet pregnancies is well described (Blickstein & Keith, 2003) though data determining the accuracy of ultrasound to predict growth abnormalities in these high-risk pregnancies is scarce. This is, to our knowledge, the largest study assessing the diagnostic accuracy to detect fetal growth anomalies in triplet gestations. Our calculated sensitivity to detect FGR in a triplet pregnancy is 55.6% [95% CI 35.6, 74.0], although the specificity is 100% [95% CI 91.3, 100]. Our findings are similar to fetal growth ultrasound studies in singleton and twin pregnancies. Diagnostic accuracy to detect FGR in singletons has been previously shown to have a sensitivity of 65–70% and specificity of 79–89% (De Jong et al., 2000; Ott, 2002). Similar studies of twins have also been performed, with sensitivities to detect FGR < 10th percentile reported from 46% to 85% and specificity ranging from 87% to 98% (Danon et al, 2008; Jensen & Jenssen, 1995; Secher et al., 1995).

This is the first study to report the accuracy of prenatal ultrasound to predict inter-triplet growth discordance, an important marker of perinatal morbidity and mortality. We found prenatal ultrasound to have a sensitivity of 80.0% [95% CI 44.2, 96.5%] and a specificity of 94.1% [95% CI 84.9, 98.1%] to detect triplets with a birth-weight discordance of at least 25%. Of note, three out of the four false positives had divergent estimation (Reberdao et al., 2010), whereby the larger triplet was overestimated and the smallest underestimated. The other false positive had an accurately measured largest triplet, but underestimated the smallest triplet significantly. Of the two false negatives, in both cases the smallest EFW was significantly overestimated, falsely affecting the discordance measurements. Twin studies evaluating sensitivity to detect birth-weight discordance of at least 25% have ranged from 11.1% to 37%, with specificities 94–98% (Caravello et al., 1997; Chamberlain et al., 1991; Reberdao et al, 2010).

It is known that ultrasound error range for EFW is the widest at the growth curves' extremities (Nahum & Stanislaw, 2003; Seimer et al., 2008). In our study, most ultrasound errors came from under- or over-estimating the smallest triplet.

EFW discordance measurements ≥25% are known to be predictive of increased mortality in twins (D'Antonio et al., 2013, Khalil et al., 2015), as has birth weight discordance ≥29% among triplets (Jacobs et al., 2003). Thus, defining the diagnostic accuracy to detect these parameters has obvious importance for decision making in the management of these pregnancies. Ideally, prenatal ultrasound should be 100% sensitive for growth anomalies, in order to offer adequate prenatal surveillance and timing of delivery. The ability to detect growth anomalies in triplets with ultrasound appears to be currently suboptimal, even in a tertiary high volume setting like ours.

In order to improve ultrasound sensitivity to diagnose growth anomalies in triplets, a change in diagnostic criteria could be considered. For example, the threshold for IUGR in triplets could be set at the 15th percentage for GA instead of the 10th; for growth discordance, the threshold could be set at 20% discordance instead of 25%. Other criteria to define IUGR could be used, such as a significant decrease in growth velocity.

Limitations of the study include its retrospective nature and limited sample size. Additionally, some may question the length of interval from last ultrasound to birth allowed for inclusion. The intervals ranged from 0–21 days, with the median length of time being 8 days (Table 1). This interval corresponds to the maximum interval frequency of ultrasound examinations for triplets at our center during the overall period of the study. Moreover, our study aims to determine whether a growth-restricted triplet can be identified by prenatal ultrasound irrespective of length of time until delivery occurred. This study does not evaluate the ultrasound accuracy to predict ABW, but rather its ability

to identify IUGR and severe growth discordance. However, the possibility exists that some fetuses could either gain or lose growth velocity just prior to delivery, thus being missed by an ultrasound performed 1–3 weeks before delivery.

Another limitation is that it is difficult to link with certainty the ultrasound order of birth to the actual order of birth, specifically when birth occurs via cesarian section. In reviewing charts, we tried to link ABWs to EFWs using common sense and described birth order; however, errors might have occurred. It is uncertain that fetus 'A' was newborn 'A'. This factor could worsen the test characteristics of our study, decreasing sensitivity, and PPV furthermore.

Additionally, singleton growth curves were used instead of twin or triplet curves, as these are the most common curves used in published studies involving multiple pregnancies. It is possible that the singleton growth curves could overestimate or underestimate the number of IUGR cases, since singletons tend to be larger than multiples.

Strengths of this study include the completeness of data for all prognostic and outcome variables, and relatively large sample size. All ultrasounds were performed by dedicated sonographers and supervised by experienced maternal fetal medicine specialists in a single tertiary center, reducing variability, and providing consistency in measurement and interpretation.

In conclusion, our study, the first to assess the ability of prenatal ultrasound to detect severe growth discordance among triplets, also demonstrates that prenatal ultrasound has less than ideal sensitivity to detect growth anomalies in triplets, resulting in a number of missed cases of FGR and severe growth discordance.

Relaxing diagnostic criteria for growth anomalies in triplets could be considered if the actual Hadlock's formula keeps being used in the future. New formulae to estimate fetal weight using advanced modalities (e.g., 3D, rate of rise) might also provide more reliable EFWs and increase prenatal ultrasound sensitivity.

Acknowledgment

Conflict of Interest

None.

References

Alexander, G. R., Kogan, M., Martin, J., & Papiernik, E. (1998). What are the growth patterns of singletons, twins, triplets in the United States?. *Clinical Obstetrics & Gynecology, 41*, 114–125.

Blickstein, I., & Kalish, R. (2003). Birth weight discordance in multiple pregnancy. *Twin Research and Human Genetics*, 6, 526–531.

Blickstein, I., & Keith, L. G. (2003). Outcome of triplets and higher-order multiple pregnancies. *Current Opinion in Obstetrics & Gynecology, 15*, 113–118.

Branum, A. M., & Schoendorf, K. C. (2003). The effect of birth weight discordance on twin neonatal mortality. *Obstetrics & Gynecology*, 101, 570–574.

Caravello, J. W., Chauhan, S. P., Morrison, J. C., Magann, E. F., Martin, J. N., & Devoe, L. D. (1997). Sonographic examination does not predict twin growth discordance accurately. *Obstetrics & Gynecology, 89*, 529–533.

Chamberlain, P., Murphy, M., & Comerford, F. R. (1991). How accurate is antenatal sonographic identification of discordant birthweight in twin?. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 40, 91–96.

Danon, D., Melamed, N., Bardin, R., & Meizner, I. (2008). Accuracy of ultrasonographic fetal weight estimation in twin pregnancies. *Obstetrics and Gynecology*, 112, 759–764.

D'Antonio, F., Khalil, A., Dias, T., & Thilaganathan, B. (2013). Weight discordance and perinatal motality in twins: Analysis of the Southwest Thames Obstetric Research Collaborative (STORK) multiple pregnancy cohort. *Ultrasound in Obstetrics & Gynecology, 41*, 643–648.

De Jong, C. L. D., Francis, A., Van Geijn, H. P., & Gardosi, J. (2000). Customized fetal weight limits for antenatal detection of fetal growth restriction. *Ultrasound in Obstetrics and Gynecology, 15*, 36–40.

Guilherme, R., Le Ray, C., Vuillard, E., Garel, C., Delezoide, A. L., Oury, J. F., & Luton, D. (2009). Ultrasound assessment of the prognosis in triplet pregnancies. *Acta Obstetricia et Gynecolica Scandinavica, 88*, 386–390.

Hadlock, F. P., Harrist, R. B., Sharman, R. S., Deter, R. L., & Par, S. K. (1985). Estimation of fetal weight with the use of head, body, and femur measurements — A prospective study. *American Journal of Obstetrics & Gynecology, 151*, 333–337.

Jacobs, A. R., Demissie, K., Neetu, J. J., & Kinzler, W. L. (2003). Birth weight discordance and adverse fetal and neonatal outcomes among triplets in the United States. *Obstetrics & Gynecology, 101*, 909–914.

Jensen, H. R., & Jenssen, H. (1995). Prediction of fetal weights in twins. *Acta Obstetricia et Gynecologica Scandinavica, 74*, 177–180.

Khalil, A. A., Khan, N., Bowe, S., Familiari, A., Papageorghiou, A., Bide, A., & Thilaganathan, B. (2015). Discordance in fetal biometry and Doppler are independent of the risk of perinatal loss in twin pregnancies. *American Journal of Obstetrics and Gynecology, 213*, 222e1–10.

Kramer, M. S., Platt, R. W., Wen, S. W., Joseph, K. S., Allen, A., Abrahamowicz, M., … Breart, G. (2001). A new and improved population-based Canadian reference for birth weight gestational age. *Pediatrics, 108*, e35.

Lynch, L., Lapinski, R., Alvarez, M., & Lockwood, C. J. (1995). Accuracy of ultrasound estimation of fetal weight in multiple pregnancies. *Ultrasound in Obstetrics & Gynecology, 6*, 349–352.

Miller, J., Chauhan, S. P., & Abuhamad, A. Z. (2012). Discordant twins: Diagnosis, evaluation, and management. *American Journal of Obstetrics & Gynecology, 206*, 10–20.

Nahum, G. G., & Stanislaw, H. (2003). Ultrasonographic prediction of term birth weight: How accurate is it?. *American Journal of Obstetrics & Gynecology, 188*, 566–574.

Ott, W. J. (2002). Diagnosis of intrauterine growth restriction: Comparison of ultrasound parameters. *American Journal of Perinatology, 19*, 133–137.

Reberdao, M. A., Martins, L., Torgal, M., Viana, R., Seminova, T., Casal, E. … Blickstein, I. (2010). The source of error in the estimation of intertwin birth weight discordance. *Journal of Perinatal Medicine, 38*, 671–674.

Secher, J. S., Kaern, J., & Hansen, P. K. (1995). Prediction of fetal weights in twins. *Acta Obstetricia et Gynecologica Scandinavica, 74*, 177–180.

Seimer, J., Egger, N., Hart, N., Meurer, B., Muller, A., Dathe, O., … Schild, R. L. (2008). Fetal weight estimation by ultrasound: Comparison of 11 different formulae and examiners with differing skill levels. *Ultraschall in Medicine, 29*, 159–164.

Weissman, A., Jakobi, P., Yoffe, N., Zimmer, E. Z., Paldi, E., & Brandes, J. M. (1990). Sonographic growth measurements in triplet pregnancies. *Obstetrics & Gynecology, 75*, 324–326.

Weissman, A., Ulanovsky, I., Burke, Y., Makhoul, I. R., Blazer, S., & Drugan, A. (1990). Sonographic growth measurements in triplet pregnancies. *Obstetrics & Gynecology, 75*, 324–326.