



Continuous independent quality control for fetal ultrasound biometry provided by the cumulative summation technique

Balsyte, D ; Schäffer, L ; Burkhardt, T ; Wissler, J ; Zimmermann, R ; Kurmanavicius, J

Abstract: **OBJECTIVE:** To apply the cumulative summation (CUSUM) technique for an evaluation of the learning process of sonographic fetal weight estimation at term in combination with the z-scores of biometry determinants and to assess the time of appearance and sources of errors. **METHODS:** Learning curve (LC-CUSUM) and double CUSUM charts for systematic error detection based on absolute and signed mean percentage error were generated to retrospectively estimate the longitudinal accuracy of sonographic fetal weight estimation conducted by three trainees and one experienced examiner. For LC-CUSUM analysis an examination was considered to be a failure when there was an absolute error in birth weight estimation $\geq 15\%$. Fetal biometry measurements (head circumference, abdominal circumference (AC) and femur length (FL)) from 227 routine ultrasound scans of one examiner were separately transformed into z-scores and double CUSUM charts were generated to assess the systematic errors for each determinant. **RESULTS:** The LC-CUSUM charts revealed that different numbers of scans are required for different examiners to achieve competence in estimating birth weight. AC and FL deviated most significantly from expected values ($P < 0.05$). The double CUSUM charts revealed exact periods of systematic errors in the measurement of biometry determinants, clearly reflecting errors of fetal weight estimation. **CONCLUSIONS:** The use of CUSUM techniques in the analysis of sonographic data allows observation of the development of an examiner's skill and maintenance of competence. The CUSUM technique not only allows the reasons for impaired fetal weight estimation to be revealed but also allows determination of the exact time when inaccurate measurements start to occur. We suggest that CUSUM charts should be implemented in routine clinical practice as a measure of objective quality evaluation of sonographic fetal biometry.

DOI: <https://doi.org/10.1002/uog.7545>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-44576>

Journal Article

Originally published at:

Balsyte, D; Schäffer, L; Burkhardt, T; Wissler, J; Zimmermann, R; Kurmanavicius, J (2010). Continuous independent quality control for fetal ultrasound biometry provided by the cumulative summation technique. *Ultrasound in Obstetrics and Gynecology*, 35(4):449-455.

DOI: <https://doi.org/10.1002/uog.7545>

Continuous independent quality control for fetal ultrasound biometry provided by the cumulative summation technique

D. BALSYTE, L. SCHÄFFER, T. BURKHARDT, J. WISSER, R. ZIMMERMANN
and J. KURMANAVICIUS

Department of Obstetrics, University Hospital of Zurich, Switzerland

KEYWORDS: cumulative summation; fetal biometry; fetal weight estimation; quality control of prenatal ultrasound; z-scores

ABSTRACT

Objective To apply the cumulative summation (CUSUM) technique for an evaluation of the learning process of sonographic fetal weight estimation at term in combination with the z-scores of biometry determinants and to assess the time of appearance and sources of errors.

Methods Learning curve (LC-CUSUM) and double CUSUM charts for systematic error detection based on absolute and signed mean percentage error were generated to retrospectively estimate the longitudinal accuracy of sonographic fetal weight estimation conducted by three trainees and one experienced examiner. For LC-CUSUM analysis an examination was considered to be a failure when there was an absolute error in birth weight estimation $\geq 15\%$. Fetal biometry measurements (head circumference, abdominal circumference (AC) and femur length (FL)) from 227 routine ultrasound scans of one examiner were separately transformed into z-scores and double CUSUM charts were generated to assess the systematic errors for each determinant.

Results The LC-CUSUM charts revealed that different numbers of scans are required for different examiners to achieve competence in estimating birth weight. AC and FL deviated most significantly from expected values ($P < 0.05$). The double CUSUM charts revealed exact periods of systematic errors in the measurement of biometry determinants, clearly reflecting errors of fetal weight estimation.

Conclusions The use of CUSUM techniques in the analysis of sonographic data allows observation of the development of an examiner's skill and maintenance of competence. The CUSUM technique not only allows the reasons for impaired fetal weight estimation to be revealed but also allows determination of the exact time when

inaccurate measurements start to occur. We suggest that CUSUM charts should be implemented in routine clinical practice as a measure of objective quality evaluation of sonographic fetal biometry. Copyright © 2010 ISUOG. Published by John Wiley & Sons, Ltd.

INTRODUCTION

The increasing impact of legal concerns and the pressure for cost-effective decisions have produced a need to implement effective quality-control systems for sonographic examinations in prenatal medicine¹. Ultrasound fetal biometry is the most frequently used technique for the assessment of fetal growth and weight estimation². Insufficient quality of ultrasound examinations, including high intra- and interobserver variability³, has a strong impact on the accuracy of fetal measurements⁴. Several methods for the quality control of fetal biometry have been proposed, the majority of which, however, are not practicable in routine clinical practice. The effectiveness of guidance by an experienced reviewer is limited by time restrictions, subjectivity and randomly occurring errors^{5,6}. Quality assurance using image-scoring methods does not provide a longitudinal evaluation of the process. Furthermore, executing an external audit may result in the selection of only the best ultrasound images obtained.

In the past, the accuracy of sonographic fetal weight estimation has been assessed by calculating the mean signed percentage error (PE) and its absolute value (APE)⁷. The disadvantage of these mathematical calculations, however, is that the temporal dynamic of the accuracy of fetal weight estimation and the experience of the examiner are not included in the process of the evaluation.

Correspondence to: Dr D. Balsyte, Department of Obstetrics, Zurich University Hospital, Frauenklinikstrasse 10, CH-8091 Zurich, Switzerland (e-mail: Dalia.Balsyte@usz.ch)

Accepted: 17 July 2009

Salomon *et al.*⁸ suggested the use of the standardized z -score system to estimate the accuracy of sonographic fetal biometry. If fetal measurements are accurate, the z -score distribution will follow a standardized normal distribution¹. In an optimal situation, measurements lying on the 50th percentile should correspond to the theoretical 50th percentile, and the resulting z -score should be equal to 0, whereas the z -scores of the 5th and 95th percentiles should be close to ± 1.645 in a non-selected population. Nevertheless, although the z -scores method compares the distribution of measured parameters with reference values adjusted for the population at the specific institution concerned, it also lacks the capability to identify the point in time when systematic errors of measurement occur.

Cumulative summation (CUSUM) charts are an established method for quality control in a number of different fields of medicine^{9–12}. We employed CUSUM, a statistical tool that graphically presents the outcomes of any consecutive procedures¹³, to estimate the putative factors diminishing the accuracy of ultrasound fetal biometry and to assess the competence of a number of examiners over a certain period of time, focusing on systematic and random errors¹⁴.

A combination of the CUSUM technique with the z -scores system allows an objective evaluation of fetal biometry measurements. We hypothesized that when CUSUM curves of fetal biometry determinants identify systematic errors of single parameters, similar errors should appear in the CUSUM curve of fetal weight assessment.

The use of the CUSUM method to analyze the learning curve of trainees in the estimation of sonographic fetal weight within the last week prior to delivery could be a valuable tool in the evaluation of their competence, as the actual birth weight is soon available for comparison. Furthermore, a combination of the CUSUM technique with the z -scores system could be used to monitor the quality of measurements of individual fetal biometry parameters and so allow the identification of the causes of inaccurate sonographic fetal weight estimations.

METHODS

Following institutional review board approval from the Committee on Human Research at the University of Zurich, fetal biometry data from a total of 1298 routine ultrasound scans performed by three trainees (Examiners 1–3) at the beginning of their ultrasound training and one experienced examiner (Examiner 4) with 10 years' experience of daily ultrasound scanning at the University Hospital of Zurich during 2004–2007 were analyzed retrospectively. All the results from each examiner were recorded sequentially. Data included sonographic estimated fetal weight (EFW) of all live-born, singleton term (≥ 37 weeks' gestation) deliveries obtained within the last week before delivery. Infants with congenital malformations were excluded from the study.

Fetal weight was estimated according to the three-parameter formula of Hadlock *et al.*, which includes

head circumference (HC), abdominal circumference (AC) and femur length (FL), or their two-parameter formula including AC and FL if the fetal head parameters biparietal diameter (BPD) and occipitofrontal diameter (OFD) could not be accurately obtained¹⁵. Fetal measurements were all made in the planes described by Campbell *et al.*^{16,17}. The BPD was measured from the outer margins of the skull (outer–outer), which is accepted as the standard method in Germany and Switzerland. HC was calculated from linear measurements of BPD and OFD, and AC from abdominal transverse diameter and anteroposterior abdominal diameter using the ellipse formula. All measurements were obtained with Sonoline Prima and Allegra ultrasound machines (Siemens Medical Systems Inc., Malvern, PA, USA) using a 3.5-MHz transducer. Real birth weight was determined as a reference standard for assessment of the accuracy of sonographic fetal weight estimation, with the APE and PE calculated for each case independently.

The learning curve (LC-CUSUM) chart for each examiner was generated using the APE according to the method described by Bolsin and Colson¹³, in which the null hypothesis states that the process is out of control and the alternative hypothesis determines the process to be in control¹⁸. The main principle of CUSUM is that each procedure is assigned a score of which the size and polarity depend on the chosen standard and actual outcome, respectively¹⁹. Each score is sequentially added to the cumulative score and plotted graphically. As a result, an increment of the graph reports a failure and shows the process being out of control (with significance reached at the upper boundary line), whereas a decrement demonstrates a success and achieved competence (with significance reached at the lower boundary line). When the CUSUM chart oscillates and remains between the two boundary lines, no statistical evaluation can be made, indicating that more observations are necessary²⁰. In our analysis an examination was considered to be a failure when there was an absolute error in birth weight estimation of $\geq 15\%$, the acceptable failure rate was set at 5% and the unacceptable failure rate at 15%.

For an evaluation of the systematic error of fetal weight estimation double CUSUM charts based on the PE were generated, with positive and negative errors summed separately, but presented on the same chart, for each examiner. The upper CUSUM detects increases in the positive failure rate (overestimation) while the lower CUSUM detects increases in the negative failure rate (underestimation) of the birth weight estimation.

To assess the distribution of individual fetal measurements in comparison with the normal distribution of the reference values, HC, AC and FL were transformed into z -scores^{21,22}. For this analysis, data from the 227 ultrasound scans performed by Examiner 2 were used. The z -scores were calculated according to the following formula:

$z\text{-score} = (X - M_{GA})/SD_{GA}$, where X is the measured value, M_{GA} is the mean value for the appropriate

gestational age and SD_{GA} is the standard deviation associated with the mean value at that gestational age²¹.

Subsequently, double CUSUM charts were constructed using the z -scores to assess the accuracy of the examiner in measuring each parameter of fetal biometry. Fetal parameters exhibit a certain amount of variation around the reference value owing to a natural variability and the individual manner of measuring. The standardized differences between the measured and expected values should approximately follow a normal distribution, with the mean value = 0 and the $SD = 1$ ¹⁸. If a tendency of over- or underestimation occurs and the standardized differences are greater than the tolerated level (here defined as 0.5), the CUSUM curve will deviate markedly from zero.

A control-limit violation occurs when either the positive or negative CUSUM curve exceeds a specified control limit. In our study the upper and lower control limits for double CUSUM charts were defined as three SDs from the expected value. This selection implies that the difference in expected and measured values lies within the control chart limits in 99.73% of all individual examinations. A more detailed description of the CUSUM techniques used is presented in the online supplement (Appendix S1). A one sample t -test was applied to compare the mean of each z -score distribution with the theoretically expected value.

RESULTS

Between 182 and 622 sonographic fetal weight estimations were performed at term by Examiners 1–4. The mean APE, PE, interval between ultrasound examination and delivery and infant birth weight for each examiner are given in Table 1.

The LC-CUSUM curves were appropriate and demonstrated the ongoing failures and successes when an unacceptable failure rate was defined as 15% (Figure 1a). Examiner 1 estimated fetal weight rather efficiently and consistently from the very beginning, with isolated inaccurate measurements. The acceptable level of accuracy had been reached after 20 scans, when the gap between two boundary lines was crossed. From the 36th to the 70th attempt the plot showed an upward trend, although not reaching an unacceptable boundary line. After this the performance improved, with the plot crossing the acceptable boundary line from above at the 107th scan and staying almost at the same level until the end of the observation period, with short periods of reduced accuracy.

The graphical presentation of the performance of Examiner 2 describes a typical learning curve. The graph shows an upward trend revealing consistent errors until the 132nd scan. Thereafter, the plot spanned the gap between two boundary lines downwards and revealed achieved competence at the 166th scan. The accuracy of Examiner 3 was very limited during the first 151 scans as the slope was constantly rising. Beginning with the 177th attempt, the LC-CUSUM plot crossed in a downward direction the gap between two boundary lines, revealing that the examiner had become proficient. The LC-CUSUM graph demonstrating sonographic fetal weight estimation by Examiner 4 is presented in Figure 1b. The graph constantly moved downwards, revealing a high degree of competence in fetal weight estimation in the frame of our predefined standards.

Figure 2a shows the double CUSUM chart based on the PE presenting persistent accuracy of fetal weight estimation performed by Examiner 1. The process was rather precise, with several deviations indicating a negative systematic error.

The performance of Examiner 2 varied less from the center line compared with Examiner 1. However, after the 180th scan a negative systematic error occurred, leading to an obvious underestimation of fetal weight (Figure 2b).

Figure 2c demonstrates a distinct underestimation of fetal weight by Examiner 3. Starting with the 7th attempt, an alarm signal was triggered, indicating a lack of competence. During most of the study period this examiner had been underestimating fetal weight, as the cumulative sum was greatly negative. An improvement in accuracy could be suspected only at the very end of the study period.

The accuracy of sonographic fetal weight estimation by Examiner 4 is presented in Figure 2d. Maximal accuracy is found after the 190th scan. The CUSUM charts and the center line converge as the cumulative score sum is low enough.

To analyze the impact of single determinants of fetal weight estimation, the fetal biometry data from 227 examinations performed by Examiner 2 were converted into z -scores and agreement with standard 5th, 50th and 95th percentiles was analyzed (Table 2). The accuracy of HC measurements seemed to be good as they were consistent with the normal Gaussian distribution and did not differ significantly from the expected standards of z -scores. In contrast, the z -scores for AC and FL were significantly different

Table 1 Characteristics of the study population and errors in birth weight estimation for each of the examiners

Variable	Examiner			
	1 (n = 226)	2 (n = 227)	3 (n = 182)	4 (n = 622)
APE (%)	7.1 ± 5.2	7.1 ± 5.6	8.3 ± 5.6	5.9 ± 4.6
PE (%)	-2.2 ± 8.6	-3.0 ± 8.6	-4.5 ± 9.0	-1.6 ± 7.3
Ultrasound scan before delivery (days)	0.9 ± 1.25	0.9 ± 1.15	0.8 ± 0.86	2.0 ± 1.54
Birth weight (g)	3365 ± 483	3375 ± 454	3397 ± 418	3466 ± 536

All data are presented as mean ± SD. APE, absolute percentage error; PE, signed mean percentage error.

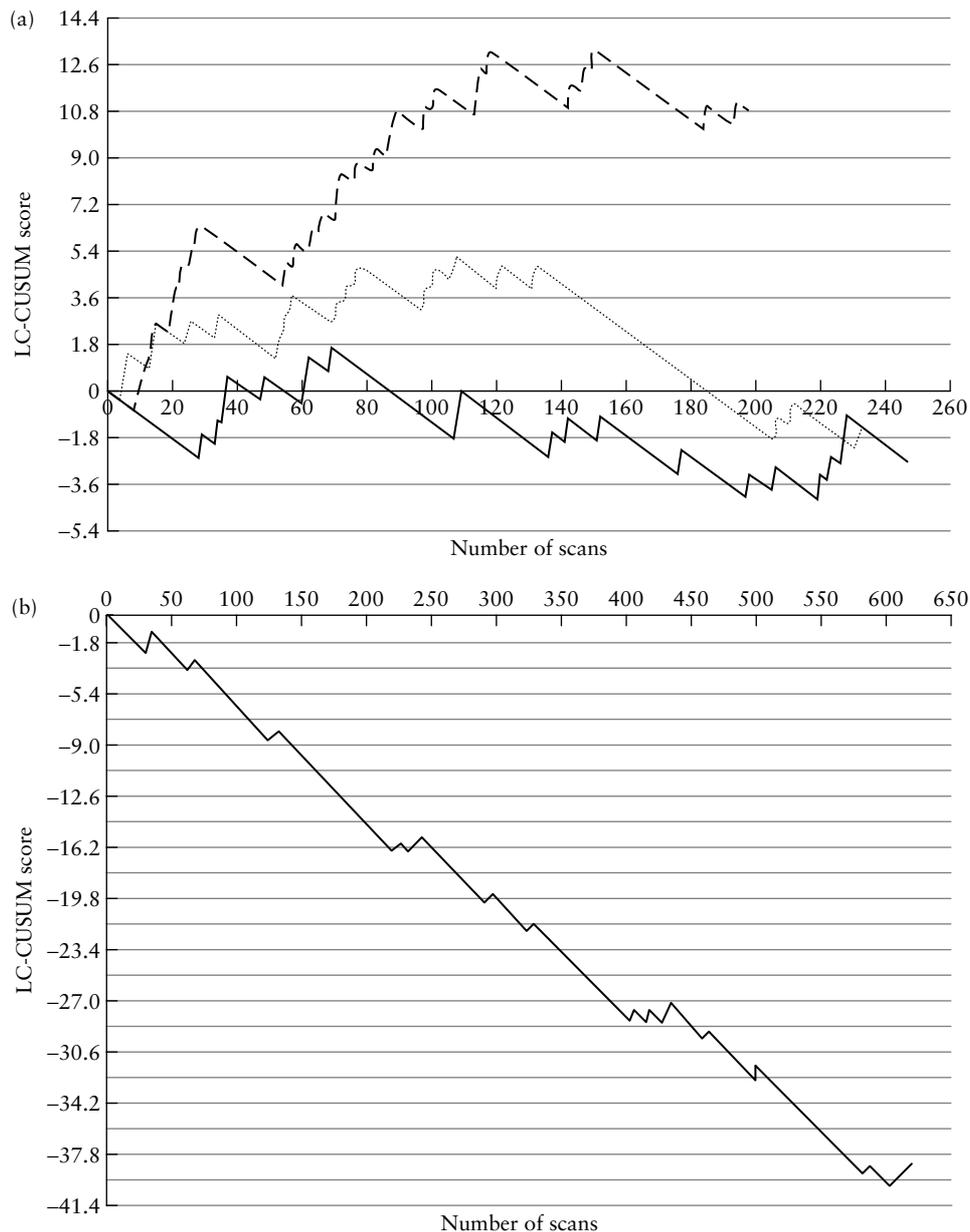


Figure 1 Learning curve cumulative summation (LC-CUSUM) charts of the accuracy of sonographic fetal weight estimation, based on absolute percentage error (APE), for Examiners 1–3 (—, 1; - - -, 2 - - -, 3) (a) and Examiner 4 (b). An examination was considered to be a failure when the birth weight estimation APE was $\geq 15\%$, with the acceptable failure rate set at 5% and the unacceptable failure rate at 15%. Boundary lines are shown at intervals of 1.8 (calculated as a standard for our clinic at the time of the examinations). Crossing two boundary lines in an upward direction signifies unacceptable performance and crossing two boundary lines in a downward direction signifies acceptable performance.

from the expected mean of zero, indicating an unsatisfactory performance. Correspondingly, the double CUSUM chart showed that the consecutive measurements of HC were mostly within the control limits during the study period, showing only two transient negative and positive deviations (Figure 3a). In contrast, the double CUSUM chart presenting the ongoing quality of AC measurements (Figure 3b) revealed poor accuracy, with a strong tendency towards overestimation between scans 61 and 118 and 135 and 183. Solely between scans 183 and 217 did the chart oscillate between the acceptable boundary lines. During the final scans AC was again systematically overestimated. The accuracy of fetal FL measurements

corresponded with the predefined standards until scan 154. Subsequently, a systematic negative error occurred and persisted until the end of the study period (Figure 3c). This underestimation corresponded closely to the fetal weight underestimation shown in Figure 2b.

DISCUSSION

The evaluation of competence and quality of measurements in sonographic fetal biometry has recently become of increasing interest in prenatal medicine. The rationale for performing monitoring is to improve the performance

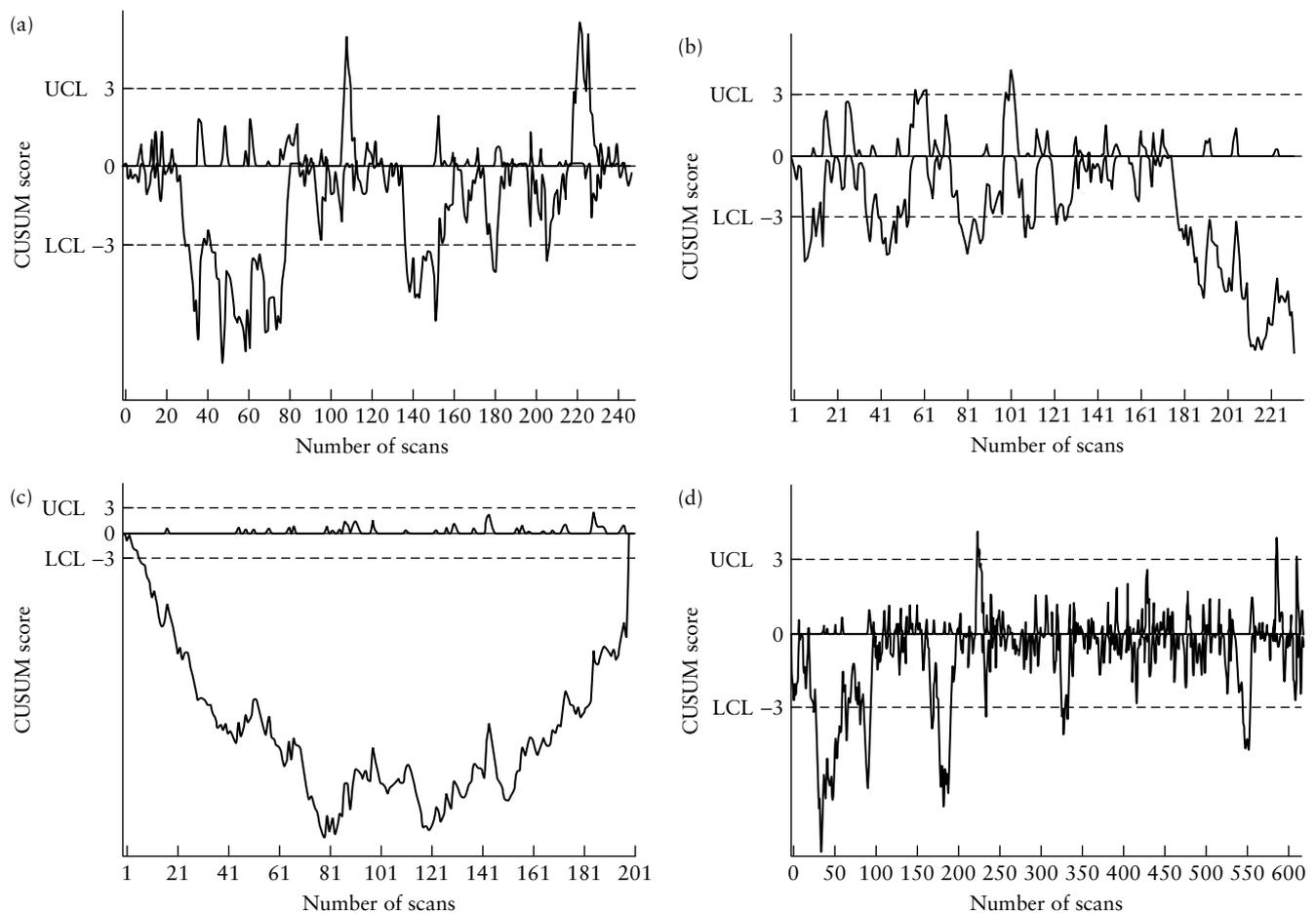


Figure 2 Double cumulative summation (CUSUM) charts demonstrating the accuracy of sonographic fetal weight estimation, based on percentage error for Examiner 1 (a), Examiner 2 (b), Examiner 3 (c) and Examiner 4 (d). LCL, lower control limit; UCL, upper control limit.

Table 2 Characteristics of the z -score distributions of ultrasound biometry measurements ($n = 227$) performed by Examiner 2

Measurement	Percentile			Mean \pm SD	p*
	5 th	50 th	95 th		
Head circumference	-1.7	-0.1	1.5	-0.027 \pm 1.069	0.7066
Abdominal circumference	-1.4	0.3	2.2	0.344 \pm 1.142	< 0.0001
Femur length	-1.8	-0.4	1.3	-0.350 \pm 0.944	< 0.0001

All data shown as z -scores. * t -tests comparing observed z -score mean with expected value of zero.

by detecting errors and implementing adequate corrective measures¹⁸.

The CUSUM technique appears to be a promising method for application in routine clinical practice. The accumulation of recent ultrasound results allows the detection of even small permanent errors, which otherwise are easily missed²³, and the appropriate corrective measures can be implemented immediately. Sonographic fetal weight estimation is an optimal variable for evaluation by the examiner because no subjective bias is expected, as the actual weight of an infant is known only after birth. To rule out a selection bias in the analyzed population, CUSUM curves of single fetal biometry determinants were compared with

CUSUM curves of weight estimation of the same fetuses.

When applying the CUSUM method, the target and the properties should be defined prospectively¹⁸. Usually unacceptable failure rates are set to be less stringent at the beginning of the learning process, and once these initial rates have been achieved the CUSUM chart might be recalculated according to stricter standards²⁰. The same principle could be used for setting control limits.

At present, in different medical fields, approval of competence is based on a certain number of procedures that the physician has to perform¹⁰. Our findings, however, show that, at least for sonographic fetal weight estimation, the number of procedures required

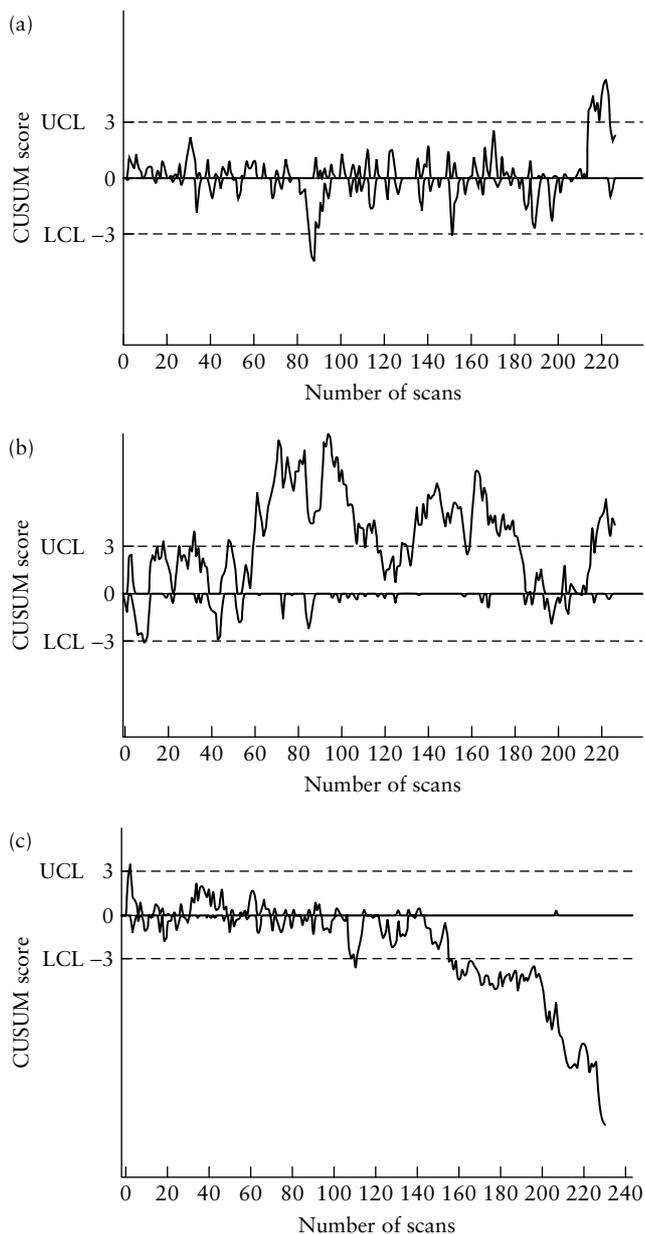


Figure 3 Double cumulative summation (CUSUM) charts demonstrating the accuracy of measurements of fetal head circumference (a), fetal abdominal circumference (b) and fetal femur length (c) for Examiner 2. UCL, upper control limit; LCL, lower control limit.

for proficiency to be achieved varies greatly between individuals (Figure 1). It has been recommended that trainees should complete 200 scans over a period of 3 years to achieve competence²⁴. According to protocols of the German Medical Society of Ultrasound²⁵, the examiner has to perform more than 300 ultrasound scans to prove sufficient competence.

Predanic *et al.*²⁶ have shown that the level of accuracy in fetal weight estimation increases with the number of examinations. Nevertheless, these investigators could not identify the exact point in time at which the improvement occurred or at what time corrective action should have been started²⁶. Thus CUSUM, being a longitudinal time-weighted control chart, provides prompt information

when the process is getting out of control and also allows continuing quality control for trained physicians, especially when new equipment or a different anatomical approach is implemented²⁷.

The LC-CUSUM chart presents the dynamic of the errors of every single examiner (Figure 1). It is sufficient to use the LC-CUSUM to observe the time when sonographic quality improves and proficiency has been achieved. The double CUSUM chart, on the other hand, also allows detection of systematic errors of fetal weight estimation in contrast to aggregate methods, where poor runs can be compensated and hidden by the existence of excellent results at other time periods²³.

Measurements of individual fetal biometric parameters can be assessed to identify the source of the errors leading to inaccurate fetal weight estimation. We therefore suggest evaluation of the competence of clinicians longitudinally using the z -scores of fetal HC, AC and FL analyzed by the CUSUM technique. This approach focuses on the direct cause and exact time point of continuously occurring systematic errors of sonographic fetal weight estimation.

Salomon *et al.*⁸ demonstrated the power of the z -scores system for evaluating the quality of sonographic fetal biometry. This system provides a valuable tool for identifying insufficient performance as shown by the abnormal distribution of z -scores of fetal biometric parameters. Indeed, fetuses with measurements located at the extremes of the normal distribution curve are considered to be at higher risk, and wrongly shifted distributions may lead to ineffective screening and even inappropriate action being taken in these pregnancies. Although this method allows statistical evaluation of the overall process, the disadvantage is that it is not able to detect the specific time points at which systematic errors occur. The identification of the exact time point, however, would be of great practical value as it would allow the direct implementation of corrective actions – or at least for a decision to be made regarding whether to continue or modify the monitored process¹⁸.

The results of Salomon *et al.*⁸ correspond with our findings and show that owing to the overestimation of fetal AC and underestimation of FL many fetuses are allocated as false-negatives with respect to their birth weight. We agree that it is essential to correctly identify fetuses with borderline biometric measurements, as these findings are often associated with intrauterine growth restriction or macrosomia⁸. However, correct measurements are difficult to achieve without an exact failure analysis of separate biometric parameters.

Our retrospective analysis of one examiner showed that several periods of fetal AC overestimation did not have a major influence on fetal weight measurement. In contrast, systematic underestimation of FL starting with the 154th attempt (Figure 3c) was associated with a large negative impact on fetal weight estimation. In fact, even the short periods of AC and HC overestimation during the final study period (Figure 3a,b) did not balance this until the end of the study. This finding can be explained by the original formulae of Hadlock

et al.²⁸, which include regression coefficients for FL that are greater than those for the other fetal parameters included.

The number of measurements analyzed in this study may be insufficient for an exact evaluation of the impact of single determinants on the accuracy of weight estimation; nevertheless, we aimed to demonstrate a trend based on systematic errors in the measurement of single fetal biometric parameters.

A prospective study would be able to evaluate the value of the CUSUM technique as a continuous audit system, allowing urgent real time feedback to improve the quality of fetal ultrasound biometry of beginners as well as of experts in ultrasound diagnostics. Furthermore, the CUSUM technique may be valuable for other sonographic assessments in perinatal medicine.

REFERENCES

- Salomon LJ, Bernard JP, Ville Y. [Quality control of prenatal ultrasound. A role for biometry.] *Gynecol Obstet Fertil* 2006; **34**: 683–691.
- Platz E, Newman R. Diagnosis of IUGR: traditional biometry. *Semin Perinatol* 2008; **32**: 140–147.
- Scioscia M, Vimercati A, Ceci O, Vicino M, Selvaggi LE. Estimation of birth weight by two-dimensional ultrasonography: a critical appraisal of its accuracy. *Obstet Gynecol* 2008; **111**: 57–65.
- Dudley NJ, Chapman E. The importance of quality management in fetal measurement. *Ultrasound Obstet Gynecol* 2002; **19**: 190–196.
- Weerasinghe S, Mirghani H, Revel A, Abu-Zidan FM. Cumulative sum (CUSUM) analysis in the assessment of trainee competence in fetal biometry measurement. *Ultrasound Obstet Gynecol* 2006; **28**: 199–203.
- Salomon LJ, Bernard JP, Duyme M, Doris B, Mas N, Ville Y. Feasibility and reproducibility of an image-scoring method for quality control of fetal biometry in the second trimester. *Ultrasound Obstet Gynecol* 2006; **27**: 34–40.
- Dudley NJ. A systematic review of the ultrasound estimation of fetal weight. *Ultrasound Obstet Gynecol* 2005; **25**: 80–89.
- Salomon LJ, Bernard JP, Ville Y. Analysis of Z-score distribution for the quality control of fetal ultrasound measurements at 20–24 weeks. *Ultrasound Obstet Gynecol* 2005; **26**: 750–754.
- Bowles TA, Watters DA. Time to CUSUM: simplified reporting of outcomes in colorectal surgery. *ANZ J Surg* 2007; **77**: 587–591.
- Naik VN, Devito I, Halpern SH. Cusum analysis is a useful tool to assess resident proficiency at insertion of labour epidurals. *Can J Anaesth* 2003; **50**: 694–698.
- Wolfe R, Bolsin S, Colson M, Stow P. Monitoring the rate of re-exploration for excessive bleeding after cardiac surgery in adults. *Qual Saf Health Care* 2007; **16**: 192–196.
- Yap CH, Colson ME, Watters DA. Cumulative sum techniques for surgeons: a brief review. *ANZ J Surg* 2007; **77**: 583–586.
- Bolsin S, Colson M. The use of the Cusum technique in the assessment of trainee competence in new procedures. *Int J Qual Health Care* 2000; **12**: 433–438.
- Williams SM, Parry BR, Schlup MM. Quality control: an application of the cusum. *BMJ* 1992; **304**: 1359–1361.
- 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.
- Campbell S, Thoms A. Ultrasound measurement of the fetal head to abdomen circumference ratio in the assessment of growth retardation. *Br J Obstet Gynaecol* 1977; **84**: 165–174.
- Campbell S, Wilkin D. Ultrasonic measurement of fetal abdomen circumference in the estimation of fetal weight. *Br J Obstet Gynaecol* 1975; **82**: 689–697.
- Biau DJ, Porcher R, Salomon LJ. CUSUM: a tool for ongoing assessment of performance. *Ultrasound Obstet Gynecol* 2008; **31**: 252–255.
- Parry BR. CUSUM: timely and trendy. *ANZ J Surg* 2007; **77**: 523.
- de Oliveira Filho GR. The construction of learning curves for basic skills in anesthetic procedures: an application for the cumulative sum method. *Anesth Analg* 2002; **95**: 411–416.
- Kurmanavicius J, Wright EM, Royston P, Wisser J, Huch R, Huch A, Zimmermann R. Fetal ultrasound biometry: 1. Head reference values. *Br J Obstet Gynaecol* 1999; **106**: 126–135.
- Kurmanavicius J, Wright EM, Royston P, Zimmermann R, Huch R, Huch A, Wisser J. Fetal ultrasound biometry: 2. Abdomen and femur length reference values. *Br J Obstet Gynaecol* 1999; **106**: 136–143.
- Sibanda T, Sibanda N. The CUSUM chart method as a tool for continuous monitoring of clinical outcomes using routinely collected data. *BMC Med Res Methodol* 2007; **7**: 46.
- Dresang LT, Rodney WM, Dees J. Teaching prenatal ultrasound to family medicine residents. *Fam Med* 2004; **36**: 98–107.
- Eichhorn KH, Schramm T, Bald R, Hansmann M, Gembruch U. [DEGUM grade I quality standards in obstetric ultrasound diagnosis during the 19th – 22nd week of pregnancy.] *Ultraschall Med* 2006; **27**: 185–187.
- Predanic M, Cho A, Ingrid F, Pellettieri J. Ultrasonographic estimation of fetal weight: acquiring accuracy in residency. *J Ultrasound Med* 2002; **21**: 495–500.
- Kestin IG. A statistical approach to measuring the competence of anaesthetic trainees at practical procedures. *Br J Anaesth* 1995; **75**: 805–809.
- Hadlock FP, Harrist RB, Carpenter RJ, Deter RL, Park SK. Sonographic estimation of fetal weight. The value of femur length in addition to head and abdomen measurements. *Radiology* 1984; **150**: 535–540.

SUPPORTING INFORMATION ON THE INTERNET

The following supporting information may be found in the online version of this article:

Appendix S1 Derivation of the main statistical formulae used for construction of the CUSUM curves, including their predefined boundary lines.