European Journal of Radiology 91 (2017) 35–40

Contents lists available at ScienceDirect

European Journal of Radiology

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

Repeatability of estimated fetal weight: Comparison between MR imaging versus 2D ultrasound in at- and near-term patients

Caroline Kadji [a], Margaux F. Camus [a], Elisa Bevilacqua [a], Mieke M. Cannie [b,c], Teresa Cos Sanchez [a], Jacques C. Jani [a,*]

[a] Department of Obstetrics and Gynecology, University Hospital Brugmann, Université Libre de Bruxelles, Brussels, Belgium
[b] Department of Radiology, University Hospital Brugmann, Université Libre de Bruxelles, Brussels, Belgium
[c] Department of Radiology, UZ Brussel, Vrije Universiteit Brussel, Brussels, Belgium

ARTICLE INFO

Article history:
Received 20 December 2016
Received in revised form 14 February 2017
Accepted 14 March 2017

Keywords:
Estimation of fetal weight
Fetal body volume
Magnetic resonance imaging
2D ultrasound
Repeatability
Operator experience

ABSTRACT

Introduction: Our aim was to evaluate the intra- and inter-observer variability and the impact of operator experience on the estimation of fetal weight (EFW) as measured by 2-dimensional ultrasound (2D-US) and magnetic resonance (MR) imaging.
Material and methods: We estimated fetal weight in 46 singleton pregnancies at 35.6–41.4 weeks gestation using 2D-US according to the Hadlock formula and using MR imaging according to the equation developed by Baker. Each examination was performed twice, once by an inexperienced operator and once by an experienced operator. The MR-EFW was derived from the planimetric measurement of fetal body volume (FBV) using an assisted semi-automated method. Intra- and inter-observer variability was evaluated by Bland-Altman analysis. Regression analysis was used to investigate the effect of maternal BMI, delivery weight, diabetes and fetal gender on the differences in US-EFW between the inexperienced and experienced operators.
Results: US-EFW showed higher intra-observer variability than MR-EFW, irrespective of operator experience. The 95% limits of agreement of MR were narrower compared with those of the US measurements. Similarly, US-EFW showed higher inter-observer variability than MR-EFW. MR-EFW improvement over 2D-US for the limits of agreement was 77.9% for intra-observer variability and 74.5% for inter-observer variability. Regression analysis showed that the differences between US-EFW measurements were not related to any of the tested variables.
Conclusions: Operator experience has a marginal impact on the variability of US-EFW and no impact on MR-EFW variability. The variability in US-EFW measurements is unpredictable.

1. Introduction

Since the introduction of ultrasound (US) in the 1960s, attempts to estimate fetal weight have led to US becoming part of obstetrical practice. In 1985, Hadlock et al. [1] developed a formula based on prenatal parameters such as the biparietal diameter (BPD), abdominal circumference (AC) and femur length (FL); the combination of which allows the indirect US estimation of fetal weight (US-EFW). Since then, many authors have considered the US-EFW as described by Hadlock et al. to be inaccurate. As a consequence, attempts have been made to improve the accuracy of the measurement using new formulas, but despite these efforts, no successful methods have been described [1–6]. In 2012, Kehl et al. [7] evaluated the biometry of 628 newborns within 1 h of delivery to investigate the limits of accuracy of US-EFW. They showed that a good sonographic weight formula will have a standard deviation (SD) of approximately 7%, and 80% of cases will be within a 10% range of discrepancy. The authors further concluded that improvement in EFW can no longer be achieved on the basis of conventional biometric parameters using two-dimensional (2D) US [7].

In 1994, Baker et al. [8] introduced magnetic resonance (MR) imaging for EFW, whereby fetal body volume (FBV) is measured, and MR-EFW is derived from a simple formula. Since then, many authors have shown that MR-EFW is more accurate than 2D ultrasound [9–11]. Unfortunately, this promising method has not been

Abbreviations;: EFW, estimated fetal weight; 2D-US, two-dimensional ultrasound; MR, magnetic resonance; FBV, fetal body volume.

* Corresponding author at: Department of Obstetrics and Gynecology University Hospital Brugmann, Université Libre de Bruxelles Place A. Van Gehuchten 4, 1020 Brussels Belgium.
E-mail address: jackjani@hotmail.com (J.C. Jani).

http://dx.doi.org/10.1016/j.ejrad.2017.03.005
0720-048X/

introduced in clinical practice, partly due to the time-consuming process of FBV measurement [12].

In a recent study, a semi-automated method for the planimetric measurement of FBV was applied to 36 female patients prior to delivery at a median gestational age of 38.6 weeks [13]. The MR-EFW was derived from the FBV measurement and compared with the actual birth weight. The study demonstrated that MR-EFW could be obtained with a relative error of approximately 2.5% when compared with the actual birth weight and, more importantly, with a measurement time consistently below 5 min [13]. The described method for MR-EFW is, therefore, quite promising and has the potential for clinical implementation. However, another step is needed before clinical implementation can demonstrate that the described method is not only accurate but also reproducible, including in inexperienced hands, which will allow the technique to be easily integrated into clinical practice. While many US studies have evaluated the reliability of 2D and 3D EFW measurements, including the impact of operator experience, no such studies have been conducted for MR-EFW [14–16].

Thus, the purpose of our study was to evaluate the intra- and inter-observer variability and the impact of operator experience on the variability of EFW as measured by 2D US and MR imaging in at- and near-term patients.

2. Material and methods

2.1. Study design

This study was approved by the local ethics committee (CE 2015/17; March 10th, 2016), and all patients gave written informed consent. This was a single-institution prospective study conducted at the Department of Radiology and the Fetal Medicine Unit of the University Hospital Brugmann, Université Libre de Bruxelles, Brussels, Belgium. During the study period between March and September 2015, 46 patients with singleton pregnancies scheduled for a planned delivery, such as the induction of labor or an elective cesarean section, underwent an MR examination for fetal weight estimation. Within 3 h of the MR examination, a prenatal US examination was also performed. The MR and US examinations were performed prior to admission to the labor ward. In all patients, gestational age was dated based on the first-trimester scan.

2.2. Prenatal US examination

All prenatal US examinations were carried out using transabdominal sonography (RAB 4-8L probe, Voluson E8; 4CRS probe, Voluson e; GE Medical Systems, Milwaukee, WI) by a fellow in fetal medicine (M.C.) with 1 year of experience in 2D ultrasound, hereafter referred to as the inexperienced operator, and by a maternal-fetal medicine specialist (E.B.) with 6 years of experience in 2D ultrasound, hereafter referred to as the experienced operator. US-EFW was determined according to Hadlock et al. [1] based on the measurements of biparietal diameter (BPD), abdominal circumference (AC) and femoral length (FL). For each patient, each operator performed two measurements for US-EFW with 10 min in between each measurement and were blinded to each other's results.

2.3. MR examination

MR imaging was performed using a clinical 1.5 T whole-body unit (Siemens Magnetom Avanto, Erlangen, Germany) with a gradient field strength of 45 mT/m. The patients were scanned in the supine position with a combination of a six-channel phased-array body and six elements of the spine coil positioned over the lower pelvic area. Following a scout scan to gather information about the orientation of the fetus, we recorded T2-weighted images using fast imaging with steady-state free procession (TrueFISP) sequences in the fetal sagittal plane. Between 10 and 32 adjacent slices on average were adjusted according to fetal size with a 4 mm slice thickness, an intersection gap of 4 and 20 mm, a field-of-view of $420 \times 336\,mm^2$, a matrix of 166×256, a TR (repetition time)/TE (echo time) = 4.65 ms/2.33 ms, a resulting voxel resolution of $2.1 \times 1.6 \times 4\,mm^3$ and a bandwidth of 399 Hz/pixel. Sequences that were degraded by fetal motion were repeated with the same parameters. Placing the patient into the magnet was 5 min or less. The total examination time was 3 min or less.

2.4. MR imaging planimetry using the assisted method

Total FBV was delineated using a home-built user interface programmed using the MATLAB (matrix laboratory) environment (MATLAB 2013a, the MathWorks, Inc., Natick, MA, US) as previously described [13]. Briefly, the program automatically searched for voxels that potentially represented the skin of the fetus by selecting only the low intensity pixels surrounded by two layers of higher intensity. When the operator selected a point in the fetus border, the program automatically searched for a path of skin pixels connected to the selected points and added new points along the path. The automatic correction was deactivated at each step to allow the operator to manually refine the region (Fig. 1).

Total FBV planimetric measurements were performed by a fellow in fetal medicine (M.C.) with less than 1 month of experience in planimetric FBV measurements, hereafter referred to as the inexperienced operator, and a maternal-fetal medicine specialist (C.K.) with more than 2 years of experience in planimetric FBV measurements, hereafter referred to as the experienced operator. For each patient, each operator performed two FBV measurements for MR-EFW with at least 2 weeks in between the measurements. The time needed to perform planimetric FBV measurements was 5 min or less.

The MR-EFW was calculated based on the equation $0.12 + 1.031 \times FBV\ (mL) = MR\ weight\ (g)$ as developed by Baker et al. [8].

2.5. Statistical analysis

Inter-observer variability (95% CI) (M.C. and C.K.; one delineation each) and intra-observer variability (95% CI) (M.C. and C.K., respectively two delineations each with at least 1 month in between) were assessed for FBV by using the intra-class correlation coefficient. Pearson's correlation coefficient was used to measure the strength of association between the 2 sets of variables.

Comparison of the two sets of measurements for the evaluation of intra- and inter-observer variability was further performed by calculating the following parameters as described by Bland and Altman and graphically presented as Bland-Altman scatter plots [17,18]:

- Bias, i.e., the mean of the proportionate difference (the difference between the first and second measurements of M.C., E.B., and C.K. or the difference between M.C. and either one of the experienced operators (E.B./C.K.) divided by the average of those measurements).
- Proportionate 95% limits of agreement, i.e., 1.96-times the standard deviation of the mean of the proportionate difference (difference between two measurements of the same 2D US-EFW or MR-EFW divided by the average of these measurements expressed in%).

Regression analysis was used to investigate the effect on the proportionate difference of US-EFW between the inexperienced (M.C.)