The Reliability of Fetal MRI in the Assessment of Brain Malformations

Nora Frick^a Claudius Fazelnia^a Kathrin Kanzian^b Wolfgang Hitzl^c
Thorsten Fischer^a Rosemarie Forstner^b Gerhard Bogner^a

^aDepartment of Obstetrics and Gynecology, ^bDepartment of Radiology, and ^cResearch Institute for Biostatistics, Paracelsus Medical University, Salzburg, Austria

Key Words

Fetal magnetic resonance imaging · Fetal ultrasound · Fetal diagnosis · Abnormalities · Fetal brain malformations

Abstract

Objectives: To assess the inter- and intraobserver reliability of different fetal MRI measurements in cases of fetal brain malformations and to examine the concordance between ultrasonography (US) and MRI findings. **Methods:** Fetal brain MRIs and US findings of 56 pregnant women were retrieved from the institutional database. Standardized fetal brain MRI measurements were performed by 4 observers, and the inter- and intraobserver reliability was determined. Additionally, US and MRI findings were retrospectively compared. **Results:** The interobserver intraclass correlation coefficient (ICC) was above 0.9 for the cerebellum and posterior horn of the lateral ventricle. The measurements regarding the third ventricle (0.50), the fourth ventricle (0.58), and the corpus callosum (0.63) showed poor reliability. Overall, the intraobserver reliability was greater than the interobserver reliability. US and MRI findings were discordant in 29% of the cases with MRI rendering an extended diagnosis in 18%, a change of diagnosis in 3.6%, and excluding pathological findings suspected on US in 7.1%. **Conclusions:** Fetal MRI is a valuable complement to US in the investigation of fetal brain malfor-

mations. The reliability of most parameters was high, except for the measurements of the third and fourth ventricles and the corpus callosum.

© 2014 S. Karger AG, Basel

Introduction

Brain malformations rank among the most common fetal abnormalities with about one out of 100 live births featuring a central nervous system abnormality [1].

Ultrasonography (US) is the first-line screening tool to diagnose fetal pathologies or to detect aberrations in fetal development in early pregnancy [2]. In inconclusive or suspicious US findings, fetal MRI can be used as an additional diagnostic tool in the second and third trimester of gestation in order to confirm or refute the findings suspected on US, to gain further information of potential malformations, and to search for associated pathologies [3–5].

Fetal brain assessment is the most common indication of fetal MRI because of the high sensitivity of the MRI to visualize the structure and microstructure of the brain tissue [6]. According to recent studies, fetal MRI appears to be a valuable addition to US in the detection of fetal brain malformations [2, 7, 8]. US is an operator skill-dependent

Table 1. Fetal MR imaging parameters for T2 sequences

Sequence	Magnet, T	TR, ms	TE, ms	Slice thickness, mm	FOV, mm	Matrix	Acquisition time, s	Slices
T2 HASTE	1	2,100	90	5	330:300	138×256	40	16–25
T2 TRUFI	1	4.07	2.04	5	400:400	205×256	35	16–20
T2 TSE SSH	3	3,122 ¹	140	4	275:225	284×186	40	16–25
T2 BTFE	3	4.18 ¹	2.09 ¹	4	280:300	284×186	34	15–20
T2 TSE SSH	1.5	980 ¹	90	4	280:300	224×148	31	16–25
T2 BTFE	1.5	4.28 ¹	2.14 ¹	4	350:350	268×215	31	16–20

¹ Shortest.

and dynamic modality and therefore subject to a significant quality variation [2, 3, 8]. MRI quality can vary as well depending on motion artefacts [9] and the imaging technique, including the accuracy of imaging plane setting, which can alter the visibility of pathoanatomical structures [8]. Similar to fetal US brain measurements, standardized linear measurements for the evaluation of fetal brain on MRI have been proposed [10]. While US reliability in the context of fetal brain measurements has been examined extensively [11–14], little information exists regarding the inter- and intraobserver reliability of MRI-based measurements of fetal brain malformations.

The main purpose of this study was to determine the interobserver and intraobserver reliability of different biometrical MRI measurements in various cases of fetal brain malformations. Additionally, the concordance and discrepancies between US and fetal MRI findings were examined.

Materials and Methods

Study Population

All fetal brain MR images performed between 2003 and 2013 were retrieved from the institutional database. A total of 56 pregnant women with abnormal or inconclusive findings of the fetal brain on US and consecutive fetal brain MRI examinations were included in the analysis. 54 of them underwent US examination at the authors' Department and 2 US results were collected from peripheral hospitals. All fetal brain MR images were performed at the institutional Department of Radiology with prior knowledge of the US results. The patients did not have contraindications to MRI. The sex of the fetuses was not taken into account in this study.

The mean maternal age was 31.2 ± 5.7 (SD) years with a range from 18 to 42 years. The MRI-necessitating US examination was obtained at a mean pregnancy age of $28 + 3$ gestational weeks ± 5.8 (range, $18 + 1$ to $37 + 2$ weeks), and the MRI examinations were performed at a mean pregnancy age of $29 + 6$ gestational weeks $\pm$

5.8 (range, $19 + 1$ to $40 + 4$ weeks). The median time interval between the two different examinations was 6 ± 7.2 days (range, 0–36 days). All fetal brain MRI and US findings were retrospectively collected and compared. In order to analyze the inter- and intraobserver reliability of fetal brain malformations, the below-listed linear measurements were performed independently by 4 observers and repeated after a minimum time interval of 4 weeks.

The local ethical committee approval was obtained for this study (415-EP/73/259-2013).

MRI Technique

The fetal MR images were obtained on a 1.0-tesla unit (Siemens Harmony, Berlin, Germany; $n = 14$) until 2008, and from then on the examination was performed on a 3.0-tesla Gyroscan (Achieva; Philips, Amsterdam, The Netherlands; $n = 35$) and 1.5-tesla Gyroscan (Ingenia; Philips; $n = 5$).

In all cases, the mother was examined in a supine position using a body coil. Fasting before the MRI examination was not required nor were spasmolytic agents or sedatives administered. After a scout scan to localize the fetus, three T2-weighted sequences were performed in the coronal, transaxial and sagittal planes followed by a T1-weighted sequence in the transaxial plane of section. To obtain orthogonal orientation of the fetal brain, double angulation was required. In all patients, the MRI was focused on the fetal brain.

In poor imaging quality, complementary TRUEFISP (true fast imaging with steady-state precession) was performed, typically in the sagittal plane. Additional sequences including FLAIR (fluid-attenuated inversion recovery), FLASH (fast low angle shot), or DWI (diffusion-weighted imaging) were rarely used.

Technical details of T2-weighted imaging sequences used for assessment of the measurements are enlisted in table 1.

MRI Measurements

The linear fetal MRI measurements analyzed for their inter- and intraobserver reliability included the posterior and anterior horn of the lateral ventricles, the third and fourth ventricles, the corpus callosum, the vermis and the cerebellum. The parameters were measured according to the method described by Garel et al. [10].

The diameter of the posterior horn of the lateral ventricles was measured on an axis perpendicular to the long axis of the ventricle,

at the level of the atria, in the coronal and axial planes. In ventricular asymmetry, the measurements of the larger ventricle were used for analysis. The diameter of the third ventricle was also assessed in the coronal and axial planes. The anteroposterior diameter of the fourth ventricle was measured on the midline sagittal plane, from the median parts of its roof to its floor. The presence or absence of corpus callosum was determined in the coronal plane. When present, the length of the corpus callosum was measured on the midline sagittal plane, from the genu to the splenium. The height of the vermis was established on the midline sagittal plane, depicting the greatest height of the vermis as parallel as possible to the axis of the brain, and the transverse cerebellar diameter on the posterior coronal plane at the level of the atria [10] (fig. 1).

The fetal MR images were presented to 4 independent observers: 2 radiologists (examiner 1 and 2) and 2 gynecologists (examiner 3 and 4). The examiners had different levels of experience. Examiner 1 was a senior radiologist experienced in fetal MRI for more than 15 years. Examiner 2 was a 2nd-year radiology resident trained in reading MRI, who did not have any specialization in reading fetal MRI. Examiner 3 was a senior gynecologist specialized in prenatal diagnostics, and examiner 4 was a final-year gynecology resident trained in prenatal ultrasound. The gynecologists did not have any experience in diagnosing fetal MRI. The observers were instructed regarding the measurement criteria prior to performing the measurements. The measurements were carried out without the presence and support of experienced imagers. All observers had to select a suitable imaging slice and perform the measurements according to the above-mentioned criteria as accurately as possible. The observers had no access to the US findings or other patient information while completing the measurements. Data were recorded independently by each of the examiners on an Excel worksheet; occurring difficulties were noted, and the image quality was defined. The same measurements were repeated after a minimum time interval of 4 weeks for intraobserver reliability analysis.

All measurements were carried out in IMPAX EE (Agfa Healthcare, Mortsel, Belgium). 47 MR images were digital, and measurements were shown in millimeters; 9 were scanned-film imports with measuring scale. To complete the measurements, zooming, rotating, and contrast scaling was allowed, but no other effects were used to influence image quality. Poor image quality was subjectively defined as vague margin delineation due to fetal motion artefacts and asymmetry of the acquisition plane due to fetal head rotation. Additionally, severe malformations hampering and replacing the structures to be measured were noted. When one observer considered a parameter of a patient not to be measurable, this linear measurement was excluded for all observers in the reliability analysis. Two out of 56 cases were completely excluded from the reliability analysis as the required measurements were impeded by a tumorous lesion and hydrocephalus internus superseding the brain structures in one case, and a twin pregnancy with insufficient depiction of the neuroanatomic structures of the fetus with a meningoencephalocele in the other case.

US Examination and Comparison of Findings

The US examinations, which were completed at the authors' institution ($n = 54$), were performed with transabdominal and transvaginal volumetric transducers (Voluson 8, GE HealthcareUK). An abdominal 2D Curved Array 4CD transducer with 3.1 MHz and a RAM38 3D imaging transducer with 4.2 MHz were

used. All patients underwent a transabdominal examination. The brain was studied in the axial, coronal, and sagittal planes. The US imaging was performed by various experienced obstetricians specialized in prenatal ultrasound, who all detected a pathological finding requiring further diagnostics. The US completed within the closest time frame from the fetal MRI was used to compare the results and pathologies. US measurements were completed on the basis of the checklist of the prenatal screening program ViewPointPIA (LB-systems, Vienna, Austria).

Pathologies for comparison of US versus MR images included hydrocephalus, ventriculomegaly, aqueductal stenosis, vermis hypoplasia/agenesis with or without associated Dandy-Walker malformation, cerebellar hypoplasia/agenesis, agenesis of the corpus callosum (ACC), missing septum pellucidum, cysts, mega cisterna magna, hemorrhages, ventricular asymmetry, encephalocele/meningocele, holoprosencephaly, missing falx, vein of Galen aneurysm, and schizencephaly (table 2).

Mild ventriculomegaly was defined as an atrial diameter between 10 and 12 mm, moderate ventriculomegaly between 12.1 and 14.9 mm, and severe ventriculomegaly, called hydrocephalus, as a diameter of ≥ 15 mm [15, 16].

Mega cisterna magna was defined as a dilatation of the fluid-filled cerebellar space exceeding 10 mm [17].

ACC was determined by the absence of a corpus callosum including typical signs as the teardrop configuration of the occipital horns of the ventricles in the coronal plane, the sunburst sign in the sagittal plane and cysts in the posterior cranial fossa [18].

Statistical Analysis

Data were carefully checked for normality and outliers. A two-way mixed ANOVA model was used to compute intra- and inter-class correlation coefficients (based on consistency) together with 95% confidence intervals (CIs). The F test was used to compare means among different observers or measurements.

A p value less than 5% was considered to indicate a significant effect. All statistical analyses were performed using PASW 19 (IBM SPSS Statistics for Windows, Armonk, N.Y., USA), Mathematica 7 (Wolfram Research, Inc., Champaign, Ill., USA), and STATISTICA 10 (StatSoft, Tulsa, Okla., USA).

According to Portney and Watkins [19], in this study an intra-class correlation coefficient (ICC) smaller than 0.75 indicated poor to moderate reliability, an ICC between 0.75 and 0.90 good reliability, and an ICC above 0.90 excellent and therefore reasonable reliability for clinical measures.

Results

Most common fetal brain malformations detected on MRI in our study population included hydrocephalus ($n = 24$, 43%), ventriculomegaly ($n = 13$, 23%), and aqueductal stenosis ($n = 12$, 21%; table 2).

The interobserver reliability was the highest for the cerebellar diameter measurements in the coronal (0.97, 95% CI 0.96–0.98) and axial (0.97, 95% CI 0.95–0.98) planes and the lowest for the diameter of the third ventricle in the coronal plane (0.50, 95% CI 0.35–0.65; ta-

Fig. 1. Example of the measurements performed to determine fetal cerebral pathologies on T2-weighted brain MR images of a fetus with ventriculomegaly at gestational week 32. On coronal views of the fetal brain, the diameter of the posterior horn of the lateral ventricles, the transverse cerebellar diameter (a), the diameter of the anterior horn of the lateral ventricles (b), and the diameter of the third ventricle (c) are measured. On axial views, the diameter of

the posterior horn of the lateral ventricles (d), the diameter of the anterior horn of the lateral ventricles, the diameter of the third ventricle (e), and the transverse cerebellar diameter (f) are measured. On the midline sagittal view, the length of corpus callosum (g), the height of the vermis (h), and the anteroposterior diameter of the fourth ventricle (i) are measured.

Table 2. Cerebral pathologies in the study population identified on fetal MRI

Cerebral pathologies	n	%
Hydrocephalus	25	43
Ventriculomegaly	13	23
Aqueductal stenosis	12	21
Vermis hypoplasia/agenesis	10	18
Cerebellar hypoplasia	9	16
ACC	9	16
Missing septum pellucidum	8	14
Cysts	7	13
Mega cisterna magna	6	11
Hemorrhage	5	8.9
Ventricular asymmetry	4	7.1
Encephalocele/meningocele	4	7.1
Holoprosencephaly	3	5.4
Missing falx	2	3.6
Vein of Galen aneurysm	1	1.8
Schizencephaly	1	1.8

ble 3). Similarly, average intraobserver reliability was the highest for the cerebellar diameter measurements (0.99) and the lowest for the measurements of the third ventricle in the coronal plane (0.69; table 4).

Regarding the interobserver reliability, 3 measurements achieved an ICC higher than 0.9 (table 3), and intraobserver ICC was higher than 0.9 for 8 measurements (table 4).

The measurements of the length of the corpus callosum featured a poor interobserver reliability (0.63, 95% CI 0.45–0.79), but achieved a high average intraobserver reliability (0.94). The interobserver concordance in the detection of ACC was 63%, while the intraobserver agreement showed an average of 71%.

Generally, the radiologists did not reach higher interobserver reliabilities than the gynecologists except for the measurements of the third ventricle in the coronal plane (0.84, 95% CI 0.72–0.91 vs. 0.49, 95% CI 0.23–0.68; table 5).

In the analysis of good versus poor quality images, the interobserver reliability of the different measurements showed a clear trend towards more reliable measurements in images of higher quality. However, only the measurement of the cerebellar diameter in the coronal and axial planes showed statistically significant differences with no overlapping of the CIs (table 6). No clear trend in the reliability of the measurements was detected when performing a subgroup analysis regarding the utilized magnetic field strength. Only the measurement of the

Table 3. Interobserver reliability in terms of ICCs

Measurement parameters	ICC (95% CI)	n	%
Lateral ventricle (post. horn, cor.)	0.91 (0.86–0.94)	54	100
Lateral ventricle (post. horn, ax.)	0.87 (0.81–0.92)	52	96
Lateral ventricle (ant. horn, cor.)	0.78 (0.69–0.86)	49	91
Lateral ventricle (ant. horn, ax.)	0.79 (0.70–0.86)	49	91
3rd ventricle (cor.)	0.50 (0.35–0.65)	42	78
3rd ventricle (ax.)	0.78 (0.67–0.86)	40	74
4th ventricle (sag.)	0.58 (0.41–0.74)	31	57
Corpus callosum length (sag.)	0.63 (0.45–0.79)	18	33
Vermis height (sag.)	0.87 (0.81–0.92)	45	83
Cerebellum (cor.)	0.97 (0.96–0.98)	50	93
Cerebellum (ax.)	0.97 (0.95–0.98)	49	91

n/% = Absolute/relative number of measurable cases. ICC <0.75 = poor, 0.75–0.9 = moderate, >0.9 = excellent reliability.

posterior horn of the lateral ventricle in the axial plane showed a significantly higher reliability on the 3-tesla images than on the 1-tesla images (0.89, 95% CI 0.88–0.96 vs. 0.68, 95% CI 0.44–0.86). However, in the 3-tesla group, 2 cases with poor image quality and non-measurable posterior horn had to be excluded from the reliability analysis.

Regarding the comparison of both imaging modalities, fetal MRI did not alter the primary suspected US diagnosis in 40 cases (71%). The 16 cases (29%) in which fetal MRI provided information differing from the US findings included 10 cases (18%) in which MRI significantly extended the diagnosis and affected clinical decision-making, 2 cases (3.6%) in which a different primary pathology was detected on MRI, and 4 cases in which no abnormal findings were detected on MRI (7.1%). The extension of diagnoses on fetal MRI compared to US mostly included cases with initially detected ventriculomegaly or hydrocephalus on US (n = 7, 12.5%) and additional fetal MRI findings including hemorrhages, lesions of brain parenchyma, ventricular asymmetry, missing septum pellucidum, septo-optic dysplasia, enlarged third and/or fourth ventricle, aqueductal stenosis, ACC and vermis hypoplasia. The two primary diagnoses altered by fetal MRI included a suspected pontocerebellar dysplasia, which was identified as an enlarged mega cisterna magna, and a suspected enlarged third ventricle, which was found to be a cavum septum pellucidum cyst on fetal MRI. Previously suspected pathological US findings, which were

Table 4. Average intraobserver reliability

Measurement parameters	Observer 2 ICC (95% CI)	Observer 3 ICC (95% CI)	Observer 4 ICC (95% CI)	Average	n	%
Lateral ventricle (post. h., cor.)	0.97 (0.94–0.98)	0.91 (0.85–0.95)	0.93 (0.88–0.96)	0.94	54	100
Lateral ventricle (post. h., ax.)	0.97 (0.95–0.98)	0.94 (0.90–0.97)	0.93 (0.88–0.96)	0.95	53	98
Lateral ventricle (ant. h., cor.)	0.84 (0.74–0.91)	0.87 (0.78–0.92)	0.91 (0.85–0.95)	0.87	48	89
Lateral ventricle (ant. h., ax.)	0.92 (0.87–0.95)	0.92 (0.87–0.95)	0.94 (0.90–0.97)	0.93	48	89
3rd ventricle (cor.)	0.79 (0.65–0.87)	0.89 (0.81–0.94)	0.39 (0.13–0.60)	0.69	42	78
3rd ventricle (ax.)	0.95 (0.91–0.97)	0.84 (0.74–0.91)	0.90 (0.80–0.93)	0.90	47	87
4th ventricle (sag.)	0.96 (0.94–0.98)	0.92 (0.85–0.96)	0.86 (0.76–0.92)	0.91	34	63
Corpus callosum length (sag.)	0.89 (0.78–0.94)	0.96 (0.93–0.98)	0.98 (0.96–0.99)	0.94	23	43
Vermis height (sag.)	0.95 (0.91–0.97)	0.92 (0.87–0.96)	0.89 (0.81–0.93)	0.92	45	83
Cerebellum (cor.)	0.99 (0.99–0.10)	0.99 (0.98–0.10)	0.99 (0.99–0.10)	0.99	50	93
Cerebellum (ax.)	0.99 (0.99–0.10)	0.99 (0.99–0.10)	0.99 (0.99–0.10)	0.99	50	93

n/% = Absolute/relative number of measurable cases. ICC <0.75 = poor, 0.75–0.9 = moderate, >0.9 = excellent reliability.

Table 5. ICCs of gynecologists vs. radiologists

Measurement parameters	Gynecologists ICC (95% CI)	Radiologists ICC (95% CI)	n	%
Lateral ventricle (post. horn, cor.)	0.93 (0.88–0.96)	0.90 (0.84–0.94)	54	100
Lateral ventricle (post. horn, ax.)	0.96 (0.93–0.98)	0.74 (0.59–0.84)	53	98
Lateral ventricle (ant. horn, cor.)	0.85 (0.75–0.91)	0.73 (0.57–0.84)	48	89
Lateral ventricle (ant. horn, ax.)	0.90 (0.83–0.94)	0.66 (0.47–0.79)	48	89
3rd ventricle (cor.)	0.49 (0.23–0.68)	0.84 (0.72–0.91)	42	78
3rd ventricle (ax.)	0.80 (0.67–0.88)	0.92 (0.85–0.95)	47	87
4th ventricle (sag.)	0.62 (0.37–0.78)	0.76 (0.58–0.87)	34	63
Corpus callosum length (sag.)	0.94 (0.89–0.97)	0.51 (0.11–0.77)	23	43
Vermis height (sag.)	0.92 (0.86–0.95)	0.82 (0.69–0.90)	45	83
Cerebellum (cor.)	0.98 (0.97–0.99)	0.98 (0.96–0.99)	50	93
Cerebellum (ax.)	0.99 (0.99–0.10)	0.94 (0.89–0.96)	50	93

n/% = Absolute/relative number of measurable cases. ICC <0.75 = poor, 0.75–0.9 = moderate, >0.9 = excellent reliability.

not confirmed on fetal MRI and where no other abnormal findings were detected, included two marginal ventriculomegalies, a ventricular asymmetry, and a slightly enlarged mega cisterna magna.

Regarding the most common diagnosis hydrocephalus, 23 of 24 positive US findings (96%) were also found on MRI. In 2 cases of negative US findings, a hydrocephalus was diagnosed on fetal MRI (6.3%). Similarly, regarding mild or moderate ventriculomegaly, MRI confirmed all cases of positive US finding, and 6 negative US findings were deemed positive on fetal MRI (13%). The

most common US finding, which was not confirmed on MRI, was mega cisterna magna (n = 8, 67%). Cerebellar hypoplasia was the most common pathology deemed negative on US but positive on MRI (n = 7, 14%; table 7).

Discussion

Fetal MRI has proven to be a valuable addition to US in the detection of fetal brain malformations. While the reliability of US-based fetal brain measurements has been

Table 6. Comparison of good-quality vs. poor-quality images (interobserver reliability)

Measurement parameters	Good quality ICC (95% CI)	Poor quality ICC (95% CI)	n	%
Lateral ventricle (post. horn, cor.)	0.95 (0.92–0.97)	0.87 (0.76–0.94)	54	100
Lateral ventricle (post. horn, ax.)	0.92 (0.87–0.96)	0.83 (0.71–0.92)	52	96
Lateral ventricle (ant. horn, cor.)	0.85 (0.75–0.92)	0.61 (0.39–0.80)	49	91
Lateral ventricle (ant. horn, ax.)	0.82 (0.72–0.90)	0.66 (0.44–0.84)	49	91
3rd ventricle (cor.)	0.48 (0.23–0.67)	0.66 (0.41–0.86)	42	78
3rd ventricle (ax.)	0.88 (0.80–0.94)	0.56 (0.29–0.81)	40	74
4th ventricle (sag.)	0.63 (0.44–0.79)	0.40 (0.07–0.77)	31	57
Corpus callosum length (sag.)	0.65 (0.45–0.82)	0.53 (0.08–0.92)	18	33
Vermis height (sag.)	0.90 (0.83–0.95)	0.81 (0.65–0.92)	45	83
Cerebellum (cor.)	0.99 (0.98–0.99)	0.92 (0.85–0.96)	50	93
Cerebellum (ax.)	0.99 (0.99–0.10)	0.80 (0.65–0.91)	49	91

n/% = Absolute/relative number of measurable cases. ICC <0.75 = poor, 0.75–0.9 = moderate, >0.9 = excellent reliability.

Table 7. Discrepancies between US and fetal MRI in the detection of cerebral pathologies

Cerebral pathologies	n	US+ (absolute %)	MRI+ (relative %)	MRI– (relative %)	US– (absolute %)	MRI– (relative %)	MRI+ (relative %)
Hydrocephalus	24	43	96	4.2	57	94	6.3
Mega cisterna magna	12	21	33	67	79	96	4.6
Aqueductal stenosis	10	18	70	30	82	89	11
Vermis hypoplasia	9	16	89	11	84	98	2.1
ACC	9	16	56	44	84	92	8.5
Ventriculomegaly	7	13	100	0	88	88	12
Cysts	6	11	50	50	89	92	8.0
Cerebellar hypoplasia	4	7.1	50	50	93	87	14
Missing septum pellucidum	4	7.1	75	25	93	90	9.6
Encephalocele	4	7.1	100	0	93	100	0
Holoprosencephaly	4	7.1	75	25	93	100	0
Hemorrhage	3	5.4	100	0	95	96	3.8
Ventricular asymmetry	1	1.8	100	0	98	95	5.5
Schizencephaly	1	1.8	confirmed	0	98	confirmed	0
Missing falx	1	1.8	confirmed	0	98	confirmed	0
Vein of Galen aneurysm	1	1.8	confirmed	0	98	confirmed	0

n = Number of positive US findings; + = pathology detected; – = pathology not detected.

examined extensively [11–14, 20], little is known regarding the inter- and intraobserver reliability of MRI measurements of fetal brain malformations.

According to our data, 3 measurements achieved an ICC for interobserver reliability above 0.9, whereas regarding the intraobserver reliability, 8 measurements achieved an ICC higher than 0.9. The highest reliability was reached in the measurements of the cerebellar diameter, and poor in-

terobserver reliability was observed for the measurements of the diameter of the third ventricle, the fourth ventricle and the corpus callosum. The difficulty of the different observers to obtain reliable corpus callosum measurements was reflected in the amount of nonmeasurable cases. Corpus callosum measurements featured the highest exclusion rate with only 18 measurable cases (33%) for all observers in the interobserver and 23 measurable cases (42%) in the

intraobserver reliability. Similarly, the poor measurability of the third and fourth ventricles was also corroborated by the amount of nonmeasurable cases (77 and 57%). In contrast, the measurements showing excellent reliability, as the diameter of the cerebellum and the posterior horn of the lateral ventricle in the coronal plane, could be performed in 92 and 100% of the cases, respectively.

Parazzini et al. [21] mentioned that the measurements of structures of a smaller size make their evaluation and exact measurements difficult, which results in poor reliability. In our study, in the third ventricle measurements outliers arose from a cavum septum pellucidum cyst in 2 cases, which were mistaken for an enlarged third ventricle by the inexperienced observers.

Not only the exact measurements but also the accuracy of imaging plane setting plays a part in contributing to reasonable measurement reliability. The poor interobserver MRI reliability of the measurements of the corpus callosum and the fourth ventricle may also be caused by the lack of presence of a perfect mid-sagittal brain section due to inadequate plane setting or fetal head rotation. In our study, the genu and the splenium were included in the measurements of the corpus callosum as advised by Garel et al. [10]. In contrast, Parazzini et al. [21] did not take into account the genu and splenium thickness, stating that it is well delineated by the lower hyperintense signal of the third ventricle. Generally, according to Garel [22], it is impossible to precisely evaluate the thickness of the corpus callosum because of the limited spatial resolution on MRI.

Of course, the reliability of measurements can also depend on the skill set of the observers. Fetal brain US is mostly performed by gynecologists and MRI by radiologists; only few perform both modalities. In this study, it was investigated whether obstetricians, experienced in the use of the dynamic imaging method US, would generate similarly reliable fetal brain MRI measurements compared to radiologists, experienced in the analysis of static fetal MR images. Generally, obstetricians are not trained in MRI examination techniques, while radiologists cannot be expected to have all the background knowledge of maternal-fetal medicine [23]. However, it was shown that obstetricians could reach similarly reliable results compared to radiologists when performing standardized fetal brain MRI measurements, which approves the usefulness of the standardized measurements.

Since in other studies comparing fetal brain MRI with US the importance of using high-quality US and MRI was highlighted [8, 24–26], we conducted a subgroup analysis of MR images with differing image quality. Poor image quality MR images were defined in common according to

the above-mentioned criteria. We believe this subgroup analysis to be important as a mere exclusion of images with poor quality would provide results based on artificial selections rather than the reality in clinical practice. According to our results, there was a trend towards more reliable measurements using images of higher quality. The main reason for poor imaging quality appeared to be fetal movement artefacts. Mere technical parameters as field strength did not sufficiently alter the imaging quality to affect the reliability of the measurements.

When one of the observers considered a parameter not to be measurable at all, this linear measurement was excluded for all observers in the reliability analysis. However, the number of excluded cases was taken into account when interpreting the different ICCs.

While results of US and MRI agreed in the primary diagnosis in 40 cases (71%), MRI provided additional information in 16 cases (29%). According to our results, discrepancies between US and MRI findings were often related to structures located in the posterior fossa. Poutamo et al. [27] reported that the anatomy of the posterior fossa remained unclear in 8 of 19 cases on US and in one of 19 cases on MRI. Also Oh et al. [28] highlighted that for these common posterior fossa anomalies understanding of the anatomy is essential to avoid pitfalls and misdiagnoses. According to Paladini et al. [29], fetal MRI is helpful in a limited number of cases with CNS abnormalities and can be of benefit to a lesser extent in posterior fossa and corpus callosum anomalies, confirming our findings.

MRI has been proven as a safe modality for fetal imaging. Nevertheless, the indication of fetal MRI needs to be justified because of the acoustic noise on the fetus, the specific absorption rate, the associated MRI-induced heating of body tissues [30, 31], the limited availability, and the costs of the procedure. It is often challenging for the obstetrician to decide whether to perform further US controls or to request a fetal MRI examination.

A limitation of this study is that the time interval between the US and MRI reached a maximum of 36 days.

Regarding the intraobserver reliability analysis, we chose a minimum time interval of 4 weeks between measurements in order to limit abiding memory. In this time, the examiners did not have any contact with names, data or images of the patients.

Since the aim of this study was to evaluate the reliability of fetal brain MRI measurements, no analysis of the neonatal outcome or fetopathologic examination was made. In order to determine the validity of the findings, a postnatal MRI or a standardized fetopathologic examination would be required.

Conclusion

MRI proved to be a valuable complementary technique to US in the investigation of fetal brain malformations providing additional information in about one third of the selected cases. The reliability of standardized MRI

measurements was reasonable in most linear measurements and especially good in the measurements of the cerebellum and the posterior horn of the lateral ventricle. However, poor reliability was shown in the measurements of the third and fourth ventricles and the length of the corpus callosum.

References

- Myrianthopoulos N: Epidemiology of central nervous system malformations; in Vinken PJ, Bruyn GW (eds): *Handbook of Clinical Neurology*. Amsterdam, Elsevier, 1977, pp 139–171.
- Sepulveda W, Wong AE, Sepulveda F, Martinez-Ten P, Ximenes R: Fetal magnetic resonance imaging and three-dimensional ultrasound in clinical practice: general aspects. *Best Pract Res Clin Obstet Gynaecol* 2012;26:575–591.
- Hibbeln JF, Shors SM, Byrd SE: MRI: is there a role in obstetrics? *Clin Obstet Gynecol* 2012;55:352–366.
- Prayer D: *Fetal MRI*. Berlin, Springer, 2011, pp 2–19.
- Coakley FV, Glenn OA, Qayyum A, Barkovich AJ, Goldstein R, Filly RA: Fetal MRI: a developing technique for the developing patient. *AJR Am J Roentgenol* 2004;182:243–252.
- Girard NJ, Chaumoitte K: The brain in the belly: what and how of fetal neuroimaging? *J Magn Reson Imaging* 2012;36:788–804.
- Griffiths PD, Porteous M, Mason G, Russell S, Morris J, Fanou EM, Reeves MJ: The use of in utero MRI to supplement ultrasound in the foetus at high risk of developmental brain or spine abnormality. *Br J Radiol* 2012;85:e1038–e1045.
- Blondiaux E, Garel C: Fetal cerebral imaging – ultrasound vs. MRI: an update. *Acta Radiol* 2013;54:1046–1054.
- Malamateniou C, Malik SJ, Counsell SJ, Allsop JM, McGuinness AK, Hayat T, Broadhouse K, Nunes RG, Ederies AM, Hajnal JV, Rutherford MA: Motion-compensation techniques in neonatal and fetal MR imaging. *AJNR Am J Neuroradiol* 2013;34:1124–1136.
- Garel C, Delezoide A, Guibaud L, Sebag G: *MRI of the fetal brain: normal development and cerebral pathologies*. Berlin, Springer, 2004, pp 5–130.
- Perni SC, Chervenak FA, Kalish RB, Magherini-Rothe S, Predanic M, Streltsoff J, Skupski DW: Intraobserver and interobserver reproducibility of fetal biometry. *Ultrasound Obstet Gynecol* 2004;24:654–658.
- Salman MS, Mousa HA, Twining P, Jones NW, James D, Momtaz M, Aboulghar M, El-Sheikhah A, Bugg G: Would gestational age and presence of brain anomalies affect interobserver reliability of fetal head biometry? Using off-line analysis of 3-D dataset. *Ultrasound Med Biol* 2012;38:69–74.
- Lima JC, Miyague AH, Filho FM, Nastri CO, Martins WP: Biometry and fetal weight estimation by two-dimensional and three-dimensional ultrasonography: an intraobserver and interobserver reliability and agreement study. *Ultrasound Obstet Gynecol* 2012;40:186–193.
- Yang F, Leung KY, Lee YP, Chan HY, Tang MH: Fetal biometry by an inexperienced operator using two- and three-dimensional ultrasound. *Ultrasound Obstet Gynecol* 2010;35:566–571.
- Pilu G, Falco P, Gabrielli S, Perolo A, Sandri F, Bovicelli L: The clinical significance of fetal isolated cerebral borderline ventriculomegaly: report of 31 cases and review of the literature. *Ultrasound Obstet Gynecol* 1999;14:320–326.
- Gaglioti P, Danelon D, Bontempo S, Mombro M, Cardaropoli S, Todros T: Fetal cerebral ventriculomegaly: outcome in 176 cases. *Ultrasound Obstet Gynecol* 2005;25:372–377.
- Levine D, Barnes PD, Madsen JR, Abbott J, Wong GP, Hulka C, Mehta T, Li W, Edelman RR: Fetal CNS anomalies revealed on ultrafast MR imaging. *AJR Am J Roentgenol* 1999;172:813–818.
- Paladini D, Pastore G, Cavallaro A, Massaro M, Nappi C: Agenesis of the fetal corpus callosum: sonographic signs change with advancing gestational age. *Ultrasound Obstet Gynecol* 2013;42:687–690.
- Portney LG, Watkins MP: *Foundations of Clinical Research: Applications to Practice*. Norwalk, Appleton and Lange, 1993.
- Sarris I, Ioannou C, Chamberlain P, Ohuma E, Roseman F, Hoch L, Altman DG, Papa-georgiou AT: Intra- and interobserver variability in fetal ultrasound measurements. *Ultrasound Obstet Gynecol* 2012;39:266–273.
- Parazzini C, Righini A, Rustico M, Consonni D, Triulzi F: Prenatal magnetic resonance imaging: brain normal linear biometric values below 24 gestational weeks. *Neuroradiology* 2008;50:877–883.
- Garel C: Fetal cerebral biometry: normal parenchymal findings and ventricular size. *Eur Radiol* 2005;15:809–813.
- Cannie M, Jani J, Dymarkowski S, Deprest J: Fetal magnetic resonance imaging: luxury or necessity? *Ultrasound Obstet Gynecol* 2006;27:471–476.
- Pugash D, Brugger PC, Bettelheim D, Prayer D: Prenatal ultrasound and fetal MRI: the comparative value of each modality in prenatal diagnosis. *Eur J Radiol* 2008;68:214–226.
- Perrone A, Savelli S, Maggi C, Di Pietro L, Di Maurizio M, Tesei J, Ballesio L, De Felice C, Giancotti A, Di Iorio R, Manganaro L: Magnetic resonance imaging versus ultrasonography in fetal pathology. *Radiol Med* 2008;113:225–241.
- Levine D, Feldman HA, Tannus JF, Estroff JA, Magnino M, Robson CD, Poussaint TY, Barnewolt CE, Mehta TS, Robertson RL: Frequency and cause of disagreements in diagnoses for fetuses referred for ventriculomegaly. *Radiology* 2008;247:516–527.
- Poutamo J, Vanninen R, Partanen K, Ryyanen, Kirkinen P: Magnetic resonance imaging supplements ultrasonographic imaging of the posterior fossa, pharynx and neck in malformed fetuses. *Ultrasound Obstet Gynecol* 1999;13:327–334.
- Oh KY, Rassner UA, Frias AE Jr, Kennedy AM: The fetal posterior fossa: clinical correlation of findings on prenatal ultrasound and fetal magnetic resonance imaging. *Ultrasound Q* 2007;23:203–210.
- Paladini D, Quarantelli M, Sglavo G, Pastore G, Cavallaro A, D'Armiento, Salvatore M, Nappi C: The role of MRI in the clinical management of fetuses with central nervous system abnormalities in a tertiary referral center. *Ultrasound Obstet Gynecol* 2013, Epub ahead of print.
- Hand JW, Li Y, Thomas EL, Rutherford MA, Hajnal JV: Prediction of specific absorption rate in mother and fetus associated with MRI examinations during pregnancy. *Magn Reson Med* 2006;55:883–893.
- Pediaditis M, Leitgeb N, Cech R: RF-EMF exposure of fetus and mother during magnetic resonance imaging. *Phys Med Biol* 2008;53:7187–7195.