Reliability of MR Imaging–Based Posterior Fossa and Brain Stem Measurements in Open Spinal Dysraphism in the Era of Fetal Surgery

M. Aertsen, J. Verduyckt, F. De Keyzer, T. Vercauteren, F. Van Calenbergh, L. De Catte, S. Dymarkowski, P. Demaerel and J. Deprest

*AJNR Am J Neuroradiol* 2019, 40 (1) 191-198
doi: https://doi.org/10.3174/ajnr.A5930
http://www.ajnr.org/content/40/1/191
Reliability of MR Imaging–Based Posterior Fossa and Brain Stem Measurements in Open Spinal Dysraphism in the Era of Fetal Surgery

M. Aertsen, J. Verduyckt, F. De Keyzer, T. Vercauteren, F. Van Calenbergh, L. De Catte, S. Dymarkowski, P. Demaerel, and J. Deprest

ABSTRACT

BACKGROUND AND PURPOSE: Fetal MR imaging is part of the comprehensive prenatal assessment of fetuses with open spinal dysraphism. We aimed to assess the reliability of brain stem and posterior fossa measurements; use the reliable measurements to characterize fetuses with open spinal dysraphism versus what can be observed in healthy age-matched controls; and document changes in those within 1 week after prenatal repair.

MATERIALS AND METHODS: Retrospective evaluation of 349 MR imaging examinations took place, including 274 in controls and 52 in fetuses with open spinal dysraphism, of whom 23 underwent prenatal repair and had additional early postoperative MR images. We evaluated measurements of the brain stem and the posterior fossa and the ventricular width in all populations for their reliability and differences between the groups.

RESULTS: The transverse cerebellar diameter, cerebellar herniation level, clivus-supraocciput angle, transverse diameter of the posterior fossa, posterior fossa area, and ventricular width showed an acceptable intra- and interobserver reliability (intraclass correlation coefficient > 0.5). In fetuses with open spinal dysraphism, these measurements were significantly different from those of healthy fetuses (all with P < .0001). Furthermore, they also changed significantly (P value range = .01 to < .0001) within 1 week after the fetal operation with an evolution toward normal, most evident for the clivus-supraocciput angle (65.9 ± 12.5°; 76.6 ± 10.9°; P < .0001) and cerebellar herniation level (−9.9 ± 4.2 mm; −0.7 ± 5.2; P < .0001).

CONCLUSIONS: In fetuses with open spinal dysraphism, brain stem measurements varied substantially between observers. However, measurements characterizing the posterior fossa could be reliably assessed and were significantly different from normal. Following a fetal operation, these deviations from normal values changed significantly within 1 week.

ABBREVIATIONS: ACI = atiocerebral index; CHL = cerebellar herniation level; CSA = clivus-supraocciput angle; GA = gestational age; ICC = intraclass correlation coefficient; OSD = open spinal dysraphism; PF = posterior fossa; TCD = transverse cerebellar diameter; TDPF = transverse diameter of the PF; VW = ventricular width

Open spinal dysraphism (OSD), subdivided into myelomeningocele and myeloschisis, is a nonlethal congenital malformation with complex physical and neurodevelopmental sequelae. Its prevalence is approximately 4.9 per 10,000 live births in Europe and 3.17 in the United States.1,3 OSD results in motor and sensory deficits, their extension being defined by the upper level of the anatomic defect. These range, as the level increases, from bladder, bowel, and sexual dysfunction to involvement of the lower and even upper extremities and secondary orthopedic disabilities.4,5 Children with OSD almost invariably have an associated Chiari II hindbrain malformation and ventriculomegaly.6 The Chiari II malformation is characterized by posterior fossa (PF) and brain stem abnormalities with downward displacement and compression of the cerebellum and brain stem.7

Geerdink et al8 demonstrated that morphometric measures reliably quantify the morphologic distortions of Chiari II malformation on postnatal MR images. The mamillopontine distance and the cerebellar width were the most sensitive and specific
de
terminants of Chiari II. Some fetuses with OSD have ventriculomegaly, and its degree is believed to be predictive of the need for postnatal shunting. In 2011, the Management of Myelomeningocele Study (MOMS) demonstrated the benefit of in utero repair of myelomeningocele because the need for ventricular shunting at 12 months was reduced and motor outcome at 30 months improved. Fetuses with the suspicion of OSD should be assessed comprehensively to counsel parents about the expected outcome and possibility of fetal surgery. In this assessment, fetal MR imaging has a crucial role to characterize the brain and spinal abnormalities and rule out additional anomalies in fetuses with OSD. For fetal surgery eligibility, the presence of Chiari II hindbrain malformation on MR imaging is a necessary finding. Many measurements have been proposed to describe the typical PF changes in fetuses with OSD, yet the reproducibility of these has rarely been studied. These parameters have also been shown to change after in utero repair of OSD in small series and at different time points after fetal surgery, yet no study has consistently reported early postoperative assessment in utero.

The aims of this study were 3-fold: 1) to assess the reproducibility of measurements of the brain stem and PF that have been suggested to be representative on postnatal and prenatal MR imaging, 2) to apply those parameters that were shown to be reproducible, to discriminate fetuses with OSD from gestational age–matched fetuses with a normal PF; and 3) to document early changes in these measurements 1 week after a fetal operation.

MATERIALS AND METHODS
This was a single-center retrospective study at University Hospitals Leuven that was approved by its ethics committee (S60814). Patients eligible for the OSD group were those having fetal MR imaging examinations for additional assessment because of a prenatal diagnosis of OSD on ultrasound. Before MR imaging, patients had an ultrasound assessment, in which the lesion, secondary changes, and, when applicable, associated anomalies were characterized. Before the MR imaging, the radiologist was informed of the ultrasound findings. For the gestational age–matched controls, we included fetuses assessed for other congenital anomalies that do not affect the central nervous system or who were scanned for suspected CNS abnormalities with normal findings on prenatal ultrasound, fetal MR imaging, and postnatal evaluation (On-line Table 1).

Fetal Imaging and Quality Criteria
The routine protocol for this condition includes acquisition on a 1.5T MR imaging system (Aera; Siemens, Erlangen, Germany) with 2 small body coils placed adjacent to each other over the maternal abdomen. The mother was positioned in the supine or left lateral decubitus position. The images used were T2-weighted HASTE or balanced steady-state gradient-echo sequences in the sagittal, axial, and coronal planes relative to the fetal head. Before September 2015, maternal sedation (flunitrazepam, 0.5 mg orally 20–30 minutes before the examination) was used when the gestational age (GA) was <30 weeks. Later, this was abandoned because we, like others, thought that this induced maternal adverse effects while not clinically required. For this study, we searched our data base for all examinations performed in the setting of spinal dysraphism assessment, as well as for appropriate gestational age–matched controls. The image quality had to be good, consisting of at least 3 orthogonal T2-weighted HASTE series of the fetal brain with limited fetal motion, allowing adequate performance of the outcome measurements. The primary selection and review of images was performed by a single pediatric radiologist (M.A.) with >3 years of experience in fetal MR imaging. The main exclusion criteria were twin pregnancy, syndromal pathology, fetal hydrops, or anhydramnios. The number of patients and individual reasons for exclusion are shown in On-line Table 2. This exclusion left data from 349 MR imaging examinations of a total of 1006, including 274 examinations in 246 control fetuses. These data illustrate that some fetuses were scanned more than once. Additionally, we included 52 MR imaging examinations in fetuses with OSD, of whom 23 had a repeat MR imaging examination after the operation. The eligibility criteria for fetal surgery were those used in the MOMS trial.

Outcome Measurements
Biometric variables included the transverse cerebellar diameter (TCD), pontine thickness, and pontine height, measured according to the standards defined by Garel and Tilea et al. The transverse diameter of the PF (TDPF) was measured according to Woitek et al, who suggested that this would be a proxy for the TCD. The midsagittal PF area was measured according to Tsai et al. The ventricular width (VW) was measured in the coronal plane according to Garel and in case of asymmetry, the largest value was taken into account. Mamillopontine distance, the level of kinking of the brain stem, medullary length, tentorial length, and width of the cisterna magna were measured as described by Geerdink et al. The width of the foramen magnum was defined as the distance between the opisthion and the basion. The cerebellar herniation level (CHL) was measured by drawing a perpendicular line from the foramen magnum to the lowest cerebellar portion. In the presence of cerebellar herniation, the deepest portion was measured. The clivus-supraoccipital angle (CSA) was measured according to D’Addario et al. The TCD, TDPF, mamillopontine distance, TL, PF area, and CSA are demonstrated in Fig 1.

Reproducibility Study
The reproducibility of measurements was determined on a randomly chosen subgroup of spinal dysraphism cases (n = 15/52; referred to as the pilot group). Images were anonymized and uploaded to a research server for assessment by M.A. and J.V., a radiologist with 1-year specific training for fetal MR imaging. This radiologist was first trained with a training dataset from 5 other fetuses with spinal dysraphism, with the help of a purposely designed training document. For intraobserver assessment, M.A. read the images twice in a random order with a 2-week interval. J.V. measured parameters with at least moderate (intraclass correlation coefficient [ICC] > 0.5) intraobserver reliability once to obtain interobserver reliability.
Posterior Fossa Characteristics
The PF characteristics were determined on the presumed healthy population to obtain normative values. To compare cases of fetal myelomeningocele with the healthy population, we used these normative curves to calculate expected values for the given gestational ages, and the fetal myelomeningocele values were then expressed as observed over the expected ratio.

Short-Term (<7 Days) Postoperative Changes
In this part of the study, we looked at the difference between PF measurements shown to be reproducible in the above part of the study in 23 fetuses who had a fetal operation at our center. These fetuses had preoperative MR imaging and were imaged again within 1 week after the operation. All measurements were performed on T2-weighted images in the coronal or sagittal plane of the fetal head. Again, the values were expressed as observed over the expected ratio to determine changes after prenatal treatment that were not attributable to normal growth.

In addition to the posterior fossa, we also evaluated the ventricular width in fetuses with OSD and the difference from the control population. To describe the differential effect of a fetal operation on the parenchyma and ventricles, one can measure changes in the so-called atriocerebral index (AGI). The AGI is the ratio of the atrial diameter and the cerebral (parenchymal) biparietal diameter. Others called this index the ventricular width index as used in the postnatal literature.

Statistics
Intraobserver and interobserver variability were analyzed with a 2-way random ICC using SPSS for Windows, Version 22.0 (Released 2004; IBM, Armonk, New York). An ICC cutoff value of 0.5 as the lowest acceptable was chosen, taking into account the guidelines for interpretation by Cicchetti and allowing some variation in view of the limited spatial resolution for assessing such small structures so that borderline parameters can be fully investigated. For the interpretation of ICC values, we followed the guidelines of Cicchetti and others called this index the ventricular width index as used in the postnatal literature.

RESULTS

Demographics
MR imaging examinations in controls were performed at a mean GA of 27.9 ± 5.3 weeks (range, 18.6–38.3 weeks). In fetuses with OSD, the mean GA at MR imaging was 23.6 ± 0.3 weeks (range, 19.3–27.3 weeks). Descriptive statistics for the study parameters in controls and cases with OSD are shown in Table 1.

Reproducibility Study
The intraobserver ICCs for the PF area, VW, TCD, and CHL were excellent. Conversely, TDPF (0.729), the pontine thickness (0.59),
Table 1: Mean and SDs of different parameters for the controls and fetuses with nonoperated and operated spinal dysraphism

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Controls (n = 274)</th>
<th>Nonoperated OSD (n = 52)</th>
<th>Postoperative OSD (n = 23)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA (wk)</td>
<td>27.9 (5.3)</td>
<td>23.6 (0.3)</td>
<td>26.0 (1.1)</td>
</tr>
<tr>
<td>VW (mm)</td>
<td>6.16</td>
<td>11.9 (4.6)</td>
<td>12.3 (3.6)</td>
</tr>
<tr>
<td>PF area (cm²)</td>
<td>7.4 (2.2)</td>
<td>3.1 (0.9)</td>
<td>4.44 (0.83)</td>
</tr>
<tr>
<td>TDPF (mm)</td>
<td>39.8 (7.1)</td>
<td>22.3 (3.4)</td>
<td>27.2 (4.3)</td>
</tr>
<tr>
<td>TCD (mm)</td>
<td>31.6 (6.4)</td>
<td>21.8 (3)</td>
<td>24.4 (3.3)</td>
</tr>
<tr>
<td>Ratio TCD/TDPF</td>
<td>0.79 (0.05)</td>
<td>1 (0.05)</td>
<td>0.9 (0.07)</td>
</tr>
<tr>
<td>CHL (mm)</td>
<td>8.8 (1.9)</td>
<td>-9.3 (5.4)</td>
<td>-0.7 (5.2)</td>
</tr>
<tr>
<td>CSA</td>
<td>87.1° (8.3°)</td>
<td>62.5° (9.7°)</td>
<td>76.6° (10.9°)</td>
</tr>
</tbody>
</table>

Note: *x indicates square root (gestational age in weeks); R², coefficient of determination.

Dissection (P = .0127), PF area (P = .0003), TDPF (P = .0127), CSA (P < .0001), and VW (P = .0002). Figure 3 shows the individual observations and boxplots in patients with OSD for all tested parameters, both pre- and postoperatively. In 18/23 (78%) fetuses, the postoperative observed over expected ratios of the PF area were improving toward normal compared with the preoperative measurement. The same was true for the observed over expected ratios of the TDPF in 15/23 (65%), for the observed over expected ratios of the TCD in 7/23 (30%), and for the observed over expected TCD/ TDPF ratios in 4/23 (17%) postoperative fetuses. The observed over expected ratio level of cerebellar herniation increased, meaning that the Chiari II–associated changes were reduced in 19/23 fetuses at 1 week after the operation. The observed over expected ratios of the CSA increased toward, the normal range in 16/23 (70%) postoperative fetuses. The VW increased in 17/23 (74%) fetuses, but there was no difference in the ACI between the pre- and postoperative examinations (P = .46), with a mean of the preoperative measurements of 0.22 and, postoperatively, of 0.23.

DISCUSSION

A large number of structural measurements on MR images of the PF have been suggested to characterize changes attributed to OSD. Some are believed to be clinically relevant due to their relation to the symptoms associated with a small PF, Chiari II malformation, and brain stem compression. These changes are represented by the degree of cerebellar herniation, mamillopontine distance, brain stem kinking, tentorial hypoplasia, and the configuration of the fourth ventricle. These features can be measured reproducibly in the postnatal period. In the era of prenatal diagnosis and fetal surgery, logically, the same measurements are also used. Yet, in case of prenatal surgery, these are measured on midgestational prenatal images at a time when the condition is still progressive and image quality is different. We therefore investigated whether those measurements were reproducible in the window of interest. For instance, we did not evaluate the fourth ventricle because it is hardly visible around 20–24 weeks in fetuses with OSD. Conversely, we looked at parameters characterizing the brain stem (mamillopontine distance, pontine thickness, pontine length, foramen magnum diameter, level of brain stem kinking, medulla length, tentorial length, and cisterna magna width) and the PF as a whole (TCD, TDPF, PF area, CHL, TCD/TDPF, and CSA). Due to its clinical relevance in fetuses with OSD and the impact on outcome after fetal surgery, we also included the VW.

Herein, we conclude that brain stem parameters in fetuses with OSD cannot be measured reliably at midgestation; thus, we were not able to objectively measure brain stem elongation and displacement in utero. For fetuses with OSD, we believe this issue is due to the limited spatial resolution of fetal MR imaging at that point in gestation when the structures involved are in the millimetric range so that the slightest measurement error has a tremendous impact on statistics. Another reason is that in OSD, there is a decrease or even absence of extra-axial CSF, further limiting the contrast resolution of MR imaging. Contrast resolution is essential for accurate evaluation of small PF structures as well as some additional cerebral lesions due to the presence of different structures (medulla oblongata, pons, vermis, cerebellum) in a small area. There is no difference between balanced steady-state free-precession sequences and half-Fourier rapid relaxation with relaxation enhancement sequences for measuring the foramen magnum. Conversely, when we looked at larger structures (PF area, TCD, and TPFD) and/or with more abundant contrast between fluid and soft tissue (CHL and VW), the ICC values were much better. We confirmed this finding for those parameters typical for the smaller posterior fossa and cerebellar descent in the pathologic subgroup, as others did. Moreover, the difference with our gestational age–matched healthy patients was highly significant, again as previously described.

Fetal surgery has been shown to reverse those posterior fossa...
Other investigators used the above parameters to measure those typically ≥4 weeks after the operation. In the present study, we acquired images within 2 weeks. In line with ultrasound observations, we quantified significant changes during that short observation period. Within 1 week, 26% of operated fetuses had a PF area within the normal range, and in 52%, the TCD was normal. Furthermore, the cerebellar herniation level was at or above the foramen magnum in 52% of fetuses, and 70% had a normal CSA. These acute changes in the PF following closure of the defect are in line with the theory of McLone and Knepper. In other words, it seems that the effects of fetal surgery on the PF are already evident and can be quantified very early postoperatively. They are very likely to persist because others observed the same effects later on and even confirmed them after birth. This outcome might be an interesting proxy for measurement of the efficacy of fetal surgery in clinical studies.

In this short-term follow-up study, we observed a postoperative increase in ventricular width within 1 week in most patients. Such increase is in line with observations made by others, though several weeks after fetal surgery. They suggest that there is still a certain degree of obstructive ventricular widening. The dynamics of CSF fluid production and resorption in OSD are still poorly understood. It may take some time after fetal surgery for CSF fluid circulation to normalize after stopping its egress.

In healthy fetuses, the ACi drops dramatically between 24 and 27 weeks, which means that there is, during that time period, a proportional increase in the parenchymal brain component. In our patients undergoing fetal surgery, the ACi remained stable.
This might be counterintuitive and contradicts the findings of Rethmann et al.\textsuperscript{18} They measured the ACi and observed a drop in the ACi; yet, that was 4 weeks after the operation and continued after birth. These contrasting findings can be explained in different ways. A drop in ACi would suggest a proportional increase in biparietal cerebral diameter, hence a larger parenchymal component. Conversely, the increase in ACi for a comparable VW in our cohort would suggest that the parenchymal component decreases. Although tempting, both groups cannot be compared because measurements were performed at different gestational ages. In healthy fetuses, the ACi spontaneously declines between 24 and 27 weeks. If fetuses with OSD follow this normal evolution, our findings may eventually align with those of Rethmann et al and can be explained by spontaneous and normal evolution. Unfortunately, we have no longitudinal follow-up MR images to further study these observations.

The brain stem, which, to our knowledge, was not performed in detail before. We based our evaluation on our own normative values. Furthermore, we documented all parameters in a relatively large pathologic population in the narrow gestational age range that is relevant to prenatal spina bifida repair. Although the number of operated fetuses was not very large, we were able to describe early postoperative changes therein. There are, however, some shortcomings. First, our control fetuses definitely had normal CNS findings on both MR imaging and ultrasound but were not truly fully healthy fetuses. Controls underwent MR imaging because of other congenital abnormalities, presumed not to be associated with CNS abnormalities.

Second, we did not report on advanced MR images, such as DWI and DTI, which may also provide relevant information. We definitely acknowledge the potential of DWI and DTI because they may detect more subtle abnormalities below the anatomic level. Woitek et al\textsuperscript{38} already showed that fetuses with spina bifida have increased fractional anisotropy compared with normally developing fetuses, however without reporting the functional im-

FIG 3. Boxplot demonstrating the minimum, first quartile, median, third quartile, and maximum of the observed over expected ratio in preoperative fetuses with open spinal dysraphism and postoperative fetuses with OSD at 1 week for the posterior fossa area, ventricular width, transverse diameter of the posterior fossa, transverse cerebellar diameter, cerebellar herniation level, and clivus-supraocciput angle.
pact. Although we acquired such sequences in fetuses with spina bifida, we were lacking those in healthy fetuses or the controls used in this study; hence, we could not interpret the findings. Third, we describe in utero findings without correlation to postnatal short- or long-term follow-up or early postnatal MR imaging confirmation. Although relevant, such a follow-up was beyond the scope of this study as was a comparison of these posterior fossa measurements with those in a cohort that underwent postnatal repair. Postnatal evaluation would most likely have identified additional findings, such as subependymal heterotopias. These are often missed in utero before as well as after fetal surgery.

CONCLUSIONS

This study showed that the brain stem cannot be reliably characterized using the current panel of measurements in fetuses with OSD. Conversely, posterior fossa measurements are demonstrated to be reliable in the evaluation of fetuses with OSD. In addition, these were significantly different from those in the healthy population and changed within 7 days after prenatal surgery. This finding advocates for their use in the evaluation of fetuses with OSD on fetal MR imaging on a routine basis before and shortly after a prenatal operation.

Disclosures: Tom Vercauteren—UNRELATED: Employment: University College London.
Comments: main employer for the duration of this work; Grant: Wellcome Trust [WT101957]/Engineering and Physical Sciences Research Council (NS/A000027/1*). Comments: Support for Travel to Meetings for the Study or Other Purposes: Wellcome Trust/Engineering and Physical Sciences Research Council, Comments: travel support from the listed grant funding; Luc De Catte—UNRELATED: Employment: University College London.
Comments: paid travel and accommodations to the Sharing Progress in Neonatology 2018 meeting.

REFERENCES

29. Doel T, Shah D, Pratt R, et al. GIFT-Cloud: A data sharing and
collaboration platform for medical imaging research. *Comput Methods Programs Biomed* 2017;139:181–90 CrossRef Medline


