Thieme

MR Pelvimetry for Breech Presentation at Term– Interobserver Reliability, Incidental Findings and Reference Values

MR-Pelvimetrie bei persistierender Beckenendlage – Interobserver-Reliabilität, Nebenbefunde und Referenzwerte

Authors

Amanda von Bismarck[1], Birgit Ertl-Wagner[2], Sophia Stöcklein[2], Franziska Schöppe[2], Christoph Hübener[1], Linda Hertlein[1], Dominique Baron-Tomlinson[1], Sven Mahner[1], Maria Delius[1], Uwe Hasbargen[1], Marie Franz[1]

Affiliations

1 Gynecology and Obstetrics, University Hospital, LMU Munich, Germany
2 Institute for Clinical Radiology, University Hospital, LMU Munich, Germany

Key words

MR pelvimetry, breech, obstetric, interobserver, reference value

received 17.01.2018
accepted 13.08.2018

Bibliography

DOI https://doi.org/10.1055/a-0715-2122
Published online: 19.9.2018
Fortschr Röntgenstr 2019; 191: 424–432

ISSN 1438-9029

Correspondence

Dr. Marie Franz
Gynecology and Obstetrics, University Hospital, LMU Munich, Marchioninistrasse 15, 81377 Munich, Germany
Tel.: ++ 49/89/4 40 07 38 00
marie.franz@med.uni-muenchen.de

ZUSAMMENFASSUNG

Hintergrund Antepartale MR-Pelvimetrie wird genutzt, um die Durchführbarkeit von vaginalen Beckenendlagengeburten einzuschätzen. Evaluiert wurden die Reliabilität von MR-pelvimetrischen Messungen sowie Nebenbefunde durch verschiedene Kliniker; zudem wurden potenzielle Referenzwerte erwogen.

Methode In dieser monozentrischen Studie wurde die radiologische Datenbank nach MR-Pelvimetrien von Schwangeren mit Fetus in Beckenendlage zwischen 1999 und 2016 durchsucht. 99 konsekutive Pelvimetrien wurden inkludiert. Eine strukturierte, unabhängige Auswertung erfolgte durch sechs Untersucher mit drei klinischen Erfahrungsstufen (Oberarzt, Facharzt, Assistenzarzt aus je der radiologischen sowie geburtshilflichen Abteilung). Die Sequenzen wurden bezüglich der quantitativen Erfassung von Conjugata vera (CV) und Diameter transversalis (DT), Bildqualität und Nebenbefunde analysiert. Für die Untersuchung der Referenzwerte wurden geburtshilfliche Daten aus der passenden Datenbank erhoben.

Ergebnisse Die Interobserver-Übereinstimmung war durchgehend stark (Reichweite der durchschnittlichen Intraklassen-Korrelationskoeffizienten: 0,889 – 0,968). Individuelle Messfehler rangierten zwischen 0 und 2 mm, durchschnittliche Übereinstimmungsgrenzen waren ± 3 mm. In Bezug auf den Geburtsmodus wies die Gruppe „Kaiserschnitt empfohlen" (rCS) signifikant kleinere CV-Messungen auf (CV: 11,37 ± 0,73, p-Wert < 0,0001) als alle anderen Gruppen. Es wurde kein statistischer Unterschied in CV zwischen den Gruppen „vaginale Geburt" und „ungeplanter Kaiserschnitt" gefunden (p-Wert 0,902). Die DT-Messungen zeigten nur einen signifikanten Unterschied zwischen rCS und „elektiver Kaiserschnitt" (p-Wert 0,039). 134 Nebenbefunde wurden erfasst.

Schlussfolgerung Eine starke Interobserver-Übereinstimmung unabhängig von der Erfahrungsstufe des Auswerters bestätigt die MR-Pelvimetrie als eine verlässliche Methode, um ein fetopelvines Missverhältnis bei Schwangeren mit Fetus in Beckenendlage zu identifizieren. Für eine vollständige Erfassung der Nebenbefunde ist eine radiologische Expertise unabdingbar.

Kernaussagen:

- MR-Pelvimetrie ist eine verlässliche Methode unabhängig von der Erfahrungsstufe des Auswerters.
- Conjugata-vera-Messungen helfen in präpartaler Einschätzung zur Durchführbarkeit einer vaginalen Beckenendlagengeburt.
- Zufallsbefunde kommen relativ häufig vor.

ABSTRACT

Purpose Antepartal MR pelvimetry is used to assess the viability of vaginal breech delivery. We evaluated the reliability of MR pelvimetric measurements as well as incidental findings noted by different clinicians and assessed potential reference values.

Methods In this monocentric study, the radiologic database was searched for obstetric MR pelvimetries with singleton

breech pregnancies between 1999 and 2016. 99 consecutive MR pelvimetries were included. A structured, independent review was performed by six observers with three clinical experience levels (attending, fellow, junior resident from the departments of radiology and obstetrics). Image analysis entailed the quantitative assessment of conjugata vera (CV) and diameter transversalis (DT), image quality and incidental findings. Obstetric data was retrieved from the obstetric database for reference value assessment.

Results Interobserver agreement was strong throughout (mean intraclass correlation coefficient range: 0.889 – 0.968). The individual measuring biases ranged between 0 – 2 mm, and the average limits of agreement were ± 3 mm. Regarding the mode of delivery, the recommended cesarean section (rCS) group showed significantly smaller CV measurements (CV: 11.37 ± 0.73, p-value < 0.0001) than any other delivery group. No statistical difference in CV between the vaginal delivery and unplanned cesarean section groups was found (p-value 0.902). DT measurements only showed a significant difference between rCS and elective cesarean section (p-value 0.039). 134 incidental findings were noted.

Conclusion Strong interobserver agreement irrespective of the reader's experience level supports MR pelvimetry as a reliable method for identifying women with fetopelvic disproportion in breech presentation. For a comprehensive appraisal of incidental findings, radiologic expertise is vital.

Key points:

- MR pelvimetry is a reliable method irrespective of the reader's experience level.
- Conjugata vera measurements aid in the prepartal viability assessment of vaginal breech delivery.
- Incidental findings are relatively common.

Citation Format

- von Bismarck A, Ertl-Wagner B, Stöcklein S et al. MR Pelvimetry for Breech Presentation at Term– Interobserver Reliability, Incidental Findings and Reference Values. Fortschr Röntgenstr 2019; 191: 424–432

Introduction

Cesarean section (CS) rates in general and in cases with fetus in breech presentation in particular have risen steadily in the last decades [1]. Considering maternal and fetal morbidity in CS [2 – 4], the viability of vaginal breech delivery (VD) should be assessed. Seen as safe for mother and fetus [5], MR pelvimetry provides objective information about maternal pelvic capacity that cannot be as accurately obtained by manual examination [6] or ultrasound [7, 8]. Persistent breech presentation and suspected cephalopelvic disproportion (CPD) in vertex presentation constitute the primary reasons for prepartal MR pelvimetry referral. While skepticism regarding the usefulness of prepartal pelvimetry has been expressed [9], several studies have reported promising results. Berger et al. [10] calculated proportions between fetal breech and maternal pelvic diameters, which were significantly less favorable in women with failure to progress in labor than in women who delivered vaginally. A randomized controlled study [11] found a significantly higher VD rate in women with previous MR pelvimetry compared to women with only manual pelvic assessment (76 % and 59 %, respectively). The Term Breech Trial [12] acknowledged that only 9.8 % of cases with a trial of labor (TOL) had previously been evaluated by prepartal radiologic pelvimetry; VD was successful in 591 (61 %) out of 967 TOL candidates. Instead, the PREMODA study found that 1794 (71 %) out of 2526 TOL cases delivered vaginally; the authors partly ascribed their favorable VD rate to the frequent use of prepartal pelvimetry, namely in 2064 TOL cases (82.5 %) [13]. Also, several studies have described differences in mean pelvimetric measurements between groups of successful and failed TOL [6, 14, 15]. Yet, no unanimously accepted reference values for clinical application have been established.

MR pelvimetry and its interpretation represent a junction of radiologic and obstetric expertise. Korhonen et al. [16] retrospectively assessed MR pelvimetric measurements between radiologic and obstetric reports and found strong interobserver reliability, investigating threshold values in a subsequent study [17]. Keller et al. [15] prospectively analyzed inter- and intraobserver reliability among four radiologists and one obstetrician. However, both authors concentrated on CPD in vertex presentation and neglected experience levels. In our study, we aimed to investigate interobserver reliability and reference values as well as incidental findings of MR pelvimetry in breech presentation at term.

Materials and Methods

This study was approved by the institutional ethics committee (#44 – 16, 12.02.2016, amended 27.02.2017) and performed according to the standards of the Helsinki Declaration of 1964 (revised 2013). For this data analysis, informed consent was waived. Our institution is a tertiary center with 1600 deliveries per year on average, including 140 breech deliveries.

Data Analysis

At our hospital, it is the clinical standard to offer external cephalic version and, if unsuccessful, MR pelvimetry to patients with a fetus in persistent breech presentation and the wish for VD. The radiology information database (Syngo Imaging, Siemens, Germany) was retrospectively searched for MR pelvimetric referrals due to breech presentation between August 1999 and May 2016. All patients had been referred after persistent breech presentation was diagnosed by routine ultrasound examination. For the interobserver study, the inclusion criteria were uncomplicated singleton breech pregnancies and the retrievability of original MR sequences.

Subsequently, patient names were matched with the obstetric database (PIA, GE-Viewpoint, Germany) to identify all women who

had delivered in-house. For the assessment of reference values, the inclusion criteria were documented breech presentation at birth and in-house delivery. Data were collected in a commercially available worksheet (Excel, Microsoft, USA) and analyzed in an anonymized fashion.

MR pelvimetry

MR pelvimetry was performed as previously described [18]. The woman was placed in a supine position in a 1.5-Tesla MR system with a body coil. In total, six different Siemens Magnetom systems were used in the 16-year period (between 2001 – 04: Vision, Harmony, Symphony; between 2005 – 16: Sonata, Avanto, Aera). T1-weighted turbo-spin-echo sequences with the following parameters were acquired: repetition time 450 – 890 ms, echo time 11 – 20 ms, slice thickness 4 – 10 mm, matrix 256 – 512 × 141 – 314, total examination time 5 – 15 minutes.

For standard pelvimetry, sequences were acquired in two orientations: in a sagittal and in a paratransverse or paracoronal (depending on the pelvic angulation; henceforth referred to as 'transverse') angulation. Subsequently, a sagittal slice through the cartilaginous symphysis pubis was identified and used to determine the conjugata vera (CV) as the shortest distance between the posterior border of the symphysis to the superior anterior aspect of the sacral promontory (▶ **Fig. 1a**). In addition, a transverse slice through the femoral heads was identified and used to yield the diameter transversalis (DT) as the largest transverse distance of the pelvic inlet (▶ **Fig. 1b**).

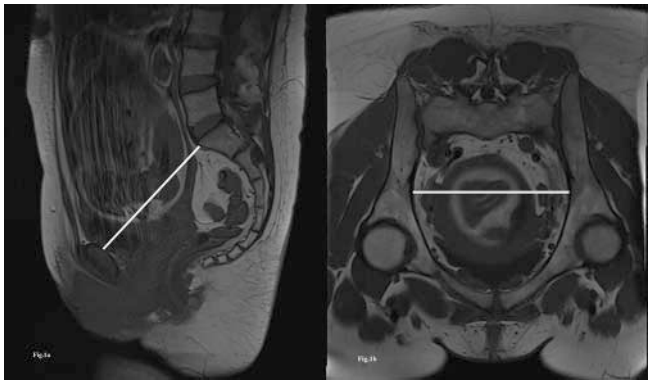
Interobserver study

Observers consisted of three readers from the departments of radiology and obstetrics, respectively: one attending (Rad./Obs.1), one fellow (Rad./Obs.2), and one junior resident (Rad./Obs.3). Initially, all observers completed the same five training cases according to a standardized protocol. Afterward, each observer independently evaluated the 99 breech study cases in a randomized order without knowledge of previous measurements or mode of delivery. Measurements were conducted at an integrated picture archiving and communication (PACS) work station (MMWP, Siemens Healthineers, Germany) with the same software package (Syngo Imaging, Siemens Healthineers, Germany).

Recorded variables were CV and DT, as they represent the standard parameters for obstetric pelvimetry in our hospital, and any incidental findings found while reviewing the MR sequences. To capture differences in DT measurement due to varying quality of MR sequences, the navigation tool was used to evaluate the angulation of the transverse sequence in relation to the CV. For correct DT measurement, the transverse plane should be parallel to the pelvic inlet; an excessive tilt may overestimate DT. Thus an angle of 0 – 20° was considered good, an angle > 20° poor and the prevalent rating among observers was taken as 'true' quality of the sequence. Mean measuring errors were then compared between both angulation groups.

In the absence of a gold standard of measurement, the mean of all six observers served as the reference value.

As previously described [16], diagnostic accuracy was assessed by applying a limit for measurement error of 0.5 cm from the mean.



▶ **Fig. 1** MR pelvimetry, **a** measurement of conjugata vera in midsagittal plane, **b** measurement of diameter transversalis in transverse plane.

▶ **Abb. 1** MR-Pelvimetrie, **a** Messung der Conjugata vera in mittsagittaler Ebene, **b** Messung des Diameter transversalis in transversaler Ebene.

Reference value assessment

Study cases were screened for complete obstetric records and categorized according to mode of delivery. Documented characteristics comprised age, parity, height, pre-pregnancy weight and body mass index (BMI) of the mother and head circumference, length and five-minute APGAR of the neonate. Successful TOL resulting in VD represented the primary aim. Failed TOL resulting in a CS was termed unplanned CS (uCS). Recommended CS (rCS) was defined as intended CS due to inadequate pelvimetric results irrespective of any spontaneous onset of labor prior to the scheduled operation date. Elective CS (eCS) referred to women whose pelvimetric results had been judged adequate for TOL but who opted for a scheduled CS. In order to assess reference values, mean pelvimetric measurements were compared among modes of delivery.

Statistics

Statistical Package for Social Sciences (SPSS version 23, IBM, USA) was used for statistical analysis.

Continuous variables were displayed as means and standard deviation (± SD) and categorical data as percentages. Normality testing was performed with Shapiro-Wilk and a p-value < 0.05 was considered statistically significant. As most variables lacked distribution of normality, non-parametric testing (Kruskal-Wallis, Mann-Whitney-U-Test) was used throughout.

To assess interobserver agreement, the intraclass correlation coefficient (ICC) was calculated for continuous variables [19]. Assuming firstly an effect of both rater and case and secondly that both are drawn randomly from larger populations, the 'two-way random, single measure' ICC (2,1) was chosen. Since systematic variability was to be treated as relevant, the focus was set to absolute agreement among raters. The resulting ICC would represent the reliability of a single typical rater compared to the mean. Regarding interpretation, Lee et al. challenged the mean ICC as a sufficient score of agreement and suggested that 'meaningful

▶ **Table 1** Intraclass correlation coefficients and 95 % confidence intervals for conjugata vera and diameter transversalis measurements.

▶ **Tab. 1** Intraklassen-Korrelationskoeffizienten und 95 %-Konfidenzintervalle für Messungen von Conjugata vera und Diameter transversalis.

Comparison among observers		CV	DT
		ICC (95 % CI)	ICC (95 % CI)
All raters		0.961 (0.939 – 0.975)	0.916 (0.890 – 0.938)
Departments	Radiology	0.968 (0.954 – 0.978)	0.932 (0.905 – 0.952)
	Obstetrics	0.957 (0.865 – 0.981)	0.895 (0.857 – 0.925)
Levels of experience	Attendings	0.967 (0.950 – 0.978)	0.943 (0.916 – 0.961)
	Fellows	0.954 (0.685 – 0.984)	0.889 (0.839 – 0.924)
	Junior residents	0.968 (0.953 – 0.979)	0.946 (0.920 – 0.963)

ICC = intraclass correlation coefficient, CI = confidence interval, CV = conjugata vera, DT = diameter transversalis.
ICC = Intraklassen-Korrelationskoeffizient, CI = Konfidenzintervall, CV = Conjugata vera, DT = Diameter transversalis.

agreement is attained if the lower limit of the 95 % confidence interval [..] is at least 0.75' [20].

Modified Bland-Altman plots were drawn to illustrate individual deviance of raters from the mean, as outlined [21]. One-sample t-tests gave systematic bias (mean of differences) as well as standard deviation, from which limits of agreement were calculated, using the following equation: mean ± 1.96 × SD.

Results

Review of MR referrals revealed 115 MR pelvimetry reports. Yet, original DICOM (Digital Imaging and Communications in Medicine) images were only retrievable from May 2001 onward (n = 101). Twin pregnancies were excluded (n = 2). This yielded 99 eligible MR datasets with singleton in breech for the interobserver study.

For the assessment of reference values, women who had delivered externally (n = 9) and cases with spontaneous cephalic version (n = 3) were excluded. Thus 87 breech cases at birth and known mode of delivery were included in the retrospective reference value assessment.

Interobserver Agreement

Measurements of CV and DT were accurate in 98.1 % and 97.1 % of the cases, respectively. The total observation time ranged from 180 to 275 minutes. The mean evaluation time per case was similar for obstetricians and radiologists (2 min 16 s and 2 min 3 s, respectively).

Interobserver agreement was strong throughout. In detail, measurements of CV and DT by all six raters yielded ICCs of 0.961 and 0.916, respectively. Inter-departmental comparison between radiologists and obstetricians showed ICCs of 0.968 and 0.957 for CV, 0.932 and 0.895 for DT, respectively. Comparison among attendings, fellows and junior residents, showed ICCs of 0.967, 0.954, and 0.968 for CV, and 0.943, 0.889, and 0.946 for DT, respectively. All ICCs are shown in ▶ **Table 1**.

Bland-Altman analysis showed systematic biases were close to zero for all observers. The highest systematic bias was found for the obstetric fellow (0.21 cm for CV, 0.07 cm for DT), whose measurements were generally larger than the mean. Bland-Altman analyses are summarized in ▶ **Table 2**. ▶ **Fig. 2a–f** show scatter plots of CV measurements by each individual observer. As corresponding plots of DT measurements closely resemble ▶ **Fig. 2a–f**, they were omitted.

Evaluation of angulation ratings revealed 86 (87 %) good and 13 (13 %) poor transverse sequences. The mean measuring error of DT was 0.14 ± 0.07 cm in good sequences and 0.18 ± 0.14 cm in poor sequences; the difference was not statistically significant (p-value: 0.053).

Incidental findings

Initially, 140 incidental findings were noted. Six findings of metal implants such as spondylodesis were excluded since they could be considered as previously known. Thus, 134 findings were noted in 69 of 99 study cases. In detail, radiologists detected 86, 24, and 17, while obstetricians detected four, one, and two findings (attending, fellow, junior resident, respectively). In a second step, duplicate findings in the same case were removed, yielding 101 individual findings. When taken together, the four most frequent findings constituted 67 % of all findings: lumbar disc herniation or bulging, suspected coccygeal fracture, pelvic or genital varicosis and epidural lipomatosis. Obstetricians reported 'coccygeal kink', lumbar disc herniation and sacralization of L5; the 15 other findings were only reported by radiologists. Regarding clinical significance, none of the incidental findings was judged suspicious. Thus no further investigation or action was required. ▶ **Table 3** lists all incidental findings and the number of cases identified.

Reference Value Assessment

The demographic data of the women and their neonates are shown in ▶ **Table 4**. Based on pelvimetric results, the obstetrician in charge had recommended CS (rCS) due to an inadequate pelvis

▶ **Table 2** Bland-Altman analysis of conjugata vera and diameter transversalis measurements.

▶ **Tab. 2** Bland-Altman-Analyse der Messungen von Conjugata vera und Diameter transversalis.

Observers		Bias	SD	Limits of agreement[1]	
				lower	upper
Conjugata vera					
Obstetric	Attending	0.044	0.154	– 0.258	0.346
	Fellow	– 0.212	0.156	– 0.518	0.094
	Junior resident	0.044	0.135	– 0.221	0.309
Radiologic	Attending	0.101	0.165	– 0.222	0.424
	Fellow	0.014	0.143	– 0.266	0.294
	Junior resident	0.009	0.180	– 0.344	0.362
Diameter transversalis					
Obstetric	Attending	0.017	0.160	– 0.297	0.331
	Fellow	– 0.069	0.300	– 0.657	0.519
	Junior resident	0.021	0.177	– 0.326	0.368
Radiologic	Attending	0.037	0.188	– 0.331	0.405
	Fellow	– 0.042	0.176	– 0.387	0.303
	Junior resident	0.036	0.176	– 0.309	0.381

SD = standard deviation.
SD = Standardabweichung.
[1] ($\bar{x} \pm 1.96$x SD).
($\bar{x} \pm 1.96$x SD).

in eleven cases and had offered the option of VD in 76 cases. Of these 76 women, nine opted for an eCS despite there being no pelvimetric objections and two had medical indications for CS, namely one case of placental insufficiency and one case of complicated gestational diabetes, and were subsequently counted in the eCS group. Of 65 women who decided on a TOL, 46 (71 %) delivered vaginally and 19 (29 %) ultimately had a uCS. The reasons for uCS were: nine (47 %) cases of labor arrest in opening phase, seven (37 %) cases of fetal distress, and three (16 %) cases of suspected FPD.

Groups of different modes of delivery and their mean pelvimetric measurements are shown in ▶ **Table 5**. In summary, the group rCS showed significantly smaller CV measurements compared to all other groups (p-value < 0.0001). DT measurements only showed significant differences between rCS compared to eCS (p-value 0.039). We found no significant differences in pelvic inlet measurements among the groups VD, uCS and eCS.

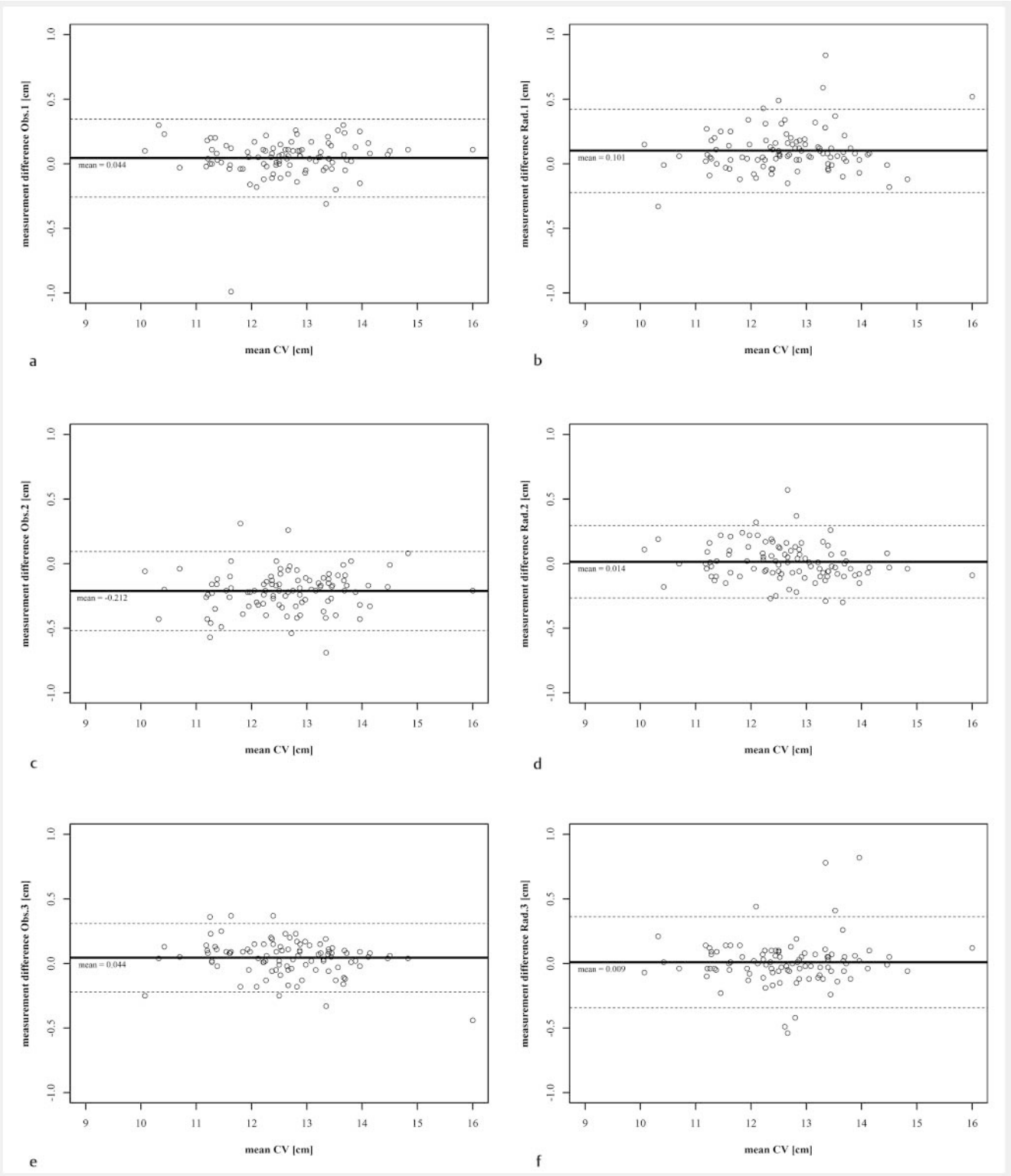
Discussion

In this study, we examined the interobserver reliability of prepartal MR pelvimetry among radiologists and obstetricians with different levels of experience. Having found strong interobserver agreement, we could confirm MR pelvimetry as a reliable tool for better patient selection for VD. Our study also revealed significant differences in CV measurements between TOL candidates and the group rCS, advancing the pursuit of clinically applicable reference values. The use of MR is generally limited by its availability, higher cost, and longer examination time compared to manual examination and ultrasound. However, previous studies indicated that prepartal MR pelvimetry is associated with lower rates of unplanned CS [11, 18]. Therefore, it could be argued that MR pelvimetry is justified when supposedly unnecessary CS and associated morbidity can be avoided.

Interobserver Agreement

The diagnostic accuracy of observers' measurements was > 97 % for both CV and DT. These results correspond to Korhonen's study of mixed vertex and breech presentations [16] who found accurate measurements in 96 % of CV and in 99 % of DT measurements. This validates the choice of CV and DT as reliable pelvimetric parameters.

Mean agreement among all observers was strong throughout and thus comparable to previous studies [15, 16]. In detail, radiologic measurements were slightly more reliable than obstetric ones and measurements of CV more reliable than those of DT. Interestingly, the mean agreement among clinicians with similar levels of experience was almost identical for attendings and junior residents of both departments; only the agreement among fellows was marginally lower. Also, all lower bounds of 95 % confi-



▶ **Fig. 2** Bland-Altman plots of observers' conjugata vera measurements. X-axis shows overall mean measurement, y-axis shows difference between overall mean and each individual observer's measurement; systematic bias (bold line), variance (circles) and limits of agreement (dashed line). Obs./Rad. 1 = obstetric/radiologic attending, Obs./Rad. 2 = obstetric/radiologic fellow, Obs./Rad. 3 = obstetric/radiologic junior resident, CV = conjugata vera.

▶ **Abb. 2** Bland-Altman-Diagramme der Conjugata-vera-Messungen durch die Auswerter. X-Achse zeigt Durchschnittsmessung aller Auswerter, Y-Achse zeigt Differenz zwischen Durchschnittsmessung und Messungen der einzelnen Auswerter; systematischer Messfehler (kräftige Linie), Streuung (Kreise) und Übereinstimmungsgrenzen (gestrichelte Linie). Obs./Rad. 1 = geburtshilflicher/radiologischer Oberarzt, Obs./Rad. 2 = geburtshilflicher/radiologischer Facharzt, Obs./Rad. 3 = geburtshilflicher/radiologischer Assistenzarzt, CV = Conjugata vera.

▶ **Table 3** List of incidental findings and times diagnosed.

▶ **Tab. 3** Liste der Nebenbefunde und Häufigkeiten.

Incidental findings	N = 101
Lumbar disc herniation or bulging	20
'Coccygeal kink' (suspected fracture)	20
Pelvic or genital varicosis	15
Epidural lipomatosis	12
Partial or total sacralization of L5	7
Signs of osteochondrosis	5
Perineural cyst	4
Enlarged inguinal lymph node	3
Disc herniation into vertebral body (persistent chorda dorsalis or suspected Scheuermann's disease)	3
Suspected sacral fracture	2
End plate deformity	2
Labial cyst	2
Loosening of symphysis pubis (suspected symphysitis)	1
Suspected hypogenesis of L1	1
Large rectal ampulla of fetus	1
Focal liponecrosis	1
Spondylophyte	1
Fetal cephalic presentation (spontaneous version)	1

dence intervals were > 0.75, except for CV measurements by fellows (95 % CI: 0.685 – 0.984). This finding corresponded to results of our Bland-Altman analyses, which revealed small systematic biases per observer overall, but a comparatively larger systematic bias (-0.2 cm) with CV measurements by the obstetric fellow. While still within the acceptable range of error, further training might clarify potential uncertainties regarding correct measurement. Korhonen [16] found similar results with the greatest bias being 0.3 cm.

While variances in DT measurements did not show a statistically significant difference between well and poorly angulated transverse sequences, the p-value of 0.053 does leave room for speculation that with a larger sample size, the difference in variance may become significant. However, the relevance of small millimeter differences in measurement has been contested [16]. Still, we consider accurate acquisition of MR sequences an important prerequisite for valuable interpretation. Overall, our findings support the notion that, after adequate training, both radiologists and obstetricians can reliably evaluate MR pelvimetry.

▶ **Table 4** Demographic data of women and their neonates.

▶ **Tab. 4** Demografische Daten der Frauen und Neugeborenen.

Characteristics	n = 87
Maternal age (mean, range) [years]	32 (22 – 43)
Parity, nulliparous/multiparous (n, %)	65/22 (75/25)
Maternal weight[1] (mean, range) [kg]	67.6 (50 – 108)
Maternal height (mean, range) [m]	1.69 (1.56 – 1.80)
Maternal body mass index[1] (mean, range) [kg/m^2]	23.7 (17.9 – 37.8)
Gestational age at MR (mean, range) [weeks]	37.1 (33.8 – 40.6)
Neonatal weight (mean, range) [g]	3222 (2140 – 4245)
Neonatal head circumference (mean, range) [cm]	35.0 (31.0 – 38.5)
Neonatal length (mean, range) [cm]	51.0 (45.0 – 57.0)
Five-minute APGAR (mean, range)	9.4 (4 – 10)

[1] before pregnancy.
vor der Schwangerschaft.

Incidental findings

Radiologists noted approximately 20 – 30 times more findings than their obstetric counterparts, demonstrating clear superiority of radiological perceptiveness for incidental findings. Yet, the incidental findings noted in our sample were of no direct clinical significance. A recent study [22] retrospectively assessing maternal incidental findings in fetal MR examinations classified the majority (90.5 %) of all findings as having little or no clinical significance. While this appears appropriate for a young and healthy study collective, case numbers are too small to rule out clinically more significant findings in general. It should be noted that in our study all findings were included without further assessment of correctness and regardless of whether they were previously known or not, rendering the term 'incidental' debatable.

The evaluation of findings in our study highlights the incidence of further pelvic diagnoses potentially relevant for delivery. Sacralization of L5 could alter angles of pelvic inlet and aperture [23], which in turn might affect fetal engagement; previous fractures might compromise pelvic capacity and subsequently limit cardinal movements. Therefore, all readers should be trained to look for these findings and, if present, document them in their report to then allow evaluation.

Reference Values

We found significantly smaller CV measurements in the rCS group compared to those of all other groups. Due to suspected pelvic inadequacy, CS was recommended to women with a mean CV of 11.37 ± 0.73 cm. We thus expect the majority of cases with suspected FPD to be found in the rCS group, rendering pelvimetric comparison between VD and rCS most appropriate. Also, exclusion of most FPD cases from a TOL might explain the lack of pelvimetric differences between successful and failed TOL.

▶ **Table 5** Modes of delivery and pelvic measurements.

▶ **Tab. 5** Geburtsmodi und Beckenmaße.

	Vaginal delivery (VD) (n = 46)	Unplanned cesarean section (uCS) (n = 19)	Elective cesarean section (eCS) (n = 11)	Recommended cesarean section (rCS) (n = 11)
Mean CV ± SD [cm]	12.70 ± 0.88	12.71 ± 0.87	13.20 ± 0.81	11.37 ± 0.73
p-value[1]	(< 0.0001)	(< 0.0001)	(< 0.0001)	
Mean DT ± SD [cm]	13.39 ± 0.77	13.41 ± 0.73	13.70 ± 0.45	13.01 ± 0.82
p-value[1]	(0.257)	(0.149)	(0.039)	

CV = conjugata vera, DT = diameter transversalis, SD = standard deviation.
CV = Conjugata vera, DT = Diameter transversalis, SD = Standardabweichung.
[1] compared to 'recommended cesarean section'.
im Vergleich zu „Kaiserschnitt empfohlen" (rCS).

A limitation of our study was the selection bias due to the partially retrospective study design. For one, women who did not consider VD or had contraindications did not undergo MR pelvimetry in the first place; for another, MR pelvimetry was used to select women with small pelvimetric measurements for CS, avoiding a TOL.

Only three of 19 uCS cases (16 %) were due to FPD. Re-examination of these three cases revealed that two women had chosen a TOL despite the caution of poor pelvimetry results (CV: 10.70 cm and 11.61 cm, respectively), and a third woman, albeit with a favorable pelvic inlet (CV: 13.96 cm, DT: 14.05 cm), demonstrated a sacrum with notable lack of curvature and in turn a small midpelvis, allowing speculation of labor arrest due to midpelvic dystocia.

Finding a mean CV of 11.37 ± 0.73 cm in rCS cases corresponds to two studies investigating FDP in vertex presentation [14, 15] that found mean CV values for their respective dystocia groups of 11.3 ± 0.9 cm and 11.4 ± 1.0 cm. In comparison, a cut-off value of 11.0 cm for CV in van Loon's study was rather permissive. Similarly, Maharaj et al. [8] stated threshold values for CV (average: 11 cm; contracted: < 10 cm) and DT (average: 13.5 cm; inadequate: < 12 cm) that hardly find application in our study collective. In contrast, a study assessing CT pelvimetry in the selection of breech cases for TOL defined CV > 12 cm as adequate [24]. Bearing in mind that our TOL candidates had a mean CV of 12.70 ± 0.87 cm and rCS cases had a mean CV of 11.37 ± 0.73 cm, a CV of 12.0 cm appears to be a likely borderline value.

In our study DT measurements added no information to the prognosis of VD. Therefore, studies on other pelvic distances and their potential as a selection marker for VD such as the midpelvic interspinous distance [25] should be further examined.

Conclusion

Strong interobserver agreement of pelvimetric measurements between and among radiologists and obstetricians with different levels of experience was found. MR pelvimetry seemed to be a reliable tool for identifying the risk for fetopelvic disproportion in women with a fetus in breech presentation. Radiologic expertise is vital for a comprehensive evaluation of incidental findings even though they were considered benign in this cohort. Our results support the usefulness of conjugata vera measurements for patient selection for vaginal breech delivery, while diameter transversalis measurements added no value. The proposed reference value of 12.0 cm for CV in our cohort should be interpreted as an approximation.

CLINICAL RELEVANCE

- Interobserver reliability was strong between and among radiologists and obstetricians.
- MR pelvimetry is a reliable method to evaluate the viability of vaginal breech delivery.
- During appraisal, it is important to look for incidental findings, as they are fairly common.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

[1] Lansac J, Crenn-Hebert C, Riviere O et al. How singleton breech babies at term are born in France: a survey of data from the AUDIPOG network. Eur J Obstet Gynecol Reprod Biol 2015; 188: 79 – 82. doi:10.1016/j.ejogrb.2015.02.029

[2] Sanchez-Ramos L, Wells TL, Adair CD et al. Route of breech delivery and maternal and neonatal outcomes. Int J Gynaecol Obstet 2001; 73: 7 – 14. doi:S0020-7292(00)00384-2 [pii]

[3] Silver RM, Landon MB, Rouse DJ et al. Maternal morbidity associated with multiple repeat cesarean deliveries. Obstet Gynecol 2006; 107: 1226 – 1232. doi:10.1097/01.AOG.0000219750.79480.84

[4] Hansen AK, Wisborg K, Uldbjerg N et al. Risk of respiratory morbidity in term infants delivered by elective caesarean section: cohort study. BMJ 2008; 336: 85 – 87. doi:10.1136/bmj.39405.539282.BE

[5] Patenaude Y, Pugash D, Lim K et al. The use of magnetic resonance imaging in the obstetric patient. J Obstet Gynaecol Can 2014; 36: 349 – 363

[6] Zaretsky MV, Alexander JM, McIntire DD et al. Magnetic resonance imaging pelvimetry and the prediction of labor dystocia. Obstet Gynecol 2005; 106: 919 – 926. doi:106/5/919 [pii]10.1097/01.AOG.0000182575.81843.e7

[7] Katanozaka M, Yoshinaga M, Fuchiwaki K et al. Measurement of obstetric conjugate by ultrasonic tomography and its significance. Am J Obstet Gynecol 1999; 180: 159 – 162

[8] Maharaj D. Assessing cephalopelvic disproportion: back to the basics. Obstet Gynecol Surv 2010; 65: 387 – 395. doi:10.1097/OGX.0b013e3181ecdf0c

[9] Griffiths M. Magnetic-resonance pelvimetry in breech presentation. Lancet 1998; 351: 912; author reply 913. doi:S0140-6736(05)70332-9 [pii]10.1016/S0140-6736(05)70332-9

[10] Berger R, Sawodny E, Bachmann G et al. The prognostic value of magnetic resonance imaging for the management of breech delivery. Eur J Obstet Gynecol Reprod Biol 1994; 55: 97 – 103

[11] van Loon AJ, Mantingh A, Serlier EK et al. Randomised controlled trial of magnetic-resonance pelvimetry in breech presentation at term. Lancet 1997; 350: 1799 – 1804. doi:S0140-6736(97)05431-7 [pii]10.1016/S0140-6736(97)05431-7

[12] Hannah ME, Hannah WJ, Hewson SA et al. Planned caesarean section versus planned vaginal birth for breech presentation at term: a randomised multicentre trial. Term Breech Trial Collaborative Group. Lancet 2000; 356: 1375 – 1383. doi:S0140673600028403 [pii]

[13] Goffinet F, Carayol M, Foidart JM et al. Is planned vaginal delivery for breech presentation at term still an option? Results of an observational prospective survey in France and Belgium. Am J Obstet Gynecol 2006; 194: 1002 – 1011. doi:S0002-9378(05)02440-3 [pii]10.1016/j.ajog.2005.10.817

[14] Sporri S, Hanggi W, Braghetti A et al. Pelvimetry by magnetic resonance imaging as a diagnostic tool to evaluate dystocia. Obstet Gynecol 1997; 89: 902 – 908. doi:S0029784497001488 [pii]

[15] Keller TM, Rake A, Michel SC et al. Obstetric MR pelvimetry: reference values and evaluation of inter- and intraobserver error and intraindividual variability. Radiology 2003; 227: 37 – 43. doi:10.1148/radiol.2271011658 2271011658 [pii]

[16] Korhonen U, Solja R, Laitinen J et al. MR pelvimetry measurements, analysis of inter- and intra-observer variation. Eur J Radiol 2010; 75: e56 – e61. doi:10.1016/j.ejrad.2009.11.018S0720-048X(09)00649-4 [pii]

[17] Korhonen U, Taipale P, Heinonen S. The diagnostic accuracy of pelvic measurements: threshold values and fetal size. Arch Gynecol Obstet 2014; 290: 643 – 648. doi:10.1007/s00404-014-3271-z

[18] Franz M, von Bismarck A, Delius M et al. MR pelvimetry: prognosis for successful vaginal delivery in patients with suspected fetopelvic disproportion or breech presentation at term. Arch Gynecol Obstet 2017; 295: 351 – 359. doi:10.1007/s00404-016-4276-6

[19] Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979; 86: 420 – 428

[20] Lee J, Koh D, Ong CN. Statistical evaluation of agreement between two methods for measuring a quantitative variable. Comput Biol Med 1989; 19: 61 – 70

[21] Jones M, Dobson A, O'Brian S. A graphical method for assessing agreement with the mean between multiple observers using continuous measures. Int J Epidemiol 2011; 40: 1308 – 1313. doi:10.1093/ije/dyr109

[22] Abdullah SB, Dietz KR, Holm TL. Fetal MRI: incidental findings in the mother. Pediatr Radiol 2016; 46: 1736 – 1743. doi:10.1007/s00247-016-3680-1

[23] Buttenberg D. Labor mechanical observations on the pelvic canal. Arch Gynakol 1962; 197: 172 – 207

[24] Jeyabalan A, Larkin RW, Landers DV. Vaginal breech deliveries selected using computed tomographic pelvimetry may be associated with fewer adverse outcomes. J Matern Fetal Neonatal Med 2005; 17: 381 – 385. doi:H734R2P133604736 [pii]10.1080/14767050500124051

[25] Hoffmann J, Thomassen K, Stumpp P et al. New MRI Criteria for Successful Vaginal Breech Delivery in Primiparae. PLoS One 2016; 11: e0161028. doi:10.1371/journal.pone.0161028