Clinical Radiology xxx (xxxx) xxx

Contents lists available at ScienceDirect

Clinical Radiology

journal homepage: www.clinicalradiologyonline.net

Post-mortem radiology in fetal and neonatal death: the diagnostic value of post-mortem MRI versus autopsy regarding non-cardiac thoracic and abdominal abnormalities

M.P.M. Tijssen [a,*,†], C.A.S. Gregoire [b,†], S.G.F. Robben [a], C.A.H. Severens-Rijvers [c], W.M. Klein [d], P.A.M. Hofman [a]

[a] Department of Radiology, Maastricht University Medical Center, Maastricht, the Netherlands
[b] Maastricht University, Faculty of Medicine, Maastricht, the Netherlands
[c] Department of Pathology, GROW, Maastricht University Medical Center, Maastricht, the Netherlands
[d] Department of Radiology, Radboud University Medical Center, Nijmegen, the Netherlands

ARTICLE INFORMATION

Article history:
Received 17 March 2023
Received in revised form
31 May 2023
Accepted 26 July 2023

AIM: To compare the diagnostic value and accuracy of post-mortem magnetic resonance imaging (PMMRI) and autopsy for non-cardiac thoracic and abdominal abnormalities in fetal death.

MATERIALS AND METHODS: This single-institution retrospective study included all consecutive cases of fetal and perinatal death between January 2015 and December 2021 for which PMMRI followed by autopsy was conducted. These cases comprised fetuses at >18 weeks of gestation and preterm and term neonates who lived for <24 h. All PMMRI and autopsy reports were re-assessed and scored for seven non-cardiac thoracic and 52 abdominal abnormalities, and concordance between autopsy and PMMRI findings was determined as the primary outcome.

RESULTS: Eighty cases were included in this study. Fetal loss was caused by termination of pregnancy in 80% of cases. Further, the mean gestational age was 166 days (23 weeks and 5 days, range 126–283 days). The concordance between PMMRI and autopsy for non-cardiac thoracic and abdominal abnormalities was 83.1% (95% confidence interval [CI] 71.3–83.3) and 76.3% (95% CI 65.8–84.2%), respectively, with a substantial and moderate strength of agreement (Cohen's kappa = 0.63 and 0.51 respectively).

CONCLUSION: PMMRI exhibited good overall diagnostic value for non-cardiac thoracic and abdominal abnormalities, specifically large structural abnormalities. PMMRI may offer parents and physicians a valuable addition to autopsy for the detection of non-cardiac thoracic and abdominal abnormalities, or even an alternative option when parents do not consent to autopsy.

* Guarantor and correspondent: M.P.M. Tijssen, Department of Radiology and Nuclear Medicine, Maastricht University Medical Center, P. Debyelaan 25, P.O. Box 5800, 6202 AZ Maastricht, the Netherlands. Tel.: +31(0)43 3874910, Fax: +31(0)43 3876909.
E-mail address: mpm.tijssen@mumc.nl (M.P.M. Tijssen).
† These authors contributed equally to this work.

https://doi.org/10.1016/j.crad.2023.07.021
0009-9260/

Introduction

Autopsy remains the reference standard for determining the cause of fetal death. The rates of perinatal autopsy has declined in recent years[2,3] with autopsies performed in only 50% of eligible cases in the UK[4] and 20.9% in the USA in the last decade.[5] Post-mortem magnetic resonance imaging (PMMRI) is gaining popularity. The less invasive nature of PMMRI makes parental consent easier to obtain, leading to an improved acceptance rate. Where consent was provided <50% of the time on average for a standard autopsy, post-mortem radiology consent rates varied between 80% and 100%[1,6,7]; however, the diagnostic value of PMMRI requires further investigation.

Several studies have examined the diagnostic accuracy and limitations of fetal PMMRI compared with those of autopsy. The largest prospective validation study to date, in which 277 fetuses were analysed, showed a concordance between minimally invasive and conventional autopsy of 94.9%.[8–11] Therefore, conventional autopsy may not have been required in these cases. Minimally invasive autopsy in this study comprised post-mortem blood sampling and examination of the placenta as well as PMMRI[11]; PMMRI alone had a lower concordance with conventional autopsy than did minimally invasive autopsy (49.5% versus 94.9%). A review study summarised the diagnostic accuracy of PMMRI across 26 studies conducted between 1996 and 2019 (including a total of 2,706 fetuses) and reported an overall diagnostic accuracy of 77–94%.[12]

Differences in diagnostic accuracy between approaches can, in part, be explained by the differences in sensitivity and specificity for specific organ systems. Respectively, the sensitivity and specificity of PMMRI for detecting brain malformations was 85.7% and 89.7% for fetuses <24 weeks, and 90% and 95.9% for those >24 weeks old[15]; however, the sensitivity of PMMRI is lower for abdominal malformations (56–90%), non-cardiac thoracic deformities (41–82%), and cardiac disorders (48–65%).[8,9,12,13]

The diagnostic value of PMMRI for non-cardiac thoracic and abdominal abnormalities has been less studied than that for cardiac or neurological abnormalities.[12] Nevertheless, thoracic and abdominal abnormalities occur in >20% of cases[16] of termination of pregnancy (TOP), intrauterine fetal death (IUFD), and perinatal death (PD). Further analysis is necessary to determine the utility of PMMRI for different organ systems. Therefore, the aim of this study was to assess the diagnostic value of PMMRI for non-cardiac thoracic and abdominal abnormalities in fetal death compared with that of conventional autopsy.

Materials and methods

Medical ethical approval

The local Medical Ethical Committee reviewed the research protocol (METC 2020–2447) and determined that, according to Dutch law, the Medical Research Involving Human Subjects Act did not apply to this study.

Study population

This single-institution retrospective study was conducted at the Department of Radiology and Pathology. All consecutive cases of fetal and perinatal death between January 2015 and December 2021 for which whole-body PMMRI was performed were included. These comprised cases of TOP, IUFD, or PD within the first 24 h after birth. The exclusion criteria were as follows: PMMRI without subsequent autopsy, gestational age (GA) < 18 weeks, and referral cases with PMMRI and/or autopsy performed elsewhere. Autopsy was performed as soon as possible after death and preceded by PMMRI. According to the protocol, the bodies were stored in refrigerated compartments at 3–5°C. Demographic data including sex, GA, manner of death (i.e., TOP, IUFD, or PD), maceration grade, post-mortem interval (PMI, defined as the time between birth and PMMRI), and body weight were extracted from medical records. Maceration was graded as follows: grade 0, red discolouration of the skin; grade 1, skin detachment and blistering; grade 2, extensive skin detachment and bloody effusions in the serous cavity; grade 3, cloudy effusions; and grade 4, fetal papyraceus.

MRI protocol

All examinations were performed using a 1.5 or 3 T magnet MRI system (Philips Healthcare, Best, Netherlands). A 3 T magnetic field strength was used in PMMRI whenever possible, with 1.5 T applied otherwise. Axial, coronal, and sagittal views of the body were acquired using T2-weighted sequences with a 2-mm section thickness. Additionally, a sagittal three-dimensional (3D) T1-weighted sequence with a voxel size of 0.65 or 1 mm was obtained.

Diagnostic scoring

Whether a particular diagnosis could be made was evaluated from the MRI images and autopsy reports that corresponded with a diagnosis in a predefined list of major pathologies (Electronic Supplementary Material Appendix S1). Based on the anatomical location and associated abnormalities, seven items were scored for the non-cardiac thorax and 52 for the abdomen. Abnormalities of the systemic thoracic vessels were considered to be within the spectrum of cardiac abnormalities and were not evaluated separately. Each item was scored according to three options: present, absent, or unknown/non-diagnostic. Major pathology was defined as a significant finding related to the final diagnosis. A non-diagnostic score, either from PMMRI or autopsy, resulted in a discordant outcome.

PMMRI data analysis

All PMMRI cases were re-assessed and scored according to the non-cardiac thorax and abdomen categories by one clinically experienced paediatric radiologist (>5 years of experience) using the Sectra IDS7 23.1 image viewer (Sectra AB, Linköping, Sweden). The radiologist was also provided with the GA and basic clinical details. Reassessment of

Please cite this article as: Tijssen MPM et al., Post-mortem radiology in fetal and neonatal death: the diagnostic value of post-mortem MRI versus autopsy regarding non-cardiac thoracic and abdominal abnormalities, Clinical Radiology, https://doi.org/10.1016/j.crad.2023.07.021

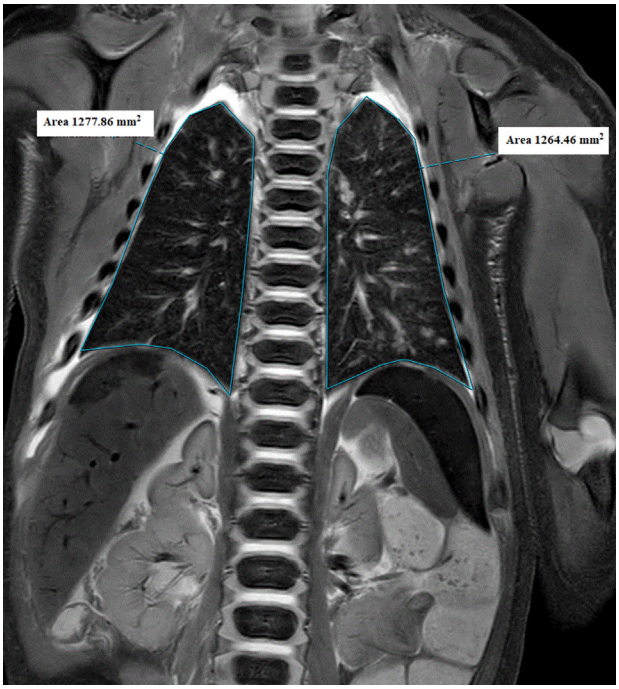

Figure 1 Coronal T2-weighted image with example of lung volume calculation. The lung area of this section is calculated for both lungs. Lung volume is calculated as the sum of the total area of all sections multiplied by section thickness.

PMMRI was performed primarily because of missing data or incomplete descriptions of scoring items in the initial reports. In cases of obvious misinterpretation of the pathology, the initial report was overruled. The radiologist was blinded to the pathology reports to simulate standard clinical practice. Additional measurements were taken to classify pulmonary hypoplasia and hepatomegaly. Lung volumes were calculated on coronal T2-weighted images (area [mm^2] $\times$ section thickness [mm]) (Fig 1) and compared with normal values for the relevant GA.[17] Pulmonary hypoplasia was defined as a lung volume >2 standard deviations (SD) below the mean. Further, hepatomegaly was classified as a midclavicular liver length >95th percentile of the normal length for that GA.[18] All remaining pathologies were scored subjectively.

Autopsy data analysis

Conventional autopsies were performed and reported by experienced paediatric pathologists (all with >5 years of experience). The pathologist was not blinded to the clinical information or initial PMMRI report, as is standard clinical practice at Maastricht University Medical Center. Autopsy reports were compared retrospectively with the same list of diagnostic items and scored according to the options used for the noncardiac thorax and abdomen categories (Electronic Supplementary Material Appendix S1). Pulmonary hypoplasia was defined as lung weight >2 SD below the mean for GA. This study was performed by two readers in cooperation with an experienced paediatric pathologist.

Statistical analyses

The primary outcome was concordance, which was defined as the percentage of cases whose outcome scores (present, absent, or unknown) agreed between PMMRI and autopsy reports. In addition to concordance, Cohen's kappa coefficient was calculated to express the probability-adjusted agreement between the ratings.

Sensitivity, specificity, and positive (PPV) and negative predictive values (NPV) were calculated as secondary outcomes for abnormalities with a prevalence of $\geq$10% (eight or more cases) in the study population, using the autopsy report as reference. Non-diagnostic values were extracted from the total data.

Following anonymisation, the data were recorded and analysed using Microsoft Excel (Microsoft 365 for Windows, version 2016; Microsoft Corporation, Redmond, WA, USA), SPSS software (IBM SPSS Statistics for Windows, version 27.0.1.0; IBM, Armonk, NY, USA), and VassarStats.[19] Statistical significance was defined as $p<0.05$.

Results

Study population

Fig 2 shows a flowchart of the study selection process. Of the 195 cases assessed by PMMRI between January 2015 and December 2021, 33 were excluded because no autopsy had been performed. A further 66 cases were excluded due to low GA, and two more cases were excluded because the neonate lived longer than 24 h. Finally, 14 second-opinion cases from other hospitals were excluded. Both PMMRI and autopsy were performed in the remaining 80 cases, which were selected for the study.

Demographic data

The patient characteristics for the 80 cases are shown in Table 1. The majority were female (56.3%), while 41% were male and 3.7% were of unknown sex. In 80% of cases, the manner of fetal loss was TOP, with IUFD and PD accounting for 11.3% and 9.7%, respectively. Most fetuses had a maceration grade of 0 (55%) or 1 (16.3%). The majority of fetuses

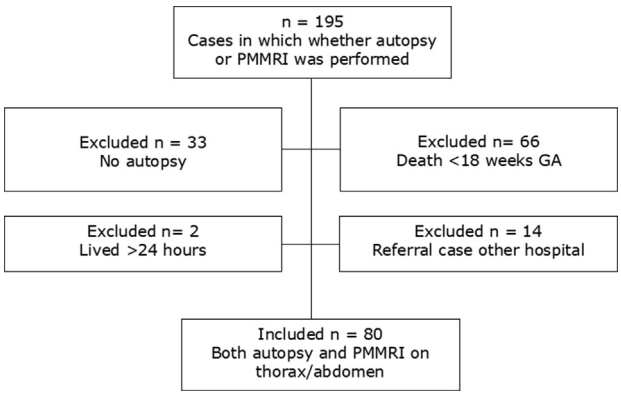

Figure 2 Flowchart of case selection.

Please cite this article as: Tijssen MPM et al., Post-mortem radiology in fetal and neonatal death: the diagnostic value of post-mortem MRI versus autopsy regarding non-cardiac thoracic and abdominal abnormalities, Clinical Radiology, https://doi.org/10.1016/j.crad.2023.07.021

Table 1
Main patient characteristics.

Patient characteristics	n=80 (%)
Gender	
Female	45 (56.3%)
Male	32 (41%)
Unknown	3 (3.7%)
Gestational age	
18–23 weeks	66 (82.5%)
24–29 weeks	5 (6.3%)
30–35 weeks	1 (1.3%)
36–40 weeks	8 (10%)
Origin of fetal loss	
Termination of pregnancy (TOP)	64 (80%)
Intrauterine fetal death (IUFD)	9 (11.3%)
Perinatal death (PD)	7 (9.7%)
Maceration grade	
Grade 0	44 (55%)
Grade 1	13 (16.3%)
Grade 2	11 (13.7%)
Grade 3	2 (2.5%)
Grade 4	0 (0%)
Missing	10 (12.5%)
MRI field strength, T	
1.5	29 (36.3%)
3	51 (63.7%)
Post-mortem interval PMMRI in hours	
<12	9 (11.3%)
12-24	34 (42.5%)
24–48	18 (22.5%)
48–72	8 (10%)
72–96	1 (1.2%)
96–120	0 (0%)
Missing[a]	10 (12.5%)

[a] Date of death is known, but the time of death is not. The mean time interval was estimated at 40 hours (average).

(63.7%) underwent PMMRI at a 3 T magnetic field strength. The PMI from time of death to PMMRI ranged from <12 to 96 h (Table 1). For 10 cases in which the date but not the exact time of death was recorded, the mean PMI was estimated at 40 h. Autopsy-specific patient characteristics, including mean body and organ weights and lengths, are listed in Table 2. The mean GA was 166 days (23 weeks and 5 days).

Table 2
Patient characteristics at autopsy including age and body parameters.

Patient characteristics, n=80	Mean	Minimum	Maximum	Missing (n)
Age				
Gestational age (total days, (weeks + days))	166 (23 + 5)	126 (18 + 0)	283 (40 + 3)	0
Maternal age (years)	31	21	40	6
Weight				
Total body (g)	725.2	142	3710	0
Liver (g)	34.2	2.2	175.1	0
Kidneys (g)	7.6	0.0	109.9	0
Lungs (g)	14.1	0.6	73.5	8
Length				
Crown-heel (cm)	29.8	19.5	54.5	0
Crown-rump (cm)	20.4	13.4	38.0	0

PMMRI data analysis

In six cases, the initial report was overruled during the PMMRI reassessment. One initial report missed a case of intestinal malrotation, two missed kidney fusion (horseshoe kidney), and three did not mention observed hydronephrosis (Fig 3). A few minor items, such as adrenal haemorrhage or ascites, were not always mentioned in the initial report but were noted during reassessment. Other items not mentioned in the initial report were scored as normal or non-diagnostic during reassessment.

PMMRI and autopsy concordance

Concordance was calculated for all 80 cases (Table 3). The mean concordance between PMMRI and autopsy for non-cardiac thoracic abnormalities was 81.3%, with a 95% confidence interval (95% CI) of 71.3–83.3%. Additionally, the concordance for abdominal abnormalities was 76.3% (95% CI 65.8–84.2%). In all 80 cases, a total of 59 diagnostic items were scored as present, absent, or unknown/non-diagnostic. The number of non-diagnostic scores was highest for the “abdominal blood vessels” item (n=11, 13.8%).

Interrater reliability was calculated using Cohen’s kappa (Table 3) to present a probability-adjusted measure of agreement between the PMMRI and autopsy ratings. Cohen’s kappa could not be calculated for items representing abnormalities that did not occur in the study population or for those with observed concordance values lower than the mean-chance concordance. Overall, non-cardiac thoracic exhibited substantial and abdominal abnormalities moderate probability-corrected agreement (kappa = 0.63 and 0.51, respectively); however, an “almost perfect” agreement was computed for renal agenesis (kappa = 1, n=4) and renal cysts (bilateral kappa = 1, unilateral kappa = 1, n=6).

Because autopsy is the reference standard, non-diagnostic items at autopsy were removed from further analysis to avoid any possible false discordance. Following removal of these items, concordance was recalculated for several items and improved; for instance, the concordance for “gallbladder agenesis” increased to 94.9% (95% CI 87.7–98%; Table 4).

Sensitivity, specificity, PPV, and NPV

The secondary outcomes, namely sensitivity, specificity, PPV, and NPV, are shown in Table 5. Unknown/non-diagnostic items were excluded from these analyses. For thoracic abnormalities, the sensitivity was 90.9% (95% CI 74.5–97.6%), specificity was 74.5% (95% CI 59.4–85.6%), PPV was 71.4% (95% CI 55.2–83.8%), and NPV was 92.1% (95% CI 77.5–97.9%). Pulmonary hypoplasia had a relatively high sensitivity of 92% (95% CI 72.4–98.6%) and a specificity of 76.4% (95% CI 62.8–86.3%). Furthermore, 23 true-positive and only two false-negative cases of pulmonary hypoplasia were identified. Abdominal abnormalities were associated with an 86.7% sensitivity (95% CI 72.5–94.5%), 62.9%

Please cite this article as: Tijssen MPM et al., Post-mortem radiology in fetal and neonatal death: the diagnostic value of post-mortem MRI versus autopsy regarding non-cardiac thoracic and abdominal abnormalities, Clinical Radiology, https://doi.org/10.1016/j.crad.2023.07.021

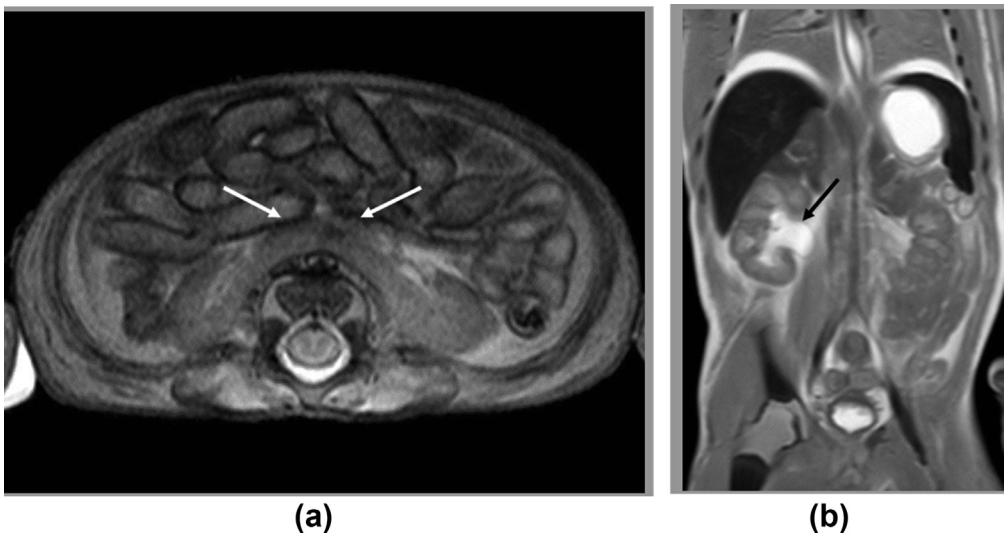

Figure 3 Two examples of misinterpretation in the initial PMMRI report, overruled at reassessment. (a) Axial T2-weighted image of a patient with a horseshoe kidney (white arrows), initially missed. (b) Coronal T2-weighted image of a patient with hydronephrosis of the right kidney (black arrow) not mentioned in the initial report.

specificity (95% CI 44.9–78%), 75% PPV (95% CI 60.1–85.5%), and 78.6% NPV (95% CI 58.8–90.9%). Specifically, the highest sensitivity was observed for renal abnormalities (94.4%; 95% CI 70.6–99.7%), with an associated specificity of 87.1% (95% CI 75.6–93.9%), PPV of 68% (95% CI 46.4–84.3%), and NPV of 98.2% (95% CI 89–99.9%). In contrast, the sensitivity for liver abnormalities was 52.6% (95% CI 29.5–74.8%). Many instances of oedematous and haemorrhagic liver abnormalities were previously missed on PMMRI (12/80 false-negative scores).

Discussion

This study investigated the concordance between PMMRI and conventional autopsy for diagnosing non-cardiac thoracic and abdominal pathologies in fetal and neonatal deaths. Overall, the mean concordance for the thoracic pathologies assessed in this study was 81.3%, with a substantial strength of agreement identified using Cohen's kappa. Previous studies have investigated the concordance between PMMRI and autopsy findings for non-cardiac thoracic abnormalities. Arthurs *et al.*[20] assessed 277 cases and identified thoracic abnormalities by autopsy in 29.8%. Concordance in the <24 weeks gestation group was 82.3% and that in the >24 weeks group was 79.3%; these results are comparable to the present study (81.3%).

To determine pulmonary hypoplasia, the calculated lung volumes were compared with normal ranges using a cut-off of 2 SD. This method has been validated for living fetuses but not for post-mortem imaging. Previous PMMRI publications on thoracic pathology have provided no distinct approaches for diagnosing pulmonary hypoplasia and likely used visual inspection.[20,21] Although the calculation method has not yet been validated for post-mortem cases, it appears to provide a more objective evaluation of pulmonary hypoplasia: with a cut-off of 2 SD, the sensitivity was 92% in the present analyses. The specificity, however, was only 76.4%. Contrastingly, previous studies have reported a lower sensitivity and higher specificity. Breeze *et al.*[21] described a sensitivity of 62.5% and specificity of 87% for total lung abnormalities including pulmonary hypoplasia. Additionally, Arthurs *et al.*[20] reported a sensitivity of 60% and a specificity of 96.6% for pulmonary hypoplasia; however, the present study showed a higher pulmonary hypoplasia prevalence of 31.1% versus 6.1% reported by Arthurs *et al.*[20] and 26.7% by Breeze *et al.*[21] (prevalence of total lung lesions including pulmonary hypoplasia). In the current study, the PPV was 63.9%, and 13 cases of pulmonary hypoplasia were overcalled. This is similar to the PPV of 52.9% Arthurs *et al.* Overcalling of pulmonary hypoplasia is possible because of the comparison of lung weight at autopsy with lung volume on PMMRI. Lung weight is relatively stable with increasing PMI and determination of lung weight remains the reference standard for diagnosing pulmonary hypoplasia. In the liver, volume decreases with increasing PMI.[22] It was suspected that the same processes also influence lung volume: therefore, overcalling of pulmonary hypoplasia may be more likely in post-mortem settings with increasing PMI, particularly, when lung volumes are calculated and no subjective adjustment can be made based on the PMI. For instance, excessive amount of post-mortem pleural fluid can influence the lung volume, whilst the volume of the thoracic cavity can suggest a normal pre-mortem lung volume. Compression of the lungs will not influence the lung weight at autopsy; however, previous studies have reported both increasing[22] and decreasing lung volumes[23] over time.

The mean concordance for abdominal abnormalities in the present study was 76.3%, with a moderate strength of agreement identified using Cohen's kappa. Several previous studies have investigated the concordance of PMMRI and autopsy findings for abdominal abnormalities. Arthurs *et al.*[10] included 250 fetuses, and observed a concordance for abdominal abnormalities of 91.1% in the group with <24

Please cite this article as: Tijssen MPM et al., Post-mortem radiology in fetal and neonatal death: the diagnostic value of post-mortem MRI versus autopsy regarding non-cardiac thoracic and abdominal abnormalities, Clinical Radiology, https://doi.org/10.1016/j.crad.2023.07.021

Table 3
Concordance per scored item, n=80.

Abnormalities	Abnormalities on PMMRI (n)	Abnormalities on autopsy (n)	Non-diagnostic (n)	Concordant cases (n)	Concordance (%, 95% CI)	Cohen's K (95% CI)
Non-cardiac thorax	42	33	0	65	81.3% (71.3–88.3)	0.63 (0.46–0.79)
Pulmonary hypoplasia	36	25	0	65	81.3% (71.3–88.3)	0.61 (0.43–0.78)
Pulmonary haemorrhage	1	5	0	76	95% (87.8–98)	0.32 (0–0.97)
Predisposition disorders	3	5	0	74	92.5% (84.5–96.5)	0.21 (0–0.82)
Pulmonary oedema	4	2	0	78	97.5% (91.3–99.3)	0.66 (0.18–1)
Diaphragmatic hernia	2	3	0	79	98.7% (93.2–99.7)	0.79 (0.39–1)
Oesophageal atresia	1	1	8	71	88.7% (79.9–93.3)	0 (0–1)
Other	1	3	0	78	97.5% (91.3–99.3)	0.49 (0–1)
Abdomen	52	45	0	61	76.3% (65.8–84.2)	0.51 (0.31–0.70)
Gastric malposition	3	3	0	80	100% (95.4–100)	1[1–1]
Intestinal[a]	14	6	0	68	87.1% (77.9–92.9)	0.44 (0.11–0.76)
Atresia[b]	7	3	0	70	92.1% (83.8–96.3)	0.36 (0–0.85)
Malrotation	7	2	5	68	85% (75.5–91.2)	0.19 (0–0.76)
Dilation	6	3	1	76	95% (87.8–98)	0.65 (0.26–1)
Renal/urinary tract	25	18	0	71	88.7% (79.9–93.3)	0.72 (0.54–0.89)
Bilateral agenesis	4	4	0	80	100% (95.4–100)	1[1–1]
Mono kidney	4	3	0	79	98.8% (93.3–99.8)	0.85 (0.56–1)
Kidney fusion	3	2	0	79	98.8% (93.3–99.8)	0.79 (0.39–1)
Malposition	1	2	0	77	96.3% (89.6–98.7)	c
Hydronephrosis	4	2	0	78	97.5% (91.3–99.3)	0.66 (0.18–1)
Bilateral	2	1	0	79	98.7% (93.2–99.7)	0.66 (0–1)
Unilateral	2	1	0	79	98.7% (93.2–99.7)	0.66 (0–1)
Hydroureter	2	3	0	79	98.7% (93.2–99.7)	0.79 (0.39–1)
Bilateral	2	3	0	79	98.7% (93.2–99.7)	0.79 (0.39–1)
Unilateral	0	0	0	80		
Cysts	6	6	0	80	100% (95.4–100)	1[1–1]
Bilateral	3	3	0	80	100% (95.4–100)	1[1–1]
Unilateral	3	3	0	80	100% (95.4–100)	1[1–1]
Tumour	0	0	0	80		
Unilateral	0	0	0	80		
Bilateral	0	0	0	80		
Acquired abnormalities	8	2	0	72	90% (81.4–94.8)	0.17 (0–0.71)
Bladder agenesis	0	2	0	78	97.5% (91.3–99.3)	0 (0–1)
Mega bladder	2	3	0	79	98.7% (93.2–99.7)	0.79 (0.39–1)
Adrenal	10	7	0	65	81.2% (71.3–88.2)	0.02 (0–0.47)
Agenesis	0	0	0	80		
Haemorrhage	10	7	0	65	81.2% (71.3–88.2)	0.02 (0–0.47)
Tumour	0	0	0	80		
Liver	34	19	0	47	58.8% (47.8–68.9)	0.10 (0–0.35)
Hypoplasia	0	0	0	80		
Hypertrophy	22	3	0	61	76.3% (65.9–84.2)	0.19 (0–0.51)
Malposition	6	3	0	75	93.7% (86.1–97.3)	0.42 (0–0.91)
Cysts	1	0	0	79	98.7% (93.2–99.7)	0 (0–1)
Tumour	0	0	0	80		
Bile duct dilation	0	0	0	80		
Haemorrhage/oedema	7	14	0	63	78.9% (68.6–86.3)	0.08 (0–0.47)
Gallbladder agenesis	2	1	7	75	93.7% (86.1–97.3)	0.65 (0.18–1)
Spleen	6	9	0	75	93.7% (86.1–97.3)	0.63 (0.32–0.94)
Agenesis	1	1	1	79	98.7% (93.2–99.7)	1 (0–1)
Polysplenia	1	1	0	80	100% (95.4–100)	1 (0–1)
Splenomegaly	0	4	0	76	95% (87.8–98)	0 (0–0.95)
Cysts	1	0	0	79	98.7% (93.2–99.7)	0 (0–1)
Tumour	0	0	0	80		
Haemorrhage	0	0	0	80		
Malposition	3	3	0	80	100% (95.4–100)	1[1–1]
Abdominal wall	1	1	0	80	100% (95.4–100)	1[1–1]
Omphalocele	1	1	0	80	100% (95.4–100)	1[1–1]
Gastroschizis	0	0	0	80		
Abdominal blood vessels	3	8	11	66	82.5% (72.7–89.2)	0.51 (0.09–0.92)
Other	0	0	0	80		

The number of abnormalities found in both PMMRI and autopsy is also shown, as is the number of non-diagnostic cases.
PMMRI, post-mortem magnetic resonance imaging.

[a] In two cases there would be no other intestinal abnormality due to the omission of anal atresia, in two other cases there would be. This brought the total number of cases to calculate concordance to 78.

[b] Four cases of anal atresia have been deleted. This brought the total number of cases to calculate concordance to 76.

[c] Kappa is not calculated for this data set because observed concordance is smaller than mean-chance concordance.

Please cite this article as: Tijssen MPM et al., Post-mortem radiology in fetal and neonatal death: the diagnostic value of post-mortem MRI versus autopsy regarding non-cardiac thoracic and abdominal abnormalities, Clinical Radiology, https://doi.org/10.1016/j.crad.2023.07.021

Table 4
Corrected concordance and interrater reliability using Cohen's kappa, based on a total of cases 80.

Abnormalities	Non-corrected Concordance (CI 95%)	Corrected concordance (CI 95%)	Corrected Cohen's K (CI 95%)	Corrected standard error	Corrected strength of agreement
Abdomen					
Gallbladder agenesis	93.7% (86.1–97.3)	94.9% (87.7–98)	0.65 (0.18–1)	0.2421	Substantial
Abdominal blood vessels	82.5% (72.7–89.2)	79.8% (69.6–87.1)	0.51 (0.09–0.92)	0.2111	Moderate

Concordance was recalculated after exclusion of non-diagnostic autopsy cases to avoid false discordance.

Table 5
Sensitivity, specificity, PPV and NPV for the items, in which prevalence was at least 10% (n=8).

Abnormalities	TP	TN	FP	FN	ND	Sensitivity (%, 95% CI)	Specificity (%, 95% CI)	PPV (%, 95% CI)	NPV (%, 95% CI)
Non-cardiac thorax	30	35	12	3	0	90.9 (74.5–97.6)	74.5 (59.4–85.6)	71.4 (55.2–83.8)	92.1 (77.5–97.9)
Pulmonary hypoplasia	23	42	13	2	0	92.0 (72.4–98.6)	76.4 (62.8–86.3)	63.9 (46.2–78.7)	95.4 (83.3–99.2)
Abdomen	39	22	13	6	0	86.7 (72.5–94.5)	62.9 (44.9–78)	75.0 (60.1–85.5)	78.6 (58.5–90.9)
Intestinal	5	63	9	1	2	83.3 (36.5–99.1)	87.5 (77.1–93.8)	35.7 (13.9–64.4)	98.4 (90.5–99.9)
Renal/urinary tract	17	54	8	1	0	94.4 (70.6–99.7)	87.1 (75.6–93.9)	68 (46.4–84.3)	98.2 (89–99.9)
Liver	10	37	24	9	0	52.6 (29.5–74.8)	60.6 (47.3–72.7)	29.4 (15.7–47.7)	80.4 (65.6–84.3)
Haemorrhage/ Oedema	2	61	5	12	0	14.3 (2.5–43.8)	92.4 (82.3–96.1)	28.6 (5.1–69.7)	83.6 (72.6–90.9)
Spleen	5	70	1	4	0	55.6 (22.7–84.7)	98.6 (91.3–99.9)	83.3 (36.5–99.1)	94.6 (86–98.3)

TP, true positives; TN=true negatives; FP, false positives; FN=false negatives; ND, number of non-diagnostic cases; PPV, positive predictive value; NPV, negative predictive value.

weeks gestation and 84.8% in the group with >24 weeks gestation. Concordance in the present study was lower (76.3%); however, abdominal abnormalities were identified in 56.3% compared with 18.5% of cases in the previous study.[10] This could mean that the prevalence of abnormalities was actually higher in the present patient group or that more minor pathologies were included and/or diagnosed. Both factors likely affected interpretation and negatively influenced the concordance rate.

The concordance in the present study was similar to that reported by Kang *et al.*[8] This study assessed 92 fetuses with a GA ranging from 12–41 weeks. At a magnetic field strength of 3 T, the concordance for non-cardiac thoracic abnormalities was 80.7% and that for abdominal abnormalities 67%; however, at a magnetic field strength of 1.5 T (used for 36.5% of cases in the present study), the concordance for abdominal abnormalities was only 53.8%.[8] Overall, the concordance results for abdominal abnormalities are consistent with those in the present study (76.3%).

Within the category of abdominal findings, the concordance differed between specific abnormalities. Larger structural lesions (e.g., diaphragmatic hernia, renal and splenic agenesis, and abdominal wall defects) showed high concordance (98.7–100%). PMMRI was less accurate for liver disease (concordance of 58.8%), especially hepatomegaly (76.3%). Many cases of oedematous and haemorrhagic liver abnormalities were missed during the initial PMMRI (12/80 false-negative scores). These cases mainly comprised subcapsular haematomas that varied in size and were diagnosed at autopsy. On PMMRI, this pathology may have been misinterpreted as haemorrhagic ascites, which is observed in physiological post-mortem processes (maceration grade 2). Alternatively, these may have been missed due to signal intensity of the haematoma approaching that of normal liver parenchyma (Fig 4). In addition, pathological oedematous liver abnormalities can be difficult to differentiate from physiological post-mortem oedematous changes.

The diagnostic accuracy for renal abnormalities was high in the present study, with almost perfect agreement for renal agenesis and cysts. This is consistent with the findings of Hagmann *et al.*[24]

In the present study, PMMRI assessment of oesophageal atresia and abnormalities of the abdominal blood vessels, including the umbilical vessels, was difficult at times because of the PMMRI resolution resulting in non-diagnostic cases (both 10%). These non-diagnostic PMMRI scores yielded discordant findings when compared with the autopsy reports. In fact, most of the discordant findings for these abnormalities resulted from non-diagnostic scores. Therefore, lower concordance and interrater reliability may be explained by the difficulty in assessing these components using fetal imaging.

The present study had some limitations. The study had a relatively small number of selected cases due to the study centre being rural university hospital located in a sparsely populated area. This may have limited the reliability of the secondary outcomes, namely the sensitivity, specificity, PPV, and NPV. Consequently, these were only calculated for abnormalities with a prevalence of at least eight cases (10%) in the present study population. A longer study period or access to additional databases could provide a larger study population to improve power.

In all cases, PMMRI and autopsy were diagnostic overall, indicating that "present" or "absent" scoring could be accurately achieved. Only a few cases had one or more components that could not be assessed and were scored as unknown/non-diagnostic. Several factors made assessment difficult or impossible in these cases, including profound post-mortem changes such as autolysis, the small dimensions of fetuses and neonates, and incomplete autopsy reports. Notably, the present study compared PMMRI with

Please cite this article as: Tijssen MPM et al., Post-mortem radiology in fetal and neonatal death: the diagnostic value of post-mortem MRI versus autopsy regarding non-cardiac thoracic and abdominal abnormalities, Clinical Radiology, https://doi.org/10.1016/j.crad.2023.07.021

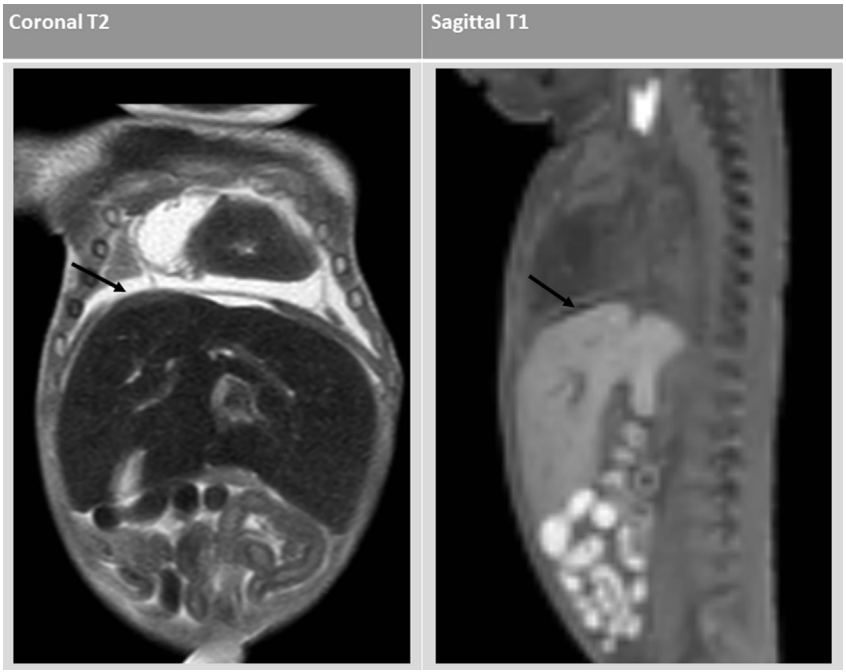

Figure 4 Example of missed subcapsular haematoma on PMMRI (discordant with autopsy, where the subcapsular haematoma is seen). Retrospectively, a small subcapsular haematoma is identified (black arrows), likely initially missed due to signal intensity on T2-weighted images similar to the intensity of normal liver parenchyma.

the autopsy reports and not the autopsies themselves. Because the scoring of autopsy reports was performed retrospectively, there was no opportunity for re-assessment. Moreover, the report descriptions may be interpreted in different ways: for example, a component may not have been described either because it was not assessed or because no abnormality was present. This meant that any unnamed part of the autopsy report automatically led to discordance with PMMRI. Additionally, semantic discrepancies in the autopsy and PMMRI reports and in the scoring according to the diagnostic items may have been present. As a result, the same clinical condition could be referred to using different terms, without negatively effecting clinical relevance, which is also a cause off spurious discordance. To account for this as much as possible, a paediatric pathologist was involved in the interpretation of the reports, although a retrospective pathological assessment of the fetus was not possible. A prospective approach could eliminate these limitations and reduce the chances of underestimation owing to possible false discordance.

Certain abnormalities, such as oedematous and haemorrhagic liver abnormalities, have a higher chance of not being detected or correctly interpreted on PMMRI, as it may be difficult to differentiate between abnormal and normal post-mortem changes. In these cases, an autopsy is required to provide a definitive answer; however, large structural abnormalities and renal abnormalities can be accurately diagnosed using PMMRI. Consequently, PMMRI is valuable for diagnosing many disorders, yielding relevant findings for establishing the final diagnosis. Although autopsy remains the reference standard for determining cause of fetal death, PMMRI is a promising addition or even an alternative when parents do not consent to autopsy, especially given the invasiveness of autopsy and declining autopsy rates.

Future studies should include the prenatal data from ultrasonography, fetal MRI, genetic testing, and obstetric assessment, as well as (minimal invasive) PMMRI and autopsy. Furthermore, the additive value of each diagnostic item and the cost-effectiveness of the diagnostic strategy should be determined in a prospective study setting.

In conclusion, this study revealed that PMMRI had good overall diagnostic value for non-cardiac thoracic and abdominal abnormalities in fetuses aged $>$18 weeks and neonates aged $<$24 h. These results emphasise the utility of PMMRI for detecting macroscopic structural abnormalities. The high accuracy and non-invasive nature of PMMRI offer parents and physicians a valuable addition to autopsy or even an alternative when parents do not consent to autopsy.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

The authors thank statistician B. Winkens for his input and help with statistical testing.

Please cite this article as: Tijssen MPM et al., Post-mortem radiology in fetal and neonatal death: the diagnostic value of post-mortem MRI versus autopsy regarding non-cardiac thoracic and abdominal abnormalities, Clinical Radiology, https://doi.org/10.1016/j.crad.2023.07.021

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.crad.2023.07.021.

References

1. Cannie M, Votino C, Moerman P, *et al*. Acceptance, reliability and confidence of diagnosis of fetal and neonatal virtuopsy compared with conventional autopsy: a prospective study. *Ultrasound Obstet Gynecol* 2012;**39**(6):659–65.
2. Boyd PA, Tondi F, Hicks NR, *et al*. Autopsy after termination of pregnancy for fetal anomaly: retrospective cohort study. *BMJ* 2004;**328**(7432):137.
3. Dickinson JE, Prime DK, Charles AK. The role of autopsy following pregnancy termination for fetal abnormality. *Aust N Z J Obstet Gynaecol* 2007;**47**(6):445–9.
4. Thayyil S, Chitty LS, Robertson NJ, *et al*. Minimally invasive fetal postmortem examination using magnetic resonance imaging and computerised tomography: current evidence and practical issues. *Prenat Diagn* 2010;**30**(8):713–8.
5. Oliver EA, Finneran MM, Rood KM, *et al*. Fetal autopsy rates in the United States: analysis of national vital statistics. *Obstet Gynecol* 2022;**140**(5):869–73.
6. Ben-Sasi K, Chitty LS, Franck LS, *et al*. Acceptability of a minimally invasive perinatal/paediatric autopsy: healthcare professionals' views and implications for practice. *Prenat Diagn* 2013;**33**(4):307–12.
7. Kang X, Cos T, Guizani M, *et al*. Parental acceptance of minimally invasive fetal and neonatal autopsy compared with conventional autopsy. *Prenat Diagn* 2014;**34**(11):1106–10.
8. Kang X, Cannie MM, Arthurs OJ, *et al*. Post-mortem whole-body magnetic resonance imaging of human fetuses: a comparison of 3-T versus 1.5-T MR imaging with classical autopsy. *Eur Radiol* 2017;**27**(8):3542–53.
9. Thayyil S, Sebire NJ, Chitty LS, *et al*. Post mortem magnetic resonance imaging in the fetus, infant and child: a comparative study with conventional autopsy (MaRIAS Protocol). *BMC Pediatr* 2011;**11**:120.
10. Arthurs OJ, Thayyil S, Owens CM, *et al*. Diagnostic accuracy of post mortem MRI for abdominal abnormalities in fetuses and children. *Eur J Radiol* 2015;**84**(3):474–81.
11. Thayyil S, Sebire NJ, Chitty LS, *et al*. Post-mortem MRI versus conventional autopsy in fetuses and children: a prospective validation study. *Lancet* 2013;**382**(9888):223–33.
12. Kang X, Carlin A, Cannie MM, *et al*. Fetal postmortem imaging: an overview of current techniques and future perspectives. *Am J Obstet Gynecol* 2020;**223**(4):493–515.
13. Sonnemans LJP, Vester MEM, Kolsteren EEM, *et al*. Dutch guideline for clinical foetal-neonatal and paediatric post-mortem radiology, including a review of literature. *Eur J Pediatr* 2018;**177**(6):791–803.
15. Arthurs OJ, Thayyil S, Pauliah SS, *et al*. Diagnostic accuracy and limitations of post-mortem MRI for neurological abnormalities in fetuses and children. *Clin Radiol* 2015;**70**(8):872–80.
16. Rossi AC, Prefumo F. Correlation between fetal autopsy and prenatal diagnosis by ultrasound: a systematic review. *Eur J Obstet Gynecol Reprod Biol* 2017;**210**:201–6.
17. Meyers ML, Garcia JR, Blough KL, *et al*. Fetal lung volumes by MRI: normal weekly values from 18 through 38 weeks' gestation. *AJR Am J Roentgenol* 2018;**211**(2):432–8.
18. Tongprasert F, Srisupundit K, Luewan S, *et al*. Normal length of the fetal liver from 14 to 40 weeks of gestational age. *J Clin Ultrasound* 2011;**39**(2):74–7.
19. VasserStats VasserStats: Statistical computation website. Available at: http://vassarstats.net/. [Accessed 13 February 2023].
20. Arthurs OJ, Thayyil S, Olsen OE, *et al*. Diagnostic accuracy of postmortem MRI for thoracic abnormalities in fetuses and children. *Eur Radiol* 2014;**24**(11):2876–84.
21. Breeze AC, Cross JJ, Hackett GA, *et al*. Use of a confidence scale in reporting postmortem fetal magnetic resonance imaging. *Ultrasound Obstet Gynecol* 2006;**28**(7):918–24.
22. Klein Tk WM, Hermans K, Bayat AR, *et al*. The common pattern of postmortem changes on whole body CT scans. *J Forensic Radiol Imaging* 2016;**4**:47–52.
23. Hyodoh H, Shimizu J, Watanabe S, *et al*. Time-related course of pleural space fluid collection and pulmonary aeration on postmortem computed tomography (PMCT). *Leg Med (Tokyo)* 2015;**17**(4):221–5.
24. Hagmann CF, Robertson NJ, Sams VR, *et al*. Postmortem magnetic resonance imaging as an adjunct to perinatal autopsy for renal-tract abnormalities. *Arch Dis Child Fetal Neonatal Ed* 2007;**92**(3):F215–8.

Please cite this article as: Tijssen MPM et al., Post-mortem radiology in fetal and neonatal death: the diagnostic value of post-mortem MRI versus autopsy regarding non-cardiac thoracic and abdominal abnormalities, Clinical Radiology, https://doi.org/10.1016/j.crad.2023.07.021