

This is a repository copy of The rate of brain abnormalities on in utero MRI studies in fetuses with normal ultrasound examinations of the brain and calculation of indicators of diagnostic performance.

White Rose Research Online URL for this paper: http://eprints.whiterose.ac.uk/143221/

Version: Accepted Version

Article:

Griffiths, P., Bradburn, M., Mandefield, L. et al. (2 more authors) (2019) The rate of brain abnormalities on in utero MRI studies in fetuses with normal ultrasound examinations of the brain and calculation of indicators of diagnostic performance. Clinical Radiology. ISSN 0009-9260

https://doi.org/10.1016/j.crad.2019.03.010

Article available under the terms of the CC-BY-NC-ND licence (https://creativecommons.org/licenses/by-nc-nd/4.0/).

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



INTRODUCTION

Recent systematic reviews¹⁻³ and the results from our previous work from the MERIDIAN 2 3 study⁴ have shown that in utero Magnetic Resonance (iuMR) imaging significantly improves the detection of fetal brain abnormalities when compared with antenatal ultrasonography 4 (USS). Specifically, iuMR improves diagnostic accuracy⁴ and diagnostic certainty⁵ when a 5 6 brain abnormality is shown or suspected on USS and those findings are likely to have substantial implications for clinical practice.⁴ An important limitation of those studies is they 7 have not evaluated the impact of iuMR imaging in cases in which no brain abnormality was 8 detected or suspected on USS. The intrinsic value of a diagnostic test relies not only its 9 ability to identify an abnormality correctly when one is present but also to exclude 10 11 abnormalities correctly when they are not present. To date, studies of iuMR for fetal brain abnormality have been undertaken among fetuses in which a brain abnormality was suspected 12 (predominantly on the basis of abnormal USS) and, whilst these strongly support the use of 13 iuMR in such cases, the benefit – if any – of iuMR in ostensibly normal pregnancies is 14 unknown. 15 16

In this study we present the results of an extension to the MERIDIAN study in which women
with low-risk pregnancies and normal fetuses on USS were recruited in order to have iuMR
imaging of the fetal brain (full protocol available at

20 https://www.sheffield.ac.uk/meridian/studysummary). This enabled the calculation of

21 negative predictive value (NPV) for both iuMR and USS imaging in order to complement the

22 positive predictive value (PPV) derived from the main MERIDIAN cohort. To our

knowledge, no previous study has evaluated the impact of iuMR in this population. We also

- 24 discuss the problems of measuring diagnostic performance for iuMR and antenatal USS with
- 25 particular reference to the inherent difficulties in estimating sensitivity and specificity.

26 MATERIALS AND METHODS

27 Participants and Recruitment

28 This work was funded by the NIHR-HTA by way of an extension to the main MERIDIAN study (ISRCTN 27626961) and conducted under the same ethics approval.⁴ We aimed to 29 30 recruit approximately 200 women carrying a fetus in whom no brain (or somatic) abnormality was detected on the 20-week anomaly USS that is routinely offered to women in the UK. 31 Any subsequent USS examinations (if performed) also had to show normal fetal anatomy. All 32 33 of the pregnancies were otherwise considered 'low-risk' with no known serological or chromosomal/genetic concerns. Potential participants were informed about the study by way 34 of posters and leaflets in 12 of the original 16 fetal medicine referral centres involved in the 35 original MERIDIAN and by press coverage in those regions. Interested pregnant women 36 contacted the central site (Academic Unit of Radiology, University of Sheffield) and were 37 38 sent a patient information leaflet by email or post, which gave full details of the study. A 39 follow-up telephone call enabled queries to be answered, initial screening questions to be assessed and eligibility for the study confirmed. A copy of the most recent antenatal USS 40 41 report was then obtained to confirm the normal development of the pregnancy. Other inclusion criteria were: the woman was at least 16 years old and the fetus a minimum of 18 42 gestational weeks (gw) at the time of iuMR imaging was to be performed. Exclusion criteria 43 were inability to give informed consent, contraindications to MR imaging, or 44 45 inability/unwillingness to travel to Sheffield for iuMR imaging. There were no set 46 requirements for the interval between considered eligible for the study and having the iuMR 47 scan.

49 Written informed consent was taken on the day of the study after further explanation of the iuMR procedure, including potential risks and the right to withdraw from the study at any 50 time. The consent procedure also confirmed willingness of the woman to inform her GP that 51 52 she had been involved in the study and to send them a copy of the iuMR report if no unexpected findings were shown. If a brain abnormality was detected on iuMR imaging the 53 woman agreed that the findings would be discussed verbally with her obstetrician who would 54 55 subsequently receive a full clinical-style report in accordance with the guidance from the Ethics Committee. Participants were not paid for volunteering for the study but a £10 gift 56 57 voucher was given, along with travel expenses, for the participant and an accompanying person. 58

59 Sample size and reference diagnoses

Starting from the assumption that no USS false negatives will be found, the study aimed to 60 recruit 200 fetuses on the basis of the 3/n rule,⁶ a large sample approximation of the upper 61 95% confidence interval for very rare events. This allowed the negative predictive value of 62 USS to be estimated to an upper confidence limit of 1.5% in the absence of any abnormal 63 64 scans, and to within a standard error of $\leq 2\%$ for an incidence of $\leq 10\%$. The brain of the fetus was assumed to be normal if both USS and iuMR were normal, an 65 approach supported by the low rate of false positive finding for iuMR in the main MERIDAN 66 study (1/570 = 0.18%). These became the True Negatives for USS and iuMR used in this 67 study. Additional tests were undertaken in the event of a brain abnormality reported on iuMR, 68 and these were intended to be the reference against which USS and iuMR were compared, 69 70 although this approach was found to be too simplistic for practical cases as discussed below.

71

72 iuMR scanning procedures and protocols

73 All of the iuMR examinations were performed at the Academic Unit of Radiology, University 74 of Sheffield on either a 1.5T whole body scanner (HDx, GE Healthcare, Milwaukee) or a 3T whole body scanner (Ingenia, Philips, Netherlands). The 3T scanner was used only when the 75 76 1.5T was not available (e.g. breakdowns) and this occurred in two cases only. The iuMR imaging targeted the fetal brain only and the woman was on the scanner for maximum of 30 77 minutes. The imaging protocol performed at 1.5T consisted of ultrafast imaging in the three 78 orthogonal planes (T2 weighted ssFSE and 2D-FIESTA), T1 weighted, FLAIR, Diffusion 79 weighted in the axial plane and T2-weighted volume acquisitions and MR cine using 3D-80 81 FIESTA. After the scan the woman and her companion(s) were shown some of the iuMR images and given the opportunity to take some images on their phone or camera. The formal 82 report on the study was issued the following day after review by a paediatric neuroradiologist 83 84 with extensive experience of fetal neuroimaging (PDG).

85 Statistical methods for assessing diagnostic performance

The accuracy of a negative USS was quantified by the NPV, the percentage of fetuses in whom no abnormality was subsequently detected. For iuMR, NPV agreement was derived separately for fetuses whose initial USS was normal and abnormal USS (i.e. USS+, iuMRand USS-, iuMR-). The PPV of USS and iuMR were derived analogously. PPVs and NPVs were presented alongside 95% binomial confidence intervals. No attempt was made to combine the PPV and NPV of iuMR with those from the main MERIDIAN study, or to estimate the sensitivity and specificity for reasons explained in the discussion.

93

94

96 **RESULTS**

Recruitment and scanning took place between November 2013 and May 2017 during which 97 98 time 225 pregnant women enquired about the study but three women did not meet the inclusion criteria because of pregnancy complications. Appointments for iuMR were made 99 for the other 222 women who did meet the entrance criteria but of those 23 did not attend. 100 One woman underwent the iuMR study but the procedure was abandoned due to the 101 participant feeling unwell before any relevant data was obtained and three women withdrew 102 from the study after iuMR imaging was performed. In total, therefore, 198 participants with 103 205 fetuses (14 twin pregnancies) were scanned successfully as shown in Figure 1. The 104 105 pregnant women recruited were from a wide geographical area, with 68 (34%) participants 106 living within 18 miles of the Sheffield MR unit and the remaining 137 from further afield (maximum 189 miles). The age range of the pregnant women was 20 - 46 years (mean 31.5 107 years) and the gestational age at the time of iuMR is shown in Figure 2 (26% between 18 and 108 109 23 gw, $74\% \ge 24$ gw). There were no reportable adverse events during the iuMR scanning of these pregnant women. IuMR studies were reported as normal for 203 cases and brain 110 111 abnormalities were reported in two fetuses (from separate pregnancies) as described below. 112

113 <u>Case 1 (Figure 3).</u>

iuMR imaging for this study was performed at 35gw following normal USS examinations in
the second trimester (3a-3c). There was focal abnormal high signal on T2-weighted images
in the right inferior/sub-central gyri with broadening of the gyri. The diagnostic confidence
of abnormality was quoted as 70% (certain) and pathology such as a focal cortical dysplasia
or cortical tuber was suggested, although the possibility of an artefact was considered. Postnatal MR imaging performed at 3 weeks (3d-3f) confirmed the antenatal findings but its
nature remained uncertain. Developmental assessment at 6 months showed plagiocephaly

and reduced central tone but otherwise a normal repertoire of movements. The Bayley Infant
Neurodevelopmental Screener (a developmental tool across four domains) put the baby in the
'middle risk' group. Genetic testing for Tuberous Sclerosis Complex was negative. The infant
remains under clinical review and a further MR examination is planned at 3 years.

125

126 <u>Case 2. (figure 4)</u>

Routine anomaly USS was performed at 20gw and showed no abnormalities. iuMR imaging
for this study was undertaken at 26gw and showed mild ventriculomegaly (trigones
measurements between 10-11mm. The rest of the brain was normal although the fetus had
macrocephaly (bi-parietal diameter >97th centile and occipito-frontal on the 97th centile).
Normal sized ventricles were confirmed on review of the USS performed at 20gw but follow
up USS confirmed non-progressive ventriculomegaly at 30gw. The child was developing
normally in all domains at 14 months.

134 Analysis

Case 1 is treated as a true brain abnormality, although the nature of the abnormality is still not 135 known, so is considered to be a False Negative for USS and a True Positive for iuMR. In 136 contrast, the appearance of VM on iuMR imaging in case 2 after retrospective confirmation 137 of normality at 20 weeks is interpreted as an evolving feature that could not be recognised at 138 20 weeks because it wasn't present. However, the confirmation of ventriculomegaly (VM) on 139 third trimester USS confirm the iuMR finding. This is taken as a True Negative for USS and 140 a True Positive for iuMR imaging. Table 1 shows the number and characteristics of correct 141 and incorrect diagnoses made by USS and iuMR using data from both this study and the 142 MERIDIAN study. Both USS and iuMR have high NPV for the normal risk pregnancies, 143 being 99.5% (95% CI 97.3 to 100.0%) for USS and 100% (98.2 to 100%). 144

145	In the main MERIDIAN cohort, 388/570 fetuses were correctly diagnosed by USS giving a
146	PPV of 68.1% (64.1 to 71.9%). Of these, iuMR found abnormalities in 513 fetuses of whom
147	39 were incorrect diagnoses giving a PPV of 92.4% (90.0 to 94.5%). The remaining 57 were
148	recorded as normal on iuMR, one of whom was subsequently found to have a brain
149	abnormality matching the original USS diagnosis, giving an NPV in this population of 98.2%
150	(90.6%-100.0%).
151	
152	
153	

154 **DISCUSSION**

155

improvement in diagnostic accuracy when iuMR imaging is used in the diagnostic pathway.¹⁻ 156 ⁴ One important implication of this finding is USS might fail to detect some brain 157 abnormalities during screening. This study shows that does not occur at high frequency and 158 supports USS being the primary screening method for brain imaging. IuMR should be used as 159 an adjunct to USS only when brain abnormalities are suspected on USS in low risk 160 pregnancies. There were two abnormalities noted on iuMR following a normal USS in 205 161 fetuses, one of which was a case of mild VM that was confidently described as an evolving 162 pathology and the original USS report was correct at the time of scanning. As such, USS has 163 164 a NPV of 99.5% (95% CI 97.3% to 100%), supporting the contention that a normal USS can safely be assumed to rule out fetal brain abnormality with very high certainty in fetuses with 165 no other risk factor. 166

The MERIDIAN study, along with published systematic reviews, demonstrate a significant

167

A review of the literature has not shown any other studies of iuMR imaging in normal 168 pregnancies as identified by USS, so there are no other comparative estimates of NPV and 169 PPV for these modalities. Our study has addressed that knowledge gap by recruiting 205 170 fetuses considered to be developing normally on USS; these were combined with the 171 MERIDIAN results to estimate NPV and PPV. Predictive values indicate the precision of a 172 diagnostic test, i.e. how likely the test is to find an abnormality when it actually exists (PPV) 173 or how likely a test is to be negative if no abnormality exists (NPV) and are arguably more 174 relevant to clinicians when making decisions on the basis of diagnostic tests.^{7,8} Traditionally, 175 sensitivity and specificity have been the preferred measures of diagnostic performance, since 176 the PPV and NPV depend on prevalence,⁹ indeed, the STARD checklist for diagnostic 177 accuracy studies made only cursory mention of predictive values until the 2015 update.^{10,11} In 178

179 this study we have not attempted to estimate the sensitivity and specificity, since our two studies have (deliberately) not recruited random samples of pregnant women. The main 180 MERIDIAN study evaluated iuMR in pregnancies where an abnormality was found on USS, 181 182 with 570 fetuses included in the primary analysis. Since abnormal brain USS occurs in less than 1% of fetuses, a prospective study of all pregnancies would have needed more than 183 57000 participants in order to recruit this number of brain abnormalities. By conducting two 184 parallel studies we were able to study fetuses with normal and abnormal USS, but combining 185 the two into one data set is inappropriate as doing so vastly over-represents by comparison to 186 187 the general population, resulting in a biased estimate of both sensitivity and specificity. Although the sensitivity could - in theory - be derived by re-weighting the two studies to 188 match population incidence, this would entail allocating a weight of less than 1% to the 189 190 original MERIDIAN study with the remainder being allocated to the two cases identified in 191 this study. A similar (though less extreme) situation applies to the specificity, and clearly this results in instable estimates which are best avoided. 192

USS is offered to all women in the UK (and taken by >95%) so sensitivity and specificity of 193 USS may be derived from routine patient notes.¹² The diagnostic capability of USS has 194 previously been quantified using sensitivity and specificity analysis by reviewing clinical 195 196 cases that have been scanned as part of the routine screening process during pregnancy. A report by the National Institute for Health and Care Excellence (NICE)¹³ described the 197 findings from those studies, showing that whilst the sensitivity of USS was variable (15% to 198 199 85%) the specificity was consistently very high (99.4% to 100%). Rossi and Perfumo¹ attempted to define the diagnostic capability of iuMR using similar sensitivity and specificity 200 201 measures but, as the vast majority of fetuses were initially suspected of being abnormal by 202 USS, the truly normal pregnancies were again greatly under-represented and their findings do 203 not adequately generalise to the wider population of pregnancies. Perhaps more importantly,

204 it is questionable whether the diagnostic ability of iuMR imaging needs to be evaluated among all pregnancies. Whilst neonatal screening relies heavily on USS, constraints on 205 resource mean it is likely that iuMR will be used more selectively as a second-line screen for 206 207 high risk pregnancies, most likely a suspected abnormality on USS- a position backed by the data from our study. There are more than 800,000 pregnancies in the UK each year,¹⁴ the 208 majority of which undergo at least one USS, and the resource implications (trained expertise 209 and financial) of providing iuMR routinely is prohibitive. It is interesting to note that the 210 results of the adequately powered study reported by NICE¹³ were comparable to the NPV 211 212 reported here.

213 There are several possible limitations to our study, which primarily stem from recruiting 'normal' participants. Firstly, there may be an element of bias within the recruitment process 214 as it was reliant on volunteers. It is unclear if the women in our sample were fully 215 216 representative of the obstetric population as, although recruited from a wide geographical area within the UK, we did not record demographics such as ethnicity. Secondly, it was not 217 possible to restrict recruitment to women who could attend for iuMR shortly after USS as we 218 were reliant on participants' availability. In theory the longer time period between USS and 219 iuMR, the greater the possibility of abnormalities evolving and hence being visible on MR 220 221 which would therefore biased the findings in favour of iuMR; in reality, the two modalities 222 agreed in all bar two cases. The advantage to not restricting the time between USS and iuMR was that a wider age range of fetuses were scanned, and allowed a greater range of 223 224 gestational age to be assessed since pregnant women are offered an anomaly screening USS between 18 and 21 weeks' gestation in the UK. Thirdly, the diagnostic accuracy of USS for 225 226 this study was based on routine USS screening rather than USS by a fetal-maternal expert, 227 which was a requirement of MERIDIAN. The availability of suitably qualified staff and the cost implications made this unattainable. It is impossible to ascertain whether the 2 cases with 228

abnormalities detected by iuMR were not present at USS or if they were missed. In the fetus
with VM, there was 6 weeks between USS and iuMR, and in the second abnormal case there
was 16 weeks. It was therefore possible that the abnormality was not present at the time of
the USS and even if it was, it is impossible to say whether a fetal-maternal expert could have
identified the abnormality.

234

235 The consequences of abnormalities being missed on ante-natal USS are variable. Detecting abnormalities allows further investigations and additional monitoring of the pregnancy, or, if 236 237 the abnormality is severe and detrimental to long term outcome allows the option of termination of pregnancy. Isolated mild VM is a common finding during pregnancy and a 238 very high proportion have a favourable outcome, but iuMR is necessary to identify additional 239 abnormalities.¹⁵⁻¹⁷ This finding therefore is perhaps less significant than the cortical 240 abnormality diagnosed by iuMR in a fetus of 35 gw. Cortical dysplasia (or cortical tubers) is 241 exceptionally difficult to identify by USS prenatally¹⁸ and can have a range of causes and 242 243 outcomes. Earlier identification of this abnormality may not have changed the outcome in terms of health of the fetus, but would have provided vital information and allowed the 244 parents to make an informed choice regarding its management. 245

246

In conclusion, our results confirm the ability of both USS and iuMR to confirm when brain
development of the fetus is normal. This highlights the validity of USS remaining as the
primary screening imaging method for pregnancy, and further supports the need for
additional iuMR imaging when abnormalities are detected on USS. However further research
on fetuses at an increased risk of brain abnormality may be appropriate.¹⁹

252

253 **REFERENCES**

Jarvis D, Mooney C, Cohen J, et al. A systematic review and meta-analysis to
 determine the contribution of MR imaging to the diagnosis of foetal brain abnormalities In
 Utero. Eur Radiol 2017;27:2367-80.

Rossi AC, Prefumo F. Additional value of fetal magnetic resonance imaging in the
 prenatal diagnosis of central nervous system anomalies: a systematic review of the literature.
 Ultrasound Obstet Gynecol 2014;44:388-93.

3. Van Doorn M, Oude Rengerink K, Newsum EA, et al. Added value of fetal MRI in
fetuses with suspected brain abnormalities on neurosonography: a systematic review and
meta-analysis. J Matern Fetal Neonatal Med 2016;29:2949-61.

4. Griffiths PD, Bradburn M, Campbell MJ, et al. Use of MRI in the diagnosis of fetal
brain abnormalities in utero (MERIDIAN): a multicentre, prospective cohort study. Lancet
2017;389:538-46.

5. Griffiths P, Bradburn M, Campbell M, et al. Change in diagnostic confidence brought
about by using in utero MRI for fetal structural brain pathology: analysis of the MERIDIAN
cohort. Clin Radiol 2017;72:451-7.

269 6. Eypasch E, Lefering R, Kum CK, Troidl H. Probability of adverse events that have
270 not yet occurred: a statistical reminder. BMJ 1995;311:619-20.

7. Naeger DM, Kohi MP, Webb EM, et al. Correctly Using Sensitivity, Specificity, and
Predictive Values in Clinical Practice: How to Avoid Three Common Pitfalls. American
Journal of Roentgenology 2013;200:W566-W570.

274 8. Lange K, Brunner E. Analysis of predictive values based on individual risk factors in
275 multi-modality trials. Diagnostics 2013;3:192-209.

276	9.	Altman DG,	Bland JM.	Diagnostic	tests 2:	Predictive	values. E	3MJ 199	94;309:102.

277 10. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting

of studies of diagnostic accuracy: the STARD Initiative. Radiology 2003;226:24–28

279 11. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: An Updated List of

Essential Items for Reporting Diagnostic Accuracy Studies. BMJ 2015;351:h5527

12. Hewison J, Green JM, Ahmed S, et al. Attitudes to prenatal testing and termination of

pregnancy for fetal abnormality: a comparison of white and Pakistani women in the UK.

283 Prenat Diagn 2007;27:419-30.

13. NICE Clinical Guidelines N. Antenatal Care: Routine Care for the Healthy Pregnant
Woman. In: (UK). NCCfWsaCsH, editor. London: RCOG Press; 2008;1-454.

286 14. Statistics OfN. Conceptions in England and Wales: 2015

287 https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/conceptio
288 nandfertilityrates/bulletins/conceptionstatistics/20152017

289 15. Salomon LJ, Ouahba J, Delezoide AL, et al. Third-trimester fetal MRI in isolated 10-

to 12-mm ventriculomegaly: is it worth it? BJOG : an international journal of obstetrics andgynaecology 2006;113:942-7.

292 16. Griffiths PD, Reeves MJ, Morris JE, et al. A prospective study of fetuses with isolated
293 ventriculomegaly investigated by antenatal sonography and in utero MR imaging. AJNR
294 American journal of neuroradiology 2010;31:106-11.

17. Manganaro L, Savelli S, Francioso A, et al. Role of fetal MRI in the diagnosis of
cerebral ventriculomegaly assessed by ultrasonography. La Radiologia medica 2009;114:
1013-23.

- 298 18. Glenn OA, Cuneo AA, Barkovich AJ, et al. Malformations of Cortical Development:
- 299 Diagnostic Accuracy of Fetal MR Imaging. Radiology 2012;263:843-55.
- 300 19. Griffiths PD, Mooney C, Bradburn M, Jarvis D. Should we perform in utero MRI on
- 301 a fetus at increased risk of a brain abnormality if ultrasonography is normal or shows non-
- 302 specific findings? Clin Radiol 2018; 73(2):123-134

TABLE LEGENDS 303 304 Table 1. Data showing the agreement between ultrasonography (table 1a) and iuMR imaging 305 (table 1b) when compared with outcome reference data. 306 307 Table 1a: 308 Agreement with ORD USS USS Test finding correct incorrect 182 388 USS abnormal* PPV=68.1% (CI 64.1%-71.9%) 204 1 USS normal NPV=99.5% (CI 97.3%-100.0%) 309 310 311 312 Table 1b: Agreement with ORD iuMR iuMR Test finding correct incorrect Following abnormal USS* 474 39 iuMR abnormal PPV=92.4% (CI 90.0%-94.5%) 1 56 NPV=98.2% (CI 90.6%-100.0%) iuMR normal Following normal USS 2 0 iuMR abnormal PPV=100% (CI 15.9%-100%) 203 0 iuMR normal NPV=100% (CI 98.2%-100%) 313 314 * taken from original MERIDIAN cohort of fetuses with brain abnormality on USS 315

316

317	FIGURE LEGENDS
318	Figure 1. Flow of participants through the study.
319	
320	Figure 2. Chart showing the number of fetuses scanned by gestational age.
321	
322	Figure 3. Single shot FSE image (3a), coronal (3b) and sagittal (3c) reconstruction from T2-
323	weighted 3D datasets show broadening of the right inferior frontal gyrus and abnormal white
324	matter signal extending into the sub-central gyrus. These features were confirmed on post-
325	natal imaging (3d-3f). See text for details.
326	
327	Figure 4. Single shot FSE images (4a sagittal, 4b axial) show mild ventriculomegaly and
328	macrocephaly (trigones of the lateral ventricles measured an axial reconstruction from a 3D
329	dataset $-4c$). See text for details.