Accuracy of in-utero MRI to detect fetal brain abnormalities and prognosticate developmental outcome: postnatal follow-up of the MERIDIAN cohort

Anthony R Hart, Nicholas D Embleton, Michael Bradburn, Daniel J A Connolly, Laura Mandefield, Cara Mooney, Paul D Griffiths

Summary
Background In utero MRI (iuMRI) detects fetal brain abnormalities more accurately than ultrasonography and provides additional clinical information in around half of pregnancies. We aimed to study whether postnatal neuroimaging after age 6 months changes the diagnostic accuracy of iuMRI and its ability to predict developmental outcome.

Methods Families enrolled in the MERIDIAN study whose child survived to age 3 years were invited to have a case note review and assessment of developmental outcome with the Bayley Scales of Infant and Toddler Development, the Ages and Stages Questionnaire, or both. A paediatric neuroradiologist, masked to the iuMRI results, reviewed the postnatal neuroimaging if the clinical report differed from iuMRI findings. Diagnostic accuracy was recalculated. A paediatric neurologist and neonatologist categorised participants’ development as normal, at risk, or abnormal, and the ability of iuMRI and ultrasonography to predict developmental outcome were assessed.

Findings 210 participants had case note review, of whom 81 (39%) had additional investigations after age 6 months. The diagnostic accuracy of iuMRI remained higher than ultrasonography (proportion of correct cases was 529 [92%] of 574 vs 387 [67%] of 574; absolute difference 25%, 95% CI 21 to 29; p=0·0001). Developmental outcome data were analysed in 156 participants, and 111 (71%) were categorised as normal or at risk. Of these 111 participants, prognosis was normal or favourable for 56 (51%) using ultrasonography and for 76 (69%) using iuMRI (difference in specificity 18%, 95% CI 7 to 29; p=0·0008). No statistically significant difference was seen in infants with abnormal outcome (difference in sensitivity 4%, 95% CI –10 to 19; p=0·73).

Interpretation iuMRI remains the optimal tool to identify fetal brain abnormalities. It is less accurate when used to predict developmental outcome, although better than ultrasonography for identifying children with normal outcome. Further work is needed to determine how the prognostic abilities of iuMRI can be improved.

Funding National Institute for Health Research Health Technology Assessment programme.

Introduction The MERIDIAN study showed that in utero MRI (iuMRI) detects fetal brain abnormalities more accurately than ultrasonography alone. iuMRI also improves diagnostic confidence and provides additional useful clinical information in 49% of women, which changes prognosis in 24%. Subsequent work in smaller cohorts of women and a meta-analysis have also confirmed the superiority of iuMRI over ultrasonography, although the numbers of additional abnormalities detected by iuMRI were smaller than the MERIDIAN study. The differences in detection are likely to reflect local antenatal care practices, including whether ultrasonography was done abdominally or transvaginally, and the MERIDIAN study reflects the role of iuMRI in UK clinical practice, where transvaginal ultrasonography is not routinely done in all pregnant women.

The outcome reference diagnoses (ORD) used in the MERIDIAN study were either post-mortem results or neuroimaging done in the first 6 months of life, which was predominately cranial ultrasound. It is possible the neuroimaging ORD were incorrect or incomplete, and there is little published data to show how accurately ultrasonography and iuMRI provide prognostic information on developmental outcomes.

Understanding the accuracy of these methods is important because women and their families want to know the significance of any abnormalities found antenatally. The MERIDIAN 3-year follow-up study was designed to address these knowledge gaps across three projects, specifically: project one was designed to recalculate the diagnostic accuracy of ultrasonography and iuMRI using diagnoses collected after 6 months of age; project two, to compare the ability of ultrasonography and iuMRI to predict developmental outcome around age 3 years; and project three, to study the developmental significance of isolated mild ventriculomegaly (IMVM), because of the large numbers of cases and perceived difficulties in antenatal prognostication in this cohort.
Research in context

Evidence before this study
In utero MRI (iuMRI) is increasingly used in antenatal care when ultrasonography detects a fetal brain abnormality. The MERIDIAN study showed that iuMRI has better diagnostic accuracy than ultrasonography, obtaining the correct diagnosis in 93% of cases compared with postnatal imaging before age 6 months or post-mortem findings. Obstetricians reported that the results of iuMRI provided information that improved their diagnostic confidence and directly informed choices about the care of half of the pregnant women whose fetus had a suspected brain abnormality on ultrasonography. There is limited evidence on how iuMRI relates to postnatal imaging after age 6 months, and its ability to predict developmental outcome in infancy. We searched PubMed from inception to Oct 11, 2019, with no language restrictions. Search terms for in utero imaging or fetal brain abnormalities included “Fetus [MESH] AND magnetic resonance imaging [MESH], “Prenatal ultrasonography [MESH], “In-utero MRI [title/abstract], “In-utero Magnetic resonance imaging [title/abstract], “Fet* magnetic resonance imaging [title/abstract], “Fet* imaging [title/abstract], “Prenatal magnetic resonance imaging [title/abstract], “Prenatal MRI [title/abstract], “Prenatal imaging [title/abstract], “Fet* brain abnormality [title/abstract], “Fet* brain disorder [title/abstract]. Search terms for child development included “Child development [MESH], “Developmental disabilities [MESH], “Developmental outcome [MESH], “Developmental impairment [title/abstract], “Developmental delay [title/abstract], “Child development [title/abstract]. These two strategies were then combined using the AND function.

Numerous articles showed the superiority of iuMRI to detect brain abnormalities compared with prenatal ultrasonography, including the MERIDIAN study. No article has directly compared iuMRI with postnatal imaging. As in the original study, obstetricians reported that the information obtained from iuMRI often changes the information they give pregnant women, both ultrasonography and iuMRI are poor predictors of outcome. Clinicians providing antenatal counselling should be aware of this fact when offering prognostication, and avoid overconfidence when discussing the developmental significance of brain abnormalities. This is true for significant structural abnormalities and of more subtle changes, such as isolated mild ventriculomegaly, that are usually associated with good outcome. It remains to be seen to whether the accuracy of prognostication improves when paediatricians with experience following up children with neurological and developmental abnormalities are involved in multidisciplinary team discussions.

Added value of all the available evidence
iuMRI detects brain abnormalities with high diagnostic accuracy, confirming its role in antenatal care, and there appears to be little additional value in routine postnatal imaging, given the chance of changing the diagnosis is small. Although obstetricians report that the information obtained from iuMRI often changes the information they give pregnant women, both ultrasonography and iuMRI are poor predictors of outcome. Clinicians providing antenatal counselling should be aware of this fact when offering prognostication, and avoid overconfidence when discussing the developmental significance of brain abnormalities. This is true for significant structural abnormalities and of more subtle changes, such as isolated mild ventriculomegaly, that are usually associated with good outcome. It remains to be seen to whether the accuracy of prognostication improves when paediatricians with experience following up children with neurological and developmental abnormalities are involved in multidisciplinary team discussions about the significance of ultrasonography and iuMRI results.

Methods

Study design and participants
The MERIDIAN 3-year follow-up study was an extension of the original MERIDIAN study, in which all families were asked to provide written consent to be contacted about future studies. Ethical approval was obtained and it was overseen by a trial steering committee. Section 251 confidentiality advisory group approval was obtained for use of the Health & Social Care Information Centre Patient Tracking system to ensure potentially eligible children were alive before approach. If the child had died, the date and cause of death were obtained but no contact was made with the family. Children were excluded if they were no longer in the care of the biological mother or if the mother was unable to understand English and no family member could translate. A letter of invitation, consent form, and a developmental screening questionnaire (the Ages and Stages Questionnaire 3rd edition; ASQ) were sent to the families of eligible children. Families could consent for either project one or all projects. The full protocol is available online.

Procedures
Families who consented to their child’s participation in project one agreed to a review of their child’s case notes and neuroimaging reports. If the latest report did not match the original ORD, or if an ORD study had not been done within the first 6 months of life, a paediatric neuroradiologist (DJAC) reviewed the neuroimaging report. As in the original study, cases in which the iuMRI was done more than 2 weeks after ultrasonography were excluded. Families who consented to project one
could also return the ASQ3, even if they did not consent to project two. Families who consented to project two were invited to a developmental assessment using the Bayley Scales of Infant and Toddler Development 3rd edition (BSID3), a self-report questionnaire of the Gross Motor Function Classification System (GMFCS), and the Strengths and Difficulties questionnaire (SDQ). If families were not willing to attend, or if the child was older than age 42 months, they were asked to complete the questionnaires only. The results of project two were combined with ASQ3 results.

Data on developmental outcome from the medical records, questionnaires, and the BSID3 assessments were reviewed independently by a paediatric neurologist (ARH) and a neonatologist (NDE), both of whom were masked to the ultrasonography iuMRI results, and ORD. Participants were categorised as normal, at risk, or abnormal (table 1). The results of the BSID3 took precedent over the ASQ3. Disagreement on developmental outcome was resolved by consensus discussion. For borderline cases, the results of the SDQ were taken into account, with a score of more than 16 taken to be abnormal (SDQ scale 0–40).

In the MERIDIAN study, obstetricians independently provided a prediction of the fetus’ developmental outcome on the basis of the ultrasonography and iuMRI results as normal (chance of abnormal development no greater than the general population), favourable (<10% chance of abnormal development), intermediate (10–50% chance of abnormal development), poor (>50% chance of abnormal development), or unknown.

In project three, the developmental outcomes of children with antenatal diagnosis of IMVM were reviewed. We included all participants in this analysis, irrespective of the time between imaging, because we were not comparing ultrasonography to iuMRI.

**Statistical analysis**

The original MERIDIAN sample size calculation was applicable for project one because the 6-month diagnosis was retained unless additional diagnostic information was obtained. For projects two and three, the sample size was constrained by the number of MERIDIAN participants whose families consented to further study. The sample size for project two was informed by the number of fetuses whose prognosis changed as a result of the iuMRI (n=312), of whom a further 63 were reclassified as the poorest prognosis, and assumed approximately 400 surviving infants. Scaling these prevalences down, 400 infants would provide 90% power to detect a 20% increase in sensitivity and a 10% increase in specificity at a two-sided significance level of 5%. For project three, with approximately 140 cases and assuming poor outcome was less than 10%, allowing estimation of the prevalence to within a standard error of 2-5%.11

For project one, the updated imaging was used to recalculate diagnostic accuracy. Developmental outcome in relation to ultrasonography and iuMRI was summarised by sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV; table 1). Specificities and sensitivities of ultrasonography and iuMRI were compared using the McNemar paired sample χ² test, with the difference and 95% CIs calculated using the Wilson score method. We studied the ability of ASQ3 to diagnose abnormal development on BSID3, to assess whether it was appropriate to include participants with only ASQ3 results in our analysis. Sensitivities and specificities were calculated. All statistical analyses were done using R, version 3.5.2, and Stata 15.

**Role of the funding source**

The funders had no role in study design, data collection, data analysis, data interpretation, or writing of the Article. The final author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

**Results**

The flow of the 829 participants from the original MERIDIAN study into each of the three projects is shown in figure 1. The characteristics of the women involved in the study are summarised in the appendix (p 1). The families of 238 children consented to
project one. Case notes were available for review in 210 of 238 children, of whom 81 (39%) had further investigations between age zero and 42 months (median 4 months), most commonly by postnatal MRI (n=38) or clinical assessment (n=34). Data on the reasons for the further investigations were not collected.

Figure 2: Study profile
ASQ=Age and Stages Questionnaire. ORD=outcome reference diagnoses.
BSID3=Bayley Scales of Infant and Toddler Development 3rd Edition. ASQ3=ASQ 3rd Edition. GMFCS=Gross Motor Function Classification Score. SDQ=Strengths and Difficulties questionnaire.*Did not complete assessment: participants who consented, but then either did not attend or cancelled; assessment incomplete: participants who attended, but the whole assessment could not be completed or a developmental score obtained. †52 completed ASQ as part of project 1 and were included in project 2. ‡11 had developmental outcome available in medical notes and were included in project 2.

200 MERIDIAN cohort with completed imaging

615 screened

576 eligible

554 sent invitation letter

317 agreed to initial consent

238 consented to MERIDIAN follow on
106 project 1 only
132 project 1 and 2

214 could not be approached
127 died in original MERIDIAN study
25 died since MERIDIAN study
2 withdrew from original MERIDIAN study
2 not screened, reason unknown

39 ineligible at screening
12 did not want to be contacted
3 unable to confirm if child alive
12 child no longer in care of mother
12 inappropriate to contact

22 invitation letter could not be sent

196 did not respond or no contact
41 declined entry
6 no time
2 not interested
18 did not want to be contacted
55 other

79 did not consent
2 not enough time
53 consent form not received
10 no reason given
13 not contacted
1 other

210 notes available for review

129 no updated neuroimaging

81 neuroimaging after age of 6 months

71 did not have information changing ORD

10 information changing ORD

65 stillbirth, in-utero demise or infant death, had ultrasonography and iuMRI within 2 weeks of each other and clinical data.
21 deaths within MERIDIAN
2 unlikely to be related to fetal brain abnormality
10 brain sufficiently abnormal that might have contributed to death
3 syndromic type diagnosis likely to be associated with short lifespan where brain abnormality was a marker of a more widespread problem
6 reason for death could not be coded

238 consented to project 1

28 notes not available for review

219 notes available for review

129 no updated neuroimaging

81 neuroimaging after age of 6 months

71 did not have information changing ORD

10 information changing ORD

132 consented to project 2

5 did not complete assessment or assessment incomplete*

179 completed at least one form of developmental assessment
88 BSID3
175 ASQ3
31 GMFCS
102 SDQ

190 data reviewed by developmental outcome panel‡

34 excluded from analysis
26 did not have ultrasonography and iuMRI within 2 weeks of each other
8 no prognosis offered by obstetrician after iuMRI

221 developmental outcome or mortality data available for analysis of project 2
156 developmental assessment
65 mortality data
development in childhood after being given a poor prognosis after ultrasonography or iuMRI (table 4).

The results for all children, irrespective of the time between ultrasonography and iuMRI, were similar (appendix p 4). The proportion of children in each prognostic category who had either abnormal or combined abnormal and at risk outcomes are shown in table 5. Normal or favourable prognostic groups from ultrasonography and iuMRI data were associated with abnormal outcomes in approximately 20% of cases (ultrasonography 18 of 85 cases; iuMR 21 of 109 cases), and either abnormal or atypical outcome in around 35% of cases (ultrasonography 30 of 85 cases; iuMR 38 of 109 cases). A normal prognosis was associated with normal outcome in around 65% of cases (ultrasonography 55 of 85; iuMR 71 of 109).

19 (22%) of 86 participants diagnosed with IMVM had abnormal development and 12 (14%) of 86 participants were considered at risk. None of the children received a poor prognosis category on antenatal imaging with either modality. Abnormal outcome was less likely in cases of resolved IMVM than persisting IMVM: IMVM resolved in 19 (22%) of 86 participants diagnosed with IMVM; IMVM resolved in 62 (72%) of 86 participants, nine (15%) of whom had abnormal outcome, and was persistent in 24 (28%) of 86 participants, ten (42%) of whom had abnormal outcome. 15 (17%) of 86 children had postnatally diagnosed structural or genetic abnormalities or significant life events, and these were proportionately more common in children with abnormal (seven [37%] of 19) and at-risk (four [33%] of 12) development compared with normal outcome (four [7%] of 54; table 1). 84 participants had both ASQ3 and BSID3

| Ultrasonography diagnosis | iuMRI diagnosis | Outcome reference diagnosis at age 6 months | Outcome reference diagnosis at age 3 years |
|---------------------------|----------------|------------------------------------------|----------------------------------------|
| 1 Ventriloculegaly, hypoplastic cerebellar hemispheres | Ventriloculegaly | Ventriloculegaly | Rhombencephaloapsis |
| 2 Ventriloculegaly | Ventriloculegaly | Ventriloculegaly | Ventriloculegaly, subependymal heterotopia |
| 3 Ventriloculegaly, hypogenesis of corpus callosum | Ventriloculegaly | Ventriloculegaly (due to hydrocephalus) | Tectal plate tumour, hydrocephalus due to aqueduct stenosis, subependymal heterotopia |
| 4 Hypoplastic cerebellar hemispheres | Hypoplastic cerebellar hemispheres | Hypoplastic cerebellar hemispheres | Hypoplastic cerebellar hemispheres, subependymal heterotopia |
| 5 Agenesis of corpus callosum | Hypogenesis of corpus callosum | Hypogenesis of corpus callosum | Hypogenesis of corpus callosum, cyst of the quadrigeminal cistern, polymicrogyria |
| 6 Ventriloculegaly, hypogenesis of corpus callosum | Ventriloculegaly, hypogenesis of corpus callosum | Not done | Ventriloculegaly |
| 7 Agenesis of corpus callosum | Agenesis of corpus callosum polymicrogyry | Not done | Agenesis of corpus callosum, polymicrogyria |
| 8 Absent cerebellum | Ventriloculegaly, Chiari II malformation | Not done | Ventriloculegaly, Chiari II malformation |
| 9 Ventriloculegaly, agenesis of corpus callosum | Ventriloculegaly | Not done | Ventriloculegaly |
| 10 Hypoplastic cerebellar vermis | Posterior fossa cyst | Not done | Dandy-Walker spectrum abnormality |

Table 2: The ten cases with updated outcome reference diagnosis

The ORD was unchanged in 71 (88%) of 81 cases. In five (6%) of 81 cases, the more recent ORD disagreed with the original ORD, and in a further five cases (6%) the post-6 month neuroimaging studies provided the only ORD (table 2). One participant had iuMRI more than 2 weeks after ultrasonography and was excluded from the recaluation. The recalculated diagnostic accuracies for ultrasonography and iuMRI, based on the refreshed ORD from 574 cases in total, were similar to those in the MERIDIAN study and the statistically significant difference in diagnostic accuracy was maintained (absolute difference 25%, 95% CI 21–29; p<0·0001; table 3).

84 participants had both ASQ3 and BSID3 development compared with normal outcome (four [7%] of 19) and at-risk (four [33%] of 12) development compared with normal outcome (four [7%] of 54; table 1)
Discussion

Clinical outcomes beyond the immediate postnatal period have not been reported in studies comparing ultrasonography and iuMRI, which is important because features like the immature myelination state of the brain can hamper the diagnostic capacity of MRI before age 6 months. Our results show that postnatal imaging after age 6 months rarely changes ORD and the improved diagnostic accuracy of iuMRI is maintained.

The developmental significance of antenatally diagnosed brain abnormalities is important for parents, who use the information to decide whether they have the ability, and emotional and financial resources to care for their child.6–8,14 Data on many fetal brain abnormalities that can be used during antenatal counselling are scarce,15,16 and whether iuMRI improves prognostication is unclear. The interpretation of our findings is not straightforward.

Among the 111 infants with normal or at risk development, more participants had good prognosis with iuMRI compared with ultrasonography, suggesting iuMRI is better at excluding abnormal developmental outcome than ultrasonography, either through better visualisation of the abnormalities or identification of false positives on ultrasonography. The percentage of infants with normal or at risk developmental outcome after a good iuMRI prognosis was similar to a good prognosis on ultrasonography. The discrepancy between specificities and NPVs is likely to be secondary for two reasons. First, iuMRI was associated with fewer unknown or intermediate prognoses than ultrasonography, often moving the fetus into normal, favourable, or poor prognosis groups. Consequently, although the NPVs were similar, iuMRI had improved prognostic accuracy compared with ultrasonography. This observation might be because iuMRI delineates or excludes other abnormalities better than ultrasonography, but there might also be a psychological factor, whereby iuMRI gives the clinician more confidence to move from an indeterminate to either

Data are n/N (%), unless otherwise indicated. The 574 fetuses include the 570 fetuses included in the original MERIDIAN primary analysis (had iuMRI within 2 weeks of ultrasonography and had outcome reference diagnosis). There were four further cases that now have an outcome reference diagnosis from the 3-year data, increasing the number to 574. Four of the original cases also had updated outcome reference diagnosis from the 3-year data. The updated diagnostic accuracy is based on new neuroimaging studies up to age 3 years and presented as a percentage along with the percentages correct from the original MERIDIAN study for reference.

Table 3: Updated diagnostic accuracy of ultrasonography and iuMRI by gestational age of fetus

| Meridian | Meridian 3-year follow-up |
|----------|---------------------------|
|          | Ultrasonography correct | iuMRI correct | Ultrasonography correct | iuMRI correct | Percentage difference (95% CI) | p value* |
| 18–23 weeks (n=372) | 256/369 (70%) | 341/369 (92%) | 257/372 (69%) | 339/372 (91%) | 22% (17–27) | <0.0001 |
| ≥24 weeks (n=202) | 129/201 (64%) | 188/201 (94%) | 130/202 (64%) | 190/202 (94%) | 30% (23–37) | <0.0001 |
| Combined (n=574) | 387/570 (68%) | 529/570 (93%) | 387/574 (67%) | 529/574 (92%) | 25% (21–29) | <0.0001 |

iuMRI=in-utero MRI. *McNemar’s test between ultrasound and iuMRI correct diagnoses.

Figure 2: Developmental outcome for cognition, language, and motor skills of iuMRI and ultrasonography by prognosis.

The box represents the first (bottom) and third (top) quartiles, the line through the centre of the box represents the median. The lower whisker extends from the smallest observation greater than or equal to the first quartile minus 1.5 × IQR and the upper whisker extends to the largest observation less than or equal to the third quartile plus 1.5 × IQR. iuMRI=in-utero MRI.
a good or poor prognosis.

When the final prognosis based on iuMRI was normal or favourable, this was often correct. The second reason relates to the effect of non-uptake of the 3-year developmental assessment. Although the reasons for non-participation are speculative, fetal and infant mortality (excluding termination of pregnancy) were marginally lower after a favourable or normal prognosis from iuMRI than ultrasonography, and higher following a poor iuMRI prognosis than a poor ultrasonography prognosis. Similar patterns were seen for frequency of pregnancy terminations. Therefore, although iuMRI was superior to ultrasonography for identifying children with normal development, the two modalities were similar in their ability to predict abnormal development.

A further consideration relates to how we defined developmental outcomes. The mean population score of the BSID3 is 100 (SD 15), so around 95% of the normal population lie between 70 to 130. These scores are in keeping with the definition of early developmental impairment—in which developmental abilities have to fall two standard deviations below the population mean—in which developmental abilities have to fall two standard deviations below the population mean in two or more domains—and the International Statistical Classification of Diseases, 10th revision, classification of mild learning disabilities when an intelligence quotient (IQ) is 50–69, moderate when 35–49, severe when 20–34, and profound when below 20. By contrast, preterm follow-up studies, such as the EPICure Study, define BSID3 scores between 70 and 85 as mild impairment, 55 and 70 as moderate, and below 55 as severe, or cognitive or language scores below 85 as moderate to severe impairment.

Using these definitions for fetal follow-up studies comes with challenges; for example, a single point could be the difference between a normal and moderate to severe impairment and might result from the child’s tiredness, boredom, or refusal. The definitions take little account of the child’s genetic or social potential: a child with a cognitive composite score of 84 whose parents’ IQs are also in the 80s might be appropriate, whereas a child with a score in the normal range at 88 whose parents’ IQs are 130–140 might not.

Developmental assessments at 2–3 years are imperfect predictors of later outcome, because children scoring normally in the early years might exhibit significant difficulties later, and children with low scores might catch up to their peers. Therefore, care is needed, especially because terminology like severe or abnormal strongly influence patient choice, and we do not wish to inflate the risk of developmental problems and inadvertently coerce women towards terminating fetuses who would be normal.

Another issue was the proportion of participants assessed only with the ASQ3, which can yield false-positive results. A number of children who were identified by the ASQ3 as borderline or abnormal, scored normally on the BSID3 (appendix p 5). For these reasons, we chose the term at risk rather than mildly abnormal or borderline because these children might have normal outcome in the long term. We were similarly careful not to classify all participants with cerebral palsy as abnormal. This might seem a controversial choice, but women report that health-care professionals focus too much on medical diagnoses and terminology during antenatal counselling, when what they really want to know is the best and worst case scenario and the likely functional outcome of their baby.

The views are consistent with a WHO report on disability, which notes that generalisations about disability are misleading because many people with a disability do not consider themselves unhealthy or to have a bad quality of life. WHO suggests clinicians’ focus should be directed less towards medical diagnosis and more towards how it impairs an individual, affecting activity, participation, and independence. With this in mind, categorising a child with mild cerebral palsy—whose GMFCS scores are low and whose motor scores on BSID3 and ASQ3 are normal—as severely abnormal is inappropriate, particularly given they are likely to have normal quality of life. Therefore, we categorised functionally able children with cerebral palsy as at risk.

Although the number of participants in project two was lower than anticipated, our findings indicate that, despite the increase in diagnostic accuracy and confi-
dence provided by iuMRI, it remains difficult to predict clinical outcomes on an individual basis. This result is unsurprising because prognostication data from antenatal studies are based on population studies rather than individualised risks for specific fetuses, and the effect of a given brain abnormality can vary markedly between individual infants.

A more appropriate way to analyse project two might be on a grouped basis: the data shown in table 5 indicate that if a poor prognosis is made on the basis of ultrasonography data alone, 69% of those fetuses have an abnormal outcome; the results using iuMRI information are similar (72%). These figures are consistent with the more than 50% chance used in our definition of poor prognosis. An intermediate prognosis based on ultrasonography data is associated with the chance that 29% have an abnormal outcome, whereas an intermediate prognosis on iuMRI is associated with a higher chance of an abnormal outcome (42%). Both fall into the 11–50% range used to define intermediate prognosis.

 Fetuses given an unknown prognostic category had clinical outcomes similar to the intermediate category, with abnormal outcomes in 38% from ultrasonography information and 31% based on iuMRI information. The results from the normal and favourable categories showed the widest discrepancy when compared with the predicted abnormal risk rates.

The predicted risk of an abnormal outcome when those two groups are merged is 10% or less but the observed abnormal outcome rates were 21% for ultrasonography and 19% for iuMRI-derived data. This finding was unexpected and concerning because they are double the expected percentage of abnormal outcomes in cases which, presumably, had normal or minor findings on ultrasonography and iuMRI. The challenges of delivering exome studies in a timely manner in antenatal care also need to be overcome, along with explaining to families the problem of phenotypical variation between individuals with the same mutation, even within a single family. Further research is needed to determine whether a paediatrician with experience in following up children with neurodevelopmental abnormalities might interpret antenatal imaging findings differently and provide better prognostication.

In many ways, the limitations of ultrasonography and iuMRI to prognosticate should not be surprising: most children with early developmental impairment have normal neuroimaging, and it is common for children with abnormalities on neuroimaging to do better developmentally than expected.

In the future, gene exome work might help with prognosis, but it is unlikely to give clinicians all the information they want. For example, the PAGE study of 610 fetuses with structural abnormalities found a diagnostic genetic result in only 8–5% and a result of uncertain clinical significance in 3–9%. The challenges of delivering exome studies in a timely manner in antenatal care also need to be overcome, along with explaining to families the problem of phenotypical variation between individuals with the same mutation, even within a single family. Project three looked at children who had IMVM. Most children with this antenatal diagnosis do not show adverse developmental outcomes, so fetal maternal clinicians usually give a favourable prognostic category. Specifically, just over half of the fetal maternal clinicians referring into the MERIDIAN study give a 90% chance of a normal outcome and the others 95% normal outcome (ie, all have >90% chance of normal outcome). Large studies and meta-analyses suggest abnormal developmental outcome rates in IMVM are 5–6–12%. Although studies differ in how IMVM is defined, with

| Prognosis based on ultrasonography | Prognosis based on iuMRI* |
|------------------------------------|--------------------------|
| Proportion with abnormal or at-risk development postnatally | Proportion with abnormal or at-risk development postnatally |
| Normal or favourable (>10%) | 18/85 (21%) | 21/109 (19%) |
| Intermediate (11–50%) | 14/49 (29%) | 15/36 (42%) |
| Poor (>50%) | 11/16 (69%) | 13/18 (72%) |
| Unknown (not known) | 15/40 (38%) | 5/15 (33%) |

iuMRI=in-utero MRI. *The decreased number of iuMRI relates to 11 who had no prognostication information offered by the obstetrician.

Table 5: Prognosis estimated by fetal medicine unit based on ultrasonography and iuMRI for all participants, irrespective of the time between ultrasonography and iuMRI.
some studies using 10–11 mm and others 10–15 mm. Furthermore, some studies only consider a fetus as having IMVM when the karyotype and congenital infection screen is negative, others rely purely on the absence of other structural abnormalities on imaging.

We took a pragmatic approach because women might refuse karyotyping if they considered the risk of miscarriage related to amniocentesis or chorionic villus sampling too high. We also reviewed the medical notes and diagnoses of children for this study, and would have recognised children with genetic conditions diagnosable by karyotype antenatally and would have excluded these, but no such cases were identified.

The results of our study show the risk of an abnormal outcome was much higher at 22%, and a further 14% were considered at risk. The high prevalence of mild ventriculomegaly cases in MERIDIAN might partially explain the high proportion of abnormal outcomes in fetuses given normal or favourable prognostic categories, with selection bias also potentially playing a role because parents with concerns about their child’s development might have been more likely to enrol.

The false-positive rate of the ASQ3 might also play a role. We found a higher rate of abnormal outcome when mild ventriculomegaly was associated with other abnormalities, such as somatic malformations or chromosomal abnormalities, and the risk of adverse developmental outcomes was lower when ventriculomegaly resolved in utero, which confirms previous data.45

The key limitation in the MERIDIAN 3-year follow-up was the low level of participation, with just under a third of eligible parents giving consent to developmental assessment. The effect of this is difficult to predict, but is likely to be non-random: previous work suggests that some parents choose not to participate in research because the burden outweighs the benefit,46,47 which in this study might arise either for children with no obvious problems or for those with clearly abnormal development who usually require multiple other hospital appointments. This reduces the statistical power and precision, but more importantly might have led to either an over-estimation or an under-estimation of the incidence of abnormal development.

We cannot be more specific without knowing the reasons parents chose not to participate. By contrast, any selection effect is less likely to affect the direct comparison between iuMRI and ultrasonography, particularly the sensitivity and specificity for predicting abnormal development.48 We also were not able to administer BSID3 to all participants, relying on the ASQ3 in a proportion. Direct comparison between the two tests shows good agreement, but false positives on the ASQ3 might have inflated the abnormal and at-risk groups.

Finally, we limited the analysis of development to surviving infants. An alternative analysis would be to incorporate both surviving and non-surviving cases using survey-based weighting to account for their differential probability for inclusion. The difficulties in interpreting these measures, in particular for cases managed by termination of pregnancy, led us to focus on outcomes conditional on the infant surviving to age 3 years. These exclusions will have underestimated the true diagnostic and prognostic ability of iuMRI, given its ability to identify severe brain conditions.

Despite a low level of participation, we confirmed that the original postnatal diagnosis remained applicable for the majority of infants at their 3-year follow-up. We did not collect information on why additional investigations were (or were not) undertaken, and it is likely that a proportion of infants had more latent complications that had either not been diagnosed or not become apparent. Despite our concerns about the veracity of ORD from neuroimaging studies before 6 months of life, the updated estimates of diagnostic accuracy of ultrasonography and iuMRI for detecting fetal brain abnormalities are exceptionally close to the original calculations and confirm the diagnostic advantage of doing iuMRI.

The findings from the clinical follow-up show iuMRI improves the confidence of the fetal medicine unit team to say a fetus will have a normal outcome, but prognostication of outcome remains similar between imaging modalities. Further work is needed to determine whether this is a problem inherent with iuMRI or whether other professionals, such as paediatricians, can help to improve prognostication.

Finally, our results indicate a much higher rate of abnormal outcomes than would be expected in fetuses with minor brain abnormalities who were estimated to have normal or favourable prognosis antenatally. This is particularly the case in fetuses with mild ventriculomegaly that does not resolve in pregnancy, and larger cohort studies are needed to determine what the exact risk is and which fetuses are at the greatest risk.

Contributors
ARH and NDE assisted in design of the methods, specifically how developmental outcome should be categorised; reviewed the developmental outcome data of participants; assisted data analysis and interpretation of results; and contributed to writing of the Article. MB and LM were study statisticians, who assisted with the study design and provided advice and input to all statistical issues; completed data analysis and data interpretation, and revised the Article. DJAC reviewed postnatal neuroimaging of participants; reviewed and commented on the final version of the Article. CM was the study manager, who coordinated study activities; provided training for research sites, assisted with and facilitated data collection, participated in data analysis, and revised the Article. PDG oversaw all study conduct: helped to develop the study design and all study materials including the study protocol; assessed participant eligibility and was the reporting radiologist for iuMRI scans completed at the Sheffield site; participated in data analysis and interpretation of the results; and drafted and revised the Article.

Declaration of interests
All authors declare no competing interests.

Acknowledgments
This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment Programme (project number 09/06/01). The views and opinions expressed herein are those of the
for fetuses with ultrasound abnormalities and a suspected Mendelian disorder. *Genome Med* 2018; 10: 74.

43 Pagani G, Thilaganathan B, Prefumo F. Neurodevelopmental outcome in isolated mild fetal ventriculomegaly: systematic review and meta-analysis. *Ultrasound Obstet Gynecol* 2014; 44: 254–60.

44 Devaseelan P, Cardwell C, Bell B, Ong S. Prognosis of isolated mild to moderate fetal cerebral ventriculomegaly: a systematic review. *J Perinat Med* 2010; 38: 401–09.

45 Thorup E, Jensen LN, Bak GS, et al. Children prenatally believed to have isolated mild ventriculomegaly. *Ultrasound Obstet Gynecol* 2019; 54: 182–89.

46 Fisher HR, McKeitt C, Bozz A. Why do parents enroll their children in research: a narrative synthesis. *J Med Ethics* 2011; 37: 544–51.

47 Sammons HM, Atkinson M, I C, Stephenson T. What motivates British parents to consent for research? A questionnaire study. *BMC Pediatr* 2007; 7: 12.

48 Straus SE, Richardson WS, Glasziou P, Haynes RB. Diagnosis and screening. Evidence-based medicine: how to practice and teach, 3rd edn. Oxford, UK: Churchill Livingstone, 2005: 89–90.