

Health Technology Assessment

Volume 23 • Issue 49 • September 2019

ISSN 1366-5278

MRI in the diagnosis of fetal developmental brain abnormalities: the MERIDIAN diagnostic accuracy study

*Paul D Griffiths, Michael Bradburn, Michael J Campbell, Cindy L Cooper,
Nicholas Embleton, Ruth Graham, Anthony R Hart, Deborah Jarvis, Mark D Kilby,
Mabel Lie, Gerald Mason, Laura Mandefield, Cara Mooney, Rebekah Pennington,
Stephen C Robson and Allan Wailoo on behalf of the MERIDIAN Collaborative Group*



MRI in the diagnosis of fetal developmental brain abnormalities: the MERIDIAN diagnostic accuracy study

Paul D Griffiths,^{1*} Michael Bradburn,²
Michael J Campbell,² Cindy L Cooper,²
Nicholas Embleton,³ Ruth Graham,⁴ Anthony R Hart,⁵
Deborah Jarvis,¹ Mark D Kilby,^{6,7} Mabel Lie,⁸
Gerald Mason,⁹ Laura Mandefield,² Cara Mooney,²
Rebekah Pennington,¹⁰ Stephen C Robson⁸
and Allan Wailoo¹⁰ on behalf of the MERIDIAN
Collaborative Group

¹Academic Unit of Radiology, University of Sheffield, Sheffield, UK

²Clinical Trials Research Unit, School of Health and Related Research, University of Sheffield, Sheffield, UK

³Newcastle Neonatal Service, Royal Victoria Infirmary, Newcastle upon Tyne, UK

⁴School of Geography, Politics and Sociology, Newcastle University, Newcastle upon Tyne, UK

⁵Department of Perinatal and Paediatric Neurology, Sheffield Children's Hospital NHS Foundation Trust, Sheffield, UK

⁶Centre for Women's and Newborn Health, Institute of Metabolism and Systems Research, University of Birmingham, Birmingham, UK

⁷Fetal Medicine Centre, Birmingham Women's and Children's NHS Foundation Trust (Birmingham Health Partners), Birmingham, UK

⁸Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne, UK

⁹Leeds Teaching Hospitals NHS Trust, Leeds, UK

¹⁰Health Economics and Decision Science, School of Health and Related Research, University of Sheffield, Sheffield, UK

*Corresponding author

Declared competing interests of authors: Paul D Griffiths was a member of Health Technology Assessment Commissioning Board (2014–18) and Cindy L Cooper is a member of National Institute for Health Research Clinical Trials Unit Standing Advisory Committee (2015 to present).

Published September 2019

DOI: 10.3310/hta23490

This report should be referenced as follows:

Griffiths PD, Bradburn M, Campbell MJ, Cooper CL, Embleton N, Graham R, *et al.* MRI in the diagnosis of fetal developmental brain abnormalities: the MERIDIAN diagnostic accuracy study. *Health Technol Assess* 2019;**23**(49).

Health Technology Assessment is indexed and abstracted in *Index Medicus/MEDLINE*, *Excerpta Medica/EMBASE*, *Science Citation Index Expanded (SciSearch®)* and *Current Contents®/Clinical Medicine*.

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 3.819

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the Clarivate Analytics Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: <http://www.nets.nihr.ac.uk/programmes/hta>

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 09/06/01. The contractual start date was in March 2011. The draft report began editorial review in March 2018 and was accepted for publication in November 2018. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Griffiths *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

NIHR Journals Library Editor-in-Chief

Professor Ken Stein Professor of Public Health, University of Exeter Medical School, UK

NIHR Journals Library Editors

Professor John Powell Chair of HTA and EME Editorial Board and Editor-in-Chief of HTA and EME journals. Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK, and Honorary Professor, University of Manchester, and Senior Clinical Researcher and Associate Professor, Nuffield Department of Primary Care Health Sciences, University of Oxford, UK

Professor Andrée Le May Chair of NIHR Journals Library Editorial Group (HS&DR, PGfAR, PHR journals) and Editor-in-Chief of HS&DR, PGfAR, PHR journals

Professor Matthias Beck Professor of Management, Cork University Business School, Department of Management and Marketing, University College Cork, Ireland

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Eugenia Cronin Senior Scientific Advisor, Wessex Institute, UK

Dr Peter Davidson Consultant Advisor, Wessex Institute, University of Southampton, UK

Ms Tara Lamont Director, NIHR Dissemination Centre, UK

Dr Catriona McDaid Senior Research Fellow, York Trials Unit, Department of Health Sciences, University of York, UK

Professor William McGuire Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads Professor of Wellbeing Research, University of Winchester, UK

Professor John Norrie Chair in Medical Statistics, University of Edinburgh, UK

Professor James Raftery Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

Dr Rob Riemsma Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

Professor Helen Roberts Professor of Child Health Research, UCL Great Ormond Street Institute of Child Health, UK

Professor Jonathan Ross Professor of Sexual Health and HIV, University Hospital Birmingham, UK

Professor Helen Snooks Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Professor Ken Stein Professor of Public Health, University of Exeter Medical School, UK

Professor Jim Thornton Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

Professor Martin Underwood Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, UK

Please visit the website for a list of editors: www.journalslibrary.nihr.ac.uk/about/editors

Editorial contact: journals.library@nihr.ac.uk

Abstract

MRI in the diagnosis of fetal developmental brain abnormalities: the MERIDIAN diagnostic accuracy study

Paul D Griffiths,^{1*} Michael Bradburn,² Michael J Campbell,² Cindy L Cooper,² Nicholas Embleton,³ Ruth Graham,⁴ Anthony R Hart,⁵ Deborah Jarvis,¹ Mark D Kilby,^{6,7} Mabel Lie,⁸ Gerald Mason,⁹ Laura Mandefield,² Cara Mooney,² Rebekah Pennington,¹⁰ Stephen C Robson⁸ and Allan Wailoo¹⁰ on behalf of the MERIDIAN Collaborative Group

¹Academic Unit of Radiology, University of Sheffield, Sheffield, UK

²Clinical Trials Research Unit, School of Health and Related Research, University of Sheffield, Sheffield, UK

³Newcastle Neonatal Service, Royal Victoria Infirmary, Newcastle upon Tyne, UK

⁴School of Geography, Politics and Sociology, Newcastle University, Newcastle upon Tyne, UK

⁵Department of Perinatal and Paediatric Neurology, Sheffield Children's Hospital NHS Foundation Trust, Sheffield, UK

⁶Centre for Women's and Newborn Health, Institute of Metabolism and Systems Research, University of Birmingham, Birmingham, UK

⁷Fetal Medicine Centre, Birmingham Women's and Children's NHS Foundation Trust (Birmingham Health Partners), Birmingham, UK

⁸Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne, UK

⁹Leeds Teaching Hospitals NHS Trust, Leeds, UK

¹⁰Health Economics and Decision Science, School of Health and Related Research, University of Sheffield, Sheffield, UK

*Corresponding author p.griffiths@sheffield.ac.uk

Background: Ultrasonography has been the mainstay of antenatal screening programmes in the UK for many years. Technical factors and physical limitations may result in suboptimal images that can lead to incorrect diagnoses and inaccurate counselling and prognostic information being given to parents. Previous studies suggest that the addition of in utero magnetic resonance imaging (iuMRI) may improve diagnostic accuracy for fetal brain abnormalities. These studies have limitations, including a lack of an outcome reference diagnosis (ORD), which means that improvements could not be assessed accurately.

Objectives: To assess the diagnostic impact, acceptability and cost consequence of iuMRI among fetuses with a suspected fetal brain abnormality.

Design: A pragmatic, prospective, multicentre, cohort study with a health economics analysis and a sociological substudy.

Setting: Sixteen UK fetal medicine centres.

Participants: Pregnant women aged ≥ 16 years carrying a fetus (at least 18 weeks' gestation) with a suspected brain abnormality detected on ultrasonography.

Interventions: Participants underwent iuMRI and the findings were reported to their referring fetal medicine clinician.

Main outcome measures: Pregnancy outcome was followed up and an ORD from postnatal imaging or postmortem autopsy/imaging collected when available. Developmental data from the Bayley Scales of Infant Development and questionnaires were collected from the surviving infants aged 2–3 years. Data on the management of the pregnancy before and after the iuMRI were collected to inform the economic evaluation. Two surveys collected data on patient acceptability of iuMRI and qualitative interviews with participants and health professionals were undertaken.

Results: The primary analysis consisted of 570 fetuses. The absolute diagnostic accuracies of ultrasonography and iuMRI were 68% and 93%, respectively [a difference of 25%, 95% confidence interval (CI) 21% to 29%]. The difference between ultrasonography and iuMRI increased with gestational age. In the 18–23 weeks group, the figures were 70% for ultrasonography and 92% for iuMRI (difference of 23%, 95% CI 18% to 27%); in the ≥ 24 weeks group, the figures were 65% for ultrasonography and 94% for iuMRI (difference of 29%, 95% CI 23% to 36%). Patient acceptability was high, with at least 95% of respondents stating that they would have iuMRI again in a similar situation. Health professional interviews suggested that iuMRI was acceptable to clinicians and that iuMRI was useful as an adjunct to ultrasonography, but not as a replacement. Across a range of scenarios, iuMRI resulted in additional costs compared with ultrasonography alone. The additional cost was consistently $< \pounds 600$ per patient and the cost per management decision appropriately changed was always $< \pounds 3000$. There is potential for reporting bias from the referring clinicians on the diagnostic and prognostic outcomes. Lower than anticipated follow-up rates at 3 years of age were observed.

Conclusions: iuMRI as an adjunct to ultrasonography significantly improves the diagnostic accuracy and confidence for the detection of fetal brain abnormalities. An evaluation of the use of iuMRI for cases of isolated microcephaly and the diagnosis of fetal spine abnormalities is recommended. Longer-term follow-up studies of children diagnosed with fetal brain abnormalities are required to fully assess the functional significance of the diagnoses.

Trial registration: Current Controlled Trials ISRCTN27626961.

Funding: This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 23, No. 49. See the NIHR Journals Library website for further project information.

Contents

List of tables	xiii
List of figures	xvii
List of abbreviations	xix
Plain English summary	xxi
Scientific summary	xxiii
Chapter 1 Introduction	1
Scientific and clinical background	1
<i>Ultrasonography and fetal screening</i>	1
<i>Development of magnetic resonance imaging</i>	1
<i>Magnetic resonance imaging in the literature: diagnostic accuracy</i>	1
<i>Magnetic resonance imaging in the literature: clinical management</i>	2
<i>Rationale for the MERIDIAN study</i>	2
Study aims and objectives	2
Chapter 2 Methods	5
Study design	5
<i>Participants and eligibility criteria</i>	5
<i>Participant identification and recruitment</i>	5
<i>Test methods</i>	5
<i>Data collection and management</i>	6
Statistical methods	7
<i>Sample size</i>	7
<i>Statistical analysis</i>	7
Health economic methods	13
<i>Background</i>	13
<i>Overview</i>	14
<i>Costs</i>	14
<i>Outcomes</i>	15
<i>Analyses</i>	15
Patient and public involvement	19
Study governance and management	19
<i>Study governance</i>	19
<i>Ethics arrangements and regulatory approvals</i>	20
<i>Protocol amendments after study initiation</i>	20
Chapter 3 Study results	21
Recruitment and participant flow	21
Characteristics	21
Uptake of in utero magnetic resonance imaging	21
Uptake of ultrasonography	21
Outcome of pregnancy	22

Diagnostic source for outcome reference diagnosis	22
Agreement	22
<i>Overall</i>	22
<i>Missing data</i>	23
<i>All participants with outcome reference diagnosis</i>	24
<i>Repeat in utero magnetic resonance imaging</i>	24
Anatomical subgroup analysis	25
<i>Isolated ventriculomegaly</i>	25
<i>Failed commissuration</i>	26
<i>Posterior fossa abnormalities</i>	27
Diagnostic confidence	28
Clinical impact	30
<i>Prognosis</i>	30
<i>Counselling and management</i>	30
Additional tests/resource use	32
Safety	33
In utero magnetic resonance imaging errors	33
Hospital Anxiety and Depression Scale	34
Health economic results	35
<i>Outcomes</i>	35
<i>Costs</i>	36
<i>Base case, scenario 1 and scenario 2</i>	36
<i>Cost-effectiveness</i>	36
<i>Imputed missing data</i>	39
<i>Probabilistic sensitivity analysis</i>	39
<i>Scenario analyses</i>	40
<i>Subgroup analyses</i>	40
<i>Comparison with repeat ultrasonography</i>	40
<i>Different costing scenarios</i>	40
<i>Longer-term horizon</i>	45
Chapter 4 Qualitative substudy	47
Introduction	47
Recruitment, sampling and data-collection methods	47
<i>Questionnaire surveys</i>	47
<i>Qualitative interviews with women</i>	49
<i>Qualitative interviews with health professionals</i>	49
<i>Qualitative data management</i>	49
Data analysis	50
<i>Data analysis approach for survey data</i>	50
<i>Data analysis for qualitative interview data</i>	50
Results	52
<i>Service user substudy participants</i>	52
<i>Patient survey findings</i>	53
<i>Patient qualitative interview findings</i>	54
<i>Health professional participant characteristics</i>	57
<i>Health professional qualitative interview findings</i>	58
Chapter 5 The 2- to 3-year follow-up study	65
Introduction	65
<i>Background</i>	65
<i>Study aims and objectives</i>	65

Methods	65
<i>Participants and eligibility criteria</i>	65
<i>Participant tracking and recruitment</i>	66
<i>Outcome measures and assessments</i>	66
<i>Sample size</i>	68
<i>Statistical methods</i>	68
<i>Protocol amendments after study initiation</i>	70
Results	70
<i>Participant flow</i>	70
<i>Participant characteristics</i>	73
<i>Project 1</i>	76
<i>Project 2</i>	76
<i>Project 3</i>	81
Chapter 6 Add-on study	83
Introduction	83
<i>Background</i>	83
<i>Study aims and objectives</i>	83
Methods	83
<i>Participants and eligibility criteria</i>	83
<i>Sample size</i>	84
<i>Statistical methods</i>	84
Results	84
Chapter 7 Discussion	85
Main findings	85
<i>Accuracy of diagnosis</i>	85
<i>Diagnostic confidence</i>	85
<i>Counselling and management</i>	86
<i>Prognostic importance</i>	86
<i>Acceptability to patients</i>	87
<i>Radiologists' experience</i>	87
<i>Health economics</i>	87
<i>Sociological substudy</i>	88
Strengths and weaknesses	90
<i>Strengths and weaknesses of the study</i>	90
<i>Strengths and weaknesses of the 2- to 3-year follow-up study</i>	91
<i>Strengths and weaknesses of the economic analysis</i>	91
<i>Strengths and weaknesses of the sociological substudy</i>	92
<i>Strengths and weaknesses of the add-on study</i>	92
Chapter 8 Conclusions	93
Implications for practice and policy	93
Recommendations for future research	94
Acknowledgements	95
References	99
Appendix 1 Unit costs (base case)	107
Appendix 2 Probability distributions for input parameters	109

Appendix 3 Recruitment and participant flow	113
Appendix 4 Ultrasonography technical factors	115
Appendix 5 Diagnostic accuracy	117
Appendix 6 Prognosis in relation to inclusion/exclusion in primary analysis	119
Appendix 7 Prognosis on ultrasonography alone versus in utero magnetic resonance imaging following ultrasonography for all mothers who completed in utero magnetic resonance imaging	121
Appendix 8 Change in prognosis from ultrasonography to in utero magnetic resonance imaging following ultrasonography by category when prognosis is unknown on one modality	123
Appendix 9 Change in counselling based on ultrasonography alone and in utero magnetic resonance imaging following ultrasonography in 774 cases with complete data	125
Appendix 10 Composite case study	127
Appendix 11 Participant characteristic tables	129
Appendix 12 Box plots of Ages and Stages Questionnaire, Third Edition, and Strengths and Difficulties Questionnaire scores	133
Appendix 13 Descriptive statistics tables for the developmental scores	137
Appendix 14 A breakdown of ultrasonography and in utero magnetic resonance imaging prognoses in 'abnormal', 'normal' and 'borderline' cases	141

List of tables

TABLE 1 Summary of analyses	16
TABLE 2 Characteristics of mothers and fetuses who were recruited to the MERIDIAN study and included in the primary analysis	22
TABLE 3 Outcome of pregnancy in 827 cases that completed iuMRI	23
TABLE 4 Diagnostic source used for ORD in 570 cases in the primary analysis	23
TABLE 5 Diagnostic accuracy of ultrasonography and iuMRI by gestational age in the 570 cases in the primary analysis population	23
TABLE 6 Diagnostic accuracy of ultrasonography and iuMRI by gestational age in the 638 cases who had scans and ORDs	24
TABLE 7 Diagnostic accuracy of ultrasonography and iuMRI in 65 cases with complete outcome data for which repeat iuMRI was used (the latest scan was used if more than two were completed)	25
TABLE 8 Diagnostic accuracy of ultrasonography and iuMRI in cases diagnosed with isolated VM on ultrasonography	25
TABLE 9 Diagnostic accuracy of ultrasonography and iuMRI in cases diagnosed with either ACC or hypogenesis of the CC on ultrasonography	26
TABLE 10 Diagnostic accuracy of ultrasonography and iuMRI in cases diagnosed with abnormalities of the posterior fossa on ultrasonography	28
TABLE 11 Diagnostic confidence categories on ultrasonography and iuMRI in the primary analysis population ($n = 570$)	29
TABLE 12 Difference in diagnostic confidence on ultrasonography alone and ultrasonography with iuMRI in the primary analysis population ($n = 570$)	29
TABLE 13 Difference in diagnostic confidence on ultrasonography and iuMRI in the primary analysis population ($n = 570$)	29
TABLE 14 Change in prognosis from ultrasonography to iuMRI, by category, for mothers who have a known prognosis on both modalities	30
TABLE 15 Counselling based on ultrasonography alone and iuMRI for all mothers who completed iuMRI	31
TABLE 16 Management of pregnancy following iuMRI in 783 cases with complete clinical feedback data	31
TABLE 17 Management of pregnancy by iuMRI contribution in 783 cases with complete data	32

TABLE 18 Summary of additional investigations performed in the 821 mothers who completed iuMRI	32
TABLE 19 Additional investigative tests and their impact on the prognosis in 559 mothers who received them	33
TABLE 20 Adverse events during iuMRI that occurred during the MERIDIAN study	34
TABLE 21 Error rates in relation to the experience category of the reporting radiologist	34
TABLE 22 Summary statistics of the HADS questionnaire completed as part of surveys 1 and 2	35
TABLE 23 'Was iuMR correct?' by 'Did iuMR change the prognosis?'	35
TABLE 24 'Did TOP happen?' by 'Was TOP offered after iuMR?'	36
TABLE 25 Disaggregated and total costs	36
TABLE 26 Deterministic results, imputed missing cost data results and probabilistic results	37
TABLE 27 Scenario analyses including personal costs	41
TABLE 28 Subgroup analyses by gestational age at iuMRI	42
TABLE 29 Scenario analysis comparing iuMRI with repeat ultrasonography	42
TABLE 30 Different costing scenarios	43
TABLE 31 Long-term scenario	46
TABLE 32 Summary of stages of thematic analysis	51
TABLE 33 Characteristics of substudy patient respondents	52
TABLE 34 Characteristics of substudy health professional participants	58
TABLE 35 Participant characteristics at entry to the MERIDIAN study	73
TABLE 36 Diagnostic accuracy of ultrasonography and iuMRI by age of fetus in the 574 cases with ORD, using updated ORD ($n = 5$) and new ORD ($n = 4$) for participants who completed iuMRI within 2 weeks	76
TABLE 37 Diagnostic accuracy of ultrasonography and iuMRI by age of fetus in 643 cases with ORDs, using updated and new ORDs	76
TABLE 38 Number and percentage of participants in each development category overall and according to each individual assessment tool	78

TABLE 39 Primary analysis of project 2: prognosis on ultrasonography and iuMRI compared with the neurodevelopmental category at 2- to 3-year follow-up in surviving infants only	79
TABLE 40 Primary analysis of project 2: prognosis on ultrasonography and iuMRI compared with the neurodevelopmental category at 2- to 3-year follow-up in surviving and non-surviving infants	80
TABLE 41 Primary analysis of project 2 repeated combining abnormal and at-risk groups	80
TABLE 42 Comparison of development categories according to BSID3 and ASQ-3 in participants who completed both assessments ($n = 84$)	81
TABLE 43 The number of cases correctly diagnosed as abnormal and normal using ASQ-3 alone when compared with BSID3	81
TABLE 44 Developmental assessment in relation to the original prognosis	81
TABLE 45 Prevalence of abnormal and at-risk neurodevelopment at 2- to 3-year follow-up among cases of isolated mild VM	82
TABLE 46 Inclusion in the follow-on study project 1 in relation to original prognosis	129
TABLE 47 Inclusion in the follow-on study project 2 in relation to the original prognosis	129
TABLE 48 Inclusion in the follow-on study project 2 in relation to the original prognosis	130
TABLE 49 Summary of developmental outcome scores from the BSID assessment	137
TABLE 50 Summary of developmental outcome scores from the ASQ-3	138
TABLE 51 Summary of developmental outcome scores from the SDQ	139
TABLE 52 Ultrasonography prognosis category compared with iuMRI prognosis category in surviving infants who completed the developmental assessment and were assessed as 'abnormal' ($n = 53$)	141
TABLE 53 Ultrasonography prognosis category compared with iuMRI prognosis category in surviving infants who completed the developmental assessment and were assessed as 'normal' or 'borderline' ($n = 125$)	141

List of figures

FIGURE 1 Flow diagram to show the pathway and recruitment numbers in the MERIDIAN study	8
FIGURE 2 Cost-effectiveness acceptability curve	39
FIGURE 3 The EVPI	40
FIGURE 4 Substudy survey flow chart	48
FIGURE 5 Responses to questions about the impact of iuMRI on understanding and decision-making in questionnaire 1 (black) and questionnaire 2 (green)	53
FIGURE 6 Participant flow in the MERIDIAN follow-on study	71
FIGURE 7 Participant flow in the project 1 component of the follow-on study	72
FIGURE 8 Participant flow in the project 2 component of the follow-on study	72
FIGURE 9 Bar chart showing the MERIDIAN study cohort ($n = 829$), by prognosis on (a) ultrasonography and (b) iuMRI for project 2	74
FIGURE 10 Bar chart showing the non-surviving infants in the MERIDIAN study cohort, by prognosis on (a) ultrasonography and (b) iuMRI ($n = 202$)	75
FIGURE 11 Box plots of BSID3 subscale composite scores, by ultrasonography and iuMRI prognosis in the MERIDIAN study	77
FIGURE 12 A waffle chart showing the MERIDIAN study cohort ($n = 829$)	131
FIGURE 13 A waffle chart showing the MERIDIAN study cohort ($n = 829$)	132
FIGURE 14 Box plots of ASQ-3 subscale scores by ultrasonography and iuMRI prognosis in the MERIDIAN study	133
FIGURE 15 Box plots of SDQ subscale scores by ultrasonography and iuMRI prognosis in the MERIDIAN study	135
FIGURE 16 Frequency of each ultrasonography prognosis category by developmental outcome	142
FIGURE 17 Frequency of each iuMRI prognosis category by developmental outcome	143

List of abbreviations

ACC	agenesis of the corpus callosum	MRI	magnetic resonance imaging
AE	adverse event	NPV	negative predictive value
ARC	Antenatal Results and Choices	OR	odds ratio
ASQ-3	Ages and Stages Questionnaire-3	ORD	outcome reference diagnosis
BSID3	Bayley Scales of Infant and Toddler Development, Third Edition	PIS	participant information sheet
CC	corpus callosum	PND	prenatal diagnosis
CI	confidence interval	PPI	patient and public involvement
CSF	cerebrospinal fluid	PPV	positive predictive value
CSP	cavum septum pellucidum	QALY	quality-adjusted life-year
CTRU	Clinical Trials Research Unit	RCOG	Royal College of Obstetricians and Gynaecologists
DMEC	Data Monitoring and Ethics Committee	REC	Research Ethics Committee
EOI	expression of interest	SAE	serious adverse event
EVPI	expected value of perfect information	SD	standard deviation
GMFCS	Gross Motor Function Classification Score	SDQ	Strengths and Difficulties Questionnaire
HADS	Hospital Anxiety and Depression Scale	SHINE	Spina Bifida, Hydrocephalus, Information, Networking, Equality
IQR	interquartile range	SOP	standard operating procedure
iuMRI	in utero magnetic resonance imaging	STARD	Standards for the Reporting of Diagnostic Accuracy Studies
MDT	multidisciplinary team	TMG	Trial Management Group
MIEP	Multidisciplinary Independent Expert Panel	TOP	termination of pregnancy
		TSC	Trial Steering Committee
		VM	ventriculomegaly

Plain English summary

Ultrasonography is routine in pregnancy to check that the baby's brain is developing as expected. However, no medical test is perfect and ultrasonography may miss some brain abnormalities, may get some brain abnormalities wrong or may diagnose an abnormality that is not really present. Magnetic resonance imaging (MRI) may help clarify difficult cases during pregnancy.

We wanted to find out if MRI was better than ultrasonography alone in making an accurate diagnosis. We recruited pregnant women whose ultrasound scan, performed by an expert, suggested that their baby had a brain abnormality, and referred them for a MRI scan. The results of the two tests were compared with each other and to the final outcome of the pregnancy.

Our results showed that using MRI in addition to ultrasonography improved the accuracy of the diagnosis in about one in four pregnancies. It changed the prediction of how the baby would develop in at least one in five cases. In many cases, the pregnancy was managed differently because of the MRI result. The MRI was acceptable to women, with 95% saying that they would have MRI again in a similar situation. Neither MRI nor ultrasonography accurately identified children who went on to have delayed development at the age of 2–3 years, but MRI was better than ultrasonography at ruling out developmental problems at this age. The MRI cost more than ultrasonography alone; therefore, whether or not it is worthwhile depends on the value placed on the decisions that changed as a result of its use.

Scientific summary

Background

For many years, fetal imaging with ultrasonography has been the mainstay of antenatal screening programmes in the UK. However, no imaging method is perfect and various technical factors and physical limitations may result in suboptimal images. These may lead to incorrect diagnoses of structural abnormalities and inaccurate counselling and prognostic information being given to parents. The fetal brain is a particular area of concern because of the relatively high frequency of developmental abnormalities (approximately 3/1000 pregnancies), many of which are associated with serious clinical morbidities.

Previous studies have suggested that in utero magnetic resonance imaging (iuMRI) may be a powerful adjunct to ultrasonography for detecting fetal brain abnormalities from 18 weeks' gestational age. However, the majority of these studies have a number of key limitations and, most notably, they lack an outcome reference diagnosis (ORD), which means that improvements in diagnostic accuracy could not be assessed accurately. Therefore, the extent of diagnostic improvement and the impact that this has on clinical management and counselling remains unclear.

Aims and objectives

The aim of the research was to assess iuMRI to aid the prenatal diagnosis (PND) of fetal developmental brain abnormalities among fetuses for which ultrasonography had suggested a brain abnormality.

The objectives were to:

- assess the diagnostic accuracy of iuMRI compared with antenatal ultrasonography using ORD as a comparator
- assess the clinical effectiveness of iuMRI through changes in clinical diagnostic confidence before and after iuMRI and the effect on prenatal counselling and management decisions
- assess the acceptability of the clinical care package with the use of magnetic resonance imaging (MRI) included
- conduct a health economics analysis to assess the cost consequences of the use of MRI.

Methods

Design

We conducted a pragmatic, prospective, multicentre, cohort study with a health economics analysis and a qualitative substudy.

Recruitment and imaging examinations

Participants were pregnant women aged > 16 years and carrying a fetus with a suspected brain abnormality on detailed ultrasonography as diagnosed by a fetal medicine consultant. They were recruited by 16 fetal medicine centres and referred for iuMRI. There were no specific requirements made for the ultrasound technique and details of suspected brain abnormalities were recorded along with a confidence rating of the diagnosis. The iuMRI was completed at one of six collaborating centres where radiologists were required to comment on each abnormality suspected by the ultrasonography and their confidence in the diagnosis. The radiologist could add a diagnosis if appropriate, or use 'diagnosis excluded' if they disagreed with an ultrasonography diagnosis. It was not possible to match protocols across all scanning centres because of different MRI manufacturing systems; however, all scans were performed at 1.5 T and an absolute requirement of imaging sequences was defined.

Outcome reference diagnoses

This was defined as postnatal neuroimaging for continued pregnancies with a surviving child and autopsy and/or postmortem MRI in cases of termination of pregnancy, stillbirth or neonatal death. Agreement between ultrasonography, iuMRI and ORD was determined by a two-level review process including a Multidisciplinary Independent Expert Panel.

Diagnostic accuracy

Participants who underwent iuMRI within 14 days of ultrasonography and for whom an ORD was available were included in the primary analysis.

In cases with multiple anatomical diagnoses, all had to be reported accurately to be classified as 'correct'. We estimated the overall diagnostic accuracy, defined as:

$$(\text{true positives})/\text{total for ultrasonography and } (\text{true positives} + \text{true negatives})/\text{total for iuMRI.} \quad (1)$$

This percentage is equivalent to the positive predictive value for ultrasonography. Diagnostic accuracy was calculated for both gestational age groups (i.e. 18–23 weeks and ≥ 24 weeks) and overall using McNemar's paired binomial test.

Diagnostic confidence and clinical impact

Data on diagnostic confidence and the clinical impact of the iuMRI on the counselling and management of pregnancies were also collected for analysis.

Economic evaluation

This determined whether or not the new intervention (iuMRI) represented good value for money compared with current practice (ultrasonography alone). Information on participant resource use and outcomes were analysed to estimate the total additional costs and change in outcomes associated with iuMRI in the diagnosis of fetal brain abnormalities. In the base case, the economic analysis compares the costs and outcomes associated with iuMRI with ultrasonography alone. In scenario analyses, iuMRI is compared with repeat ultrasonography.

Sociological substudy

The aim of the patient and health professional perspectives substudy was to assess the acceptability of the clinical care package including iuMRI using qualitative and quantitative methods. This substudy drew on quantitative data from patients and qualitative data from patients and health professionals to assess the acceptability of iuMRI in the PND for fetal brain abnormality. This mixed-method approach provided an in-depth analysis of key issues and experiences that matter to those with first-hand knowledge of iuMRI in practice.

Three data sets (i.e. women's survey responses, women's interviews and professionals' interviews) were used to explore service user and health professional experiences of engaging with the clinical study. Survey 1 was administered after the fetal medicine consultation following iuMRI (when a management plan had been made), and survey 2 was administered 3–6 months after the pregnancy outcome. Participants who responded to survey 2 formed the sampling frame for the qualitative interviews with women, and purposive sampling identified the pool for health professional interviews.

Add-on study

Recruitment into the MERIDIAN study included only women with fetuses thought to have had brain abnormalities on ultrasonography, which allowed us to assess the diagnostic accuracy of ultrasonography and iuMRI in terms of the false-positive rate of detecting a fetal brain abnormality; however, the false-negative rate cannot be estimated. The add-on study addressed this by recruiting pregnant women carrying a fetus without any form of abnormality detected on detailed ultrasonography.

Participants were recruited via the MERIDIAN study fetal medicine centres and local media advertisement. Participants could either be referred directly or self-refer. Potential participants were initially screened by telephone and, if deemed eligible, attended the unit for the iuMRI study where further screening took place prior to obtaining written informed consent.

The 2- to 3-year follow-up study

Surviving infants were followed up at 2–3 years of age in order to assess three aims: (1) whether or not additional diagnoses had become apparent, (2) what impact their diagnosis had on their neurodevelopment (i.e. the accuracy of the prognosis following ultrasonography or iuMRI) and (3) the extent to which these matter in the specific subgroup of isolated, mild ventriculomegaly (VM), which is the most frequent single diagnosis. The issue of prognosis is particularly relevant because it is the functional significance of the diagnosis that matters most to parents rather than the brain abnormality per se. Research has shown that following the diagnosis of a fetal abnormality, parents want to hear all possible outcomes. Preparation for the potential outcomes allows families to build a picture of what those difficulties would be like to live with and helps them consider other aspects, such as the emotional impact and the implications for family life.

The surviving infants who underwent ultrasonography and iuMRI as part of the original study, and whose families gave consent to be approached about future studies, were screened and the family was approached to participate in the follow-up study. The first aim involved a review of the child's medical case notes. This review allowed us to refine our estimates of diagnostic accuracy based on more recent imaging of the infant's brain. We also collected information about follow-up care and life events that may have had an impact on the child's development.

For the second aim, parents were asked to complete the Ages and Stages Questionnaire and their children were invited to participate in a developmental assessment, the Bayley Scales of Infant Development, Third Edition (BSID3). The assessments were completed either in the local recruiting hospital or, when possible, in the child's home. The cognitive, language and motor components were completed.

In addition to the BSID3, or when a face-to-face appointment could not be arranged, a further two questionnaires were completed: the Strengths and Difficulties Questionnaire and the MERIDIAN Gross Motor Skills Questionnaire. Following these assessments, each child's neurodevelopment was classified as 'normal', 'at risk' or 'abnormal'.

The third objective was to assess outcomes of children diagnosed with an isolated, mild VM on either ultrasonography or iuMRI.

Results

Between July 2011 and August 2014, 903 participants (911 fetuses) were recruited to the study. A total of 823 participants (829 fetuses) successfully underwent iuMRI. In total, 64% of all participants attended the Academic Unit of Radiology at the University of Sheffield for their iuMRI, whereas the remaining 36% of participants attended one of the five collaborating centres. An ORD was available for 638 out of 829 (77%) fetuses; of these 638, 570 (89%) had the iuMRI performed within 2 weeks of the referral ultrasonography. A total of 369 fetuses (65%) were in the 18–23 weeks group (110% of required participants) and 201 fetuses (35%) were in the ≥ 24 weeks group (120% of required participants). The three commonest ultrasonography diagnoses were isolated VM (306/570, 54%), an abnormality restricted to the contents of the posterior fossa (83/570, 15%) and failed commissuration (79/570, 14%).

The overall diagnostic accuracies of ultrasonography and iuMRI were 68% and 93%, respectively, with a difference of 25% [95% confidence interval (CI) 21% to 29%]. The difference between ultrasonography and iuMRI increased with gestational age. In the 18–23 weeks group, the figures were 70% for ultrasonography and 92% for iuMRI (difference of 23%, 95% CI 18% to 27%); in the ≥ 24 weeks group, the figures were 65% for ultrasonography and 94% for iuMRI (difference of 29%, 95% CI 23% to 36%).

In 386 out of 570 cases (67.7%) both ultrasonography and iuMRI reports were correct, and in 144 out of 570 cases (25.3%), the ultrasonography report was incorrect but the iuMRI report was correct. There were two fetuses (0.4%) in whose cases the ultrasonography was correct and the iuMRI was incorrect, and 38 fetuses (6.8%) in whose cases both the ultrasonography and iuMRI were incorrect.

In total, 198 participants (205 fetuses) were recruited to the add-on study and underwent iuMRI. Although the results confirm the ability of both ultrasonography and iuMRI to correctly confirm when brain development of the fetus is normal, there was one instance of a brain abnormality that ultrasonography had not detected but which was apparent on iuMRI. This highlights the validity of ultrasonography as the primary screening imaging method for pregnancy.

The 2- to 3-year follow-up study recruited 238 participants, 10 of whom had a new ORD available (five previously had no ORD and another five had a diagnosis that differed from the original ORD). The overall difference in diagnostic accuracy was revised to 24.8% (95% CI 21.0% to 28.5%; $p < 0.0001$), which was very similar to the original difference of 24.9% and still statistically significant. Analysis of the developmental data showed minimal difference between ultrasonography and iuMRI in prognosticating abnormal development. However, iuMRI demonstrated a higher number of correct prognoses in surviving infants who were developmentally assessed as 'normal' or 'at risk'.

In total, 19 (22%) participants diagnosed with isolated, mild VM on ultrasonography had abnormal neurodevelopment, with a further 12 (14%) participants being considered 'at risk'. Among those with a diagnosis of isolated, mild VM on iuMRI, the incidences were 18% (13/71 of patients) among those with abnormal neurodevelopment and 13% (9/71 of patients) among those 'at risk'. None of the children had received a poor prognosis on either modality. The higher than expected incidence of abnormal neurodevelopment in some of the children could be attributed to additional brain diagnoses (some of which the iuMRI detected correctly) other (non-brain) diagnoses and life events (e.g. accidents) leading to delayed development.

Health economic evaluation

The use of iuMRI resulted in additional costs compared with ultrasonography alone, owing to the cost of iuMRI itself and the costs from the resulting management decisions. Across a range of scenarios, the incremental cost was consistently $< \pounds 600$ per participant and the cost per management decision appropriately changed was always $< \pounds 3000$.

Sociological substudy

Patient acceptability was high, with 97% of questionnaire 1 and 95% of questionnaire 2 respondents stating that they would have iuMRI as part of their care if they were ever in a similar situation again. In response to the question about whether or not iuMRI was undertaken at an acceptable time and place, 73% of questionnaire 1 and 68% of questionnaire 2 respondents strongly agreed or agreed. Interviews suggested that participants were usually 'information hungry' and prepared to tolerate significant discomfort to maximise information prior to decision-making about their pregnancy. Health professional interviews suggested that iuMRI was broadly acceptable to clinicians and it was consistently noted that iuMRI was useful as an adjunct to ultrasonography, but not as a replacement.

Conclusions

Implications for health care

Our study provides robust data indicating that there would be significant benefits from the routine use of iuMRI as an adjunct to ultrasonography in the diagnostic pathway for identifying fetal brain abnormalities.

The development of a service that includes iuMRI in the diagnostic pathway may consider providing the service at a small number of supraregional centres and/or promoting collaboration between radiologists to form expert panels for cases, provide formal training to improve radiologists' knowledge base and having the iuMRI study supervised by the radiologist and the images double reported.

Recommendations for research

- Evaluate the use of iuMRI for cases of isolated microcephaly identified on ultrasonography, as this is often associated with other brain abnormalities and the severity correlates with the risk of poor outcome.
- Evaluate the use of iuMRI in the diagnosis of fetal spine abnormalities that, to date, have not been well studied; however, there is a small body of research that suggests that iuMRI may have a role in the diagnosis of these abnormalities.
- Longer-term follow-up studies of children diagnosed with fetal brain abnormalities to fully assess the functional significance of the diagnoses.

Trial registration

This trial is registered as ISRCTN27626961.

Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.

Chapter 1 Introduction

Scientific and clinical background

Ultrasonography and fetal screening

For many years, fetal imaging with ultrasonography has been the mainstay of antenatal screening programmes in the UK. However, no imaging method is perfect and various technical factors and physical limitations of ultrasonography may result in suboptimal images of the fetus being obtained. Suboptimal images may lead to incorrect diagnoses of structural abnormalities and inaccurate counselling and prognostic information being given to parents. The fetal brain is a particular area of concern because of the relatively high frequency of developmental abnormalities (approximately 3/1000 pregnancies).¹ A wide spectrum of neuropathologies can be shown, many of which are associated with serious clinical morbidities.

Development of magnetic resonance imaging

Magnetic resonance imaging (MRI) technology has advanced over the years and because of significant improvements in spatial and contrast resolution, highly reliable and accurate diagnoses of comparable pathology can now be made for children. In the 1990s, MRI became a realistic clinical possibility because of advances in hardware and software. The Academic Unit of Radiology at the University of Sheffield was a pioneer in this field.² Clinical magnetic resonance sequences have been developed that do not require maternal sedation or fetal neuromuscular blockade, allowing several groups, including our own, to believe that in utero magnetic resonance imaging (iuMRI) may be a powerful adjunct to ultrasonography for detecting fetal brain abnormalities from 18 weeks' gestational age.

Magnetic resonance imaging in the literature: diagnostic accuracy

Much of the early literature focused on describing the techniques required to perform iuMRI and relied on anecdotal cases to compare the additional information provided with ultrasonography alone.^{3–8} A key limitation of the majority of the literature was the lack of an outcome reference diagnosis (ORD), with discrepancies between iuMRI and ultrasonography generally assumed as errors on the part of the latter. Doing so inevitably leads to bias in favour of iuMRI, and fetal ultrasonography experts have criticised these studies for the artificially high detection rates for iuMRI that resulted from biased patient selection,^{9,10} although the extent of this bias is unknown. One such study was published by our group;¹¹ it was a study of 100 singleton pregnancies in which a fetal brain abnormality was suspected on ultrasonography but for which optimal diagnostic information had not been obtained. When iuMRI was used in conjunction with ultrasonography, an increase in diagnostic accuracy of 48% was observed. Although these findings are clearly prone to the aforementioned bias, they nevertheless demonstrate the potential for a significant improvement in diagnoses based on ultrasonography and provide an upper limit on the improvement attributable to the introduction of iuMRI.

Ventriculomegaly (VM) has been the focus of much research in the field of fetal neuroimaging with iuMRI because of the high prevalence of the finding (1 or 2 out of 1000 pregnancies) following a screening ultrasonography. In our more recent study, 147 fetuses were diagnosed with isolated VM on ultrasonography, with high confidence and no technical limitations.¹² When iuMRI was performed in these cases, additional clinically important brain findings were detected in 17% of cases. In a second study of 61 fetuses with cerebral VM, iuMRI was found to provide more information than ultrasonography in 33% of cases and was able to identify the cause of the VM in 21% of cases.¹³ A further study of 185 fetuses with isolated VM in the third trimester found that 11 out of 185 (5.94%) fetuses had other brain abnormalities.¹⁴ The use of iuMRI as an adjunct to ultrasonography for detecting fetal brain abnormalities has been further supported by recent studies and systematic reviews;^{15–18} however, the extent of diagnostic improvement remains unclear.

Magnetic resonance imaging in the literature: clinical management

The literature on the clinical impact of iuMRI on the counselling and management of pregnancies complicated by fetal brain abnormalities is limited. One prospective study of fetuses with a suspected central nervous system abnormality observed that 24 out of 52 (46%) pregnancies were managed differently after iuMRI,¹⁹ and a further prospective study found that the counselling and management of pregnancies were changed after iuMRI in 49.7% and 13.5% of cases, respectively.²⁰ Our study of fetuses with isolated VM found significant changes in clinical management after iuMRI in the majority of cases when additional brain abnormalities were identified. Changes in management more frequently happened between the gestational ages of 20 and 24 weeks.¹² Our group conducted a systematic review and meta-analysis¹⁸ including 34 studies, of which only 11 reported on the changes in counselling or management of pregnancies brought about by the iuMRI. These 11 studies^{11,21–30} comprised 186 fetuses and changes were observed in 78 (41.9%) cases. Our review concluded that the clinical impact of iuMRI through changes to the counselling and management of pregnancies remains unclear owing to being reported in only a small proportion of studies.

It is important to recognise that iuMRI is unlikely to be incorporated as a routinely used diagnostic tool for pregnancies. Constraints on resources, both financial and human, mean that iuMRI is likely to be used selectively, in combination with ultrasonography and other tests, for fetuses that are considered to be at a high risk of severe developmental abnormalities. The term ‘high risk’ can be interpreted in many different ways and may legitimately incorporate genetic features and/or a previous pregnancy with a fetal abnormality, but the most obvious risk factor is when the fetus has an abnormality diagnosed by the ultrasonography. Therefore, the project set out to determine whether or not adding iuMRI to the usual diagnostic pathway helps to improve the diagnosis and appropriate management of fetuses with a suspected brain abnormality based on previous ultrasonography.

Rationale for the MERIDIAN study

In the light of the uncertainties surrounding the extent of improvement in diagnostic accuracy associated with iuMRI as an adjunct to ultrasonography and its impact on clinical management and counselling, a large prospective study with a representative population of fetal brain abnormalities was required to inform clinical practice in the UK.

For a full evaluation of the use of iuMRI in clinical practice, it was essential that patient acceptability was assessed. New technologies in fetal medicine can raise ethical and social dilemmas for the patients involved and, therefore, the views of patients and their partners are important when considering the implementation of iuMRI.³¹ A survey of 227 women who were not pregnant but underwent MRI suggests that it is perceived positively in terms of comfort and impact on care;³² however, other studies have shown a potential association between MRI and anxiety^{33,34} and that patients may find it less acceptable than ultrasonography.³⁵ A study published in 2007 found that overall satisfaction with the prenatal diagnosis (PND) is high in women carrying a fetus with a congenital abnormality.³⁶ Satisfaction with prenatal care has been found to be associated with the attitude of the medical staff,^{37,38} the amount of information provided³⁸ and the patient’s involvement in decision-making.³⁷ Data on the use of MRI in fetomaternal medicine are limited but do suggest that MRI may cause additional distress, especially in women for whom the prognosis for the fetus is poor,^{39,40} and can lead to more anxiety than that of undergoing an ultrasonography.⁴¹

The research described in this report was a logical extension of our previous work and essential in providing definitive data on the diagnostic accuracy, clinical impact and acceptability of iuMRI in clinical practice in the UK.

Study aims and objectives

The aim of the research was to assess iuMRI as a technology to aid the PND of fetal developmental brain abnormalities.

The objectives were to:

- assess diagnostic accuracy of iuMRI compared with antenatal ultrasonography through the –
 - measurement of the diagnostic accuracy of antenatal ultrasonography alone (i.e. prior to iuMRI) relative to an ORD (postnatal imaging or postmortem examination)
 - measurement of the diagnostic accuracy of iuMRI (following antenatal ultrasonography) relative to an ORD (postnatal imaging or postmortem examination).
- assess the clinical effectiveness of iuMRI through the –
 - change in clinical diagnostic confidence before and after an iuMRI study
 - effect of iuMRI on prenatal counselling and management intent.
- assess the acceptability of the clinical care package with the use of MRI included via quantitative and qualitative psychosocial measures
- undertake an economic analysis to evaluate the cost consequence of the use of iuMRI scans.

Chapter 2 Methods

This report is concordant with the Standards for the Reporting of Diagnostic Accuracy Studies (STARD) guidelines updated in 2015.⁴²

Study design

The MERIDIAN study was a prospective, multicentre, observational cohort study.

Participants and eligibility criteria

Participants were pregnant women aged ≥ 16 years and carrying a fetus with a suspected brain abnormality on detailed ultrasonography as diagnosed by a fetal medicine consultant. The potential participants were screened against the following criteria.

Inclusion criteria

A participant was eligible for the trial if the following criteria were met:

- The participant had an ongoing singleton or multifetal pregnancy of $\geq 18^{+0}$ weeks' gestation by ultrasonography dating.
- The participant was thought to be carrying a fetus with a brain abnormality following detailed specialist ultrasonography examination.

Exclusion criteria

A participant was excluded from the trial if any of the following criteria were met:

- The participant was unable to give informed consent.
- The participant had a cardiac pacemaker, had an intraorbital metallic foreign body or recently had surgery with metallic sutures or an implant.
- The participant had previously experienced or is likely to suffer from severe anxiety or claustrophobia in relation to a MRI examination.
- The participant was unable or unwilling to travel to Belfast, Birmingham, Leeds, Newcastle, Nottingham or Sheffield for specialist MRI.
- The participant was unable to understand the English language (except where satisfactory translation services were available).
- The participant was aged < 16 years.

Participant identification and recruitment

Potentially eligible participants were identified from 16 participating fetal medicine units that acted as research sites. Following a detailed ultrasonography examination that confirmed the suspicion of a fetal brain abnormality, consecutive patients were screened by the referring clinician or local research midwife according to the inclusion and exclusion criteria listed above. Written informed consent was obtained after the confirmation of eligibility and after the patient had been given the opportunity to consider participation and ask any questions. *Data collection and management* provides details of the referral process once consent had been obtained.

Test methods

Detailed ultrasonography

The protocol did not specify the exact requirements for the ultrasound technique. All ultrasonographies were performed by fetal medicine consultants working within the UK NHS who reported the suspected fetal brain abnormalities using nomenclature from the ViewPoint™ antenatal ultrasonography-reporting software

(GE Healthcare, Chalfont St Giles, UK). Recruiting clinicians were asked to record their certainty of the diagnosis alongside each brain abnormality using a five-point Likert scale⁴³ as follows: very unsure (10% certain), unsure (30% certain), equivocal (50% certain), confident (70% certain) and highly confident (90% certain). Following the detailed ultrasonography, participants underwent iuMRI at one of the six collaborating sites.

In utero magnetic resonance imaging

Magnetic resonance imaging protocols could not be matched exactly across all the scanning sites because of the different manufacturers of MRI systems being used. However, all examinations were performed at 1.5 T and there was an absolute requirement to obtain T2-weighted images of the fetal brain in the three orthogonal planes using the best, ultrafast method available (maximum slice thickness of 5 mm) and a T1-weighted ultrafast sequence in at least one plane (usually axial).⁴⁴ The attending radiologist could add further sequences to the case as appropriate to obtain thorough diagnostic information. To reflect clinical practice, the radiologist was aware of the diagnoses and the level of certainty made by the fetal medicine consultant prior to the iuMRI being performed and had access to the full clinical ultrasonography report. The radiologist was required to comment on each brain abnormality diagnosed by the ultrasonography. If the radiologist did not agree with a diagnosis made by ultrasonography then 'diagnosis excluded' could be used and an additional diagnosis could be added if appropriate. The radiologists were also required to indicate their certainty for each diagnosis using the same five-point Likert scale⁴³ as the ultrasonography assessment.

Data collection and management

After the detailed ultrasonography performed by a fetal medicine consultant, participants were recruited into the study and referred for an iuMRI scan at the Academic Unit of Radiology at the University of Sheffield or at one of the five collaborating sites (Belfast, Birmingham, Leeds, Newcastle or Nottingham).

The referral form was completed by the referring clinician and a copy was sent to the MRI site through the standard referral pathway. The referral form collected details of the type of brain abnormalities suspected, confidence of diagnosis and the prognosis for the outcome of the pregnancy. It also collected data on whether or not termination of pregnancy (TOP) was discussed or offered because the abnormalities on ultrasonography were sufficient to consider that option under Ground E of the Abortion Act 1967 [section 1(1)(d) – substantial risk of serious mental or physical disability].⁴⁵

The reporting radiologist completed a similar iuMRI feedback form and was required to comment on all suspected diagnoses along with their diagnostic confidence. Additional diagnoses could be added or excluded when required, as detailed previously.

Following the iuMRI examination, the participant attended a follow-up appointment with the referring fetal medicine consultant. The iuMRI report was available to the clinician at this appointment and the clinical feedback form was completed. The purpose of the clinical feedback form was to assess the changes in the clinical management and prognosis of the pregnancy after the iuMRI. In addition, further details of diagnostic and prognostic information were collected, along with the management plan for the pregnancy. From a diagnostic perspective, participants were asked if iuMRI had provided additional information and, if so, if the 'new' pathology was visible when the clinician performed a follow-up ultrasonography. To assess the prognosis, the clinicians were asked if iuMRI had changed the prognostic information given to the patient and to regrade the prognosis in the light of the iuMRI information using the same five categories used at referral. The clinicians also recorded whether or not TOP had been discussed or offered. Finally, the clinicians were requested to state if the iuMRI had altered the counselling and management of the pregnancy and, if so, what the degree of change was.

The pregnancy was then followed up for up to 6 months post delivery and data were collected on the outcome of the pregnancy, as described in *Statistical analysis*. The outcome data form was used to collect relevant information.

Statistical methods

Sample size

A recruitment target of 750 pregnant women was set, from whom we anticipated 504 fetuses would have a complete ORD. It was assumed that 336 of these fetuses would be 18 to < 24 weeks' gestation at the time of iuMRI. A diagnostic accuracy of 70% in ultrasonography^{46–54} and 80% in iuMRI was assumed, and it was further assumed that iuMRI would correctly contradict the ultrasonography in 20% of cases and erroneously contradict the ultrasonography in 10% of cases. A sample size of 336 participants allowed the confidence interval (CI) for the difference to be estimated to within a standard error of 3%, and ensured that any improvement in diagnostic accuracy of < 5% with 90% power and at 85% confidence could be ruled out. A change of 5% was deemed to be clinically important as it would lead to changes in fetal prognostic information and management intent in $\approx 5\%$ of all cases. A total recruitment target of 504 fetuses with complete ORDs was set, assuming that one woman with a fetus of > 24 weeks' gestation would be scanned for every two fetuses in the 18 to < 24 weeks' gestation age group.

Statistical analysis

The diagnostic strategies are referred to as ultrasonography (diagnostic ultrasonography alone) and iuMRI (iuMRI following diagnostic ultrasonography).

Statistical analysis was conducted using R version 3.3.3 (The R Foundation for Statistical Computing, Vienna, Austria) and Stata® version 14 (StataCorp LP, College Station, TX, USA).

Analysis populations

Owing to the nature of the study, some outcomes were collected at the mother level (e.g. other diagnostic tests, participant feedback). In some instances, more than one record was available for the same mother because of repeat ultrasonography and/or MRI scans, multiple pregnancies or repeat pregnancies. The denominator used is outlined throughout the methods and the results.

The primary analysis of diagnostic accuracy was undertaken on all fetuses for which a successful iuMRI scan had been undertaken within 2 weeks of their diagnostic ultrasonography and for whom an ORD was available. The 2-week time window was chosen so as to minimise the bias arising because of the evolving brain and, in some cases, the evolving diagnosis. A long delay between scans may favour iuMRI because diagnoses may become more apparent as the brain matures. Secondary analyses of diagnostic accuracy were undertaken on fetuses:

- with successful ultrasonography and iuMRI scans (irrespective of their timing) and with ORD
- with successful ultrasonography and iuMRI scans undertaken within 2 weeks, with or without ORD
- with successful ultrasonography and iuMRI scans, irrespective of timing and availability of ORD.

For all these analysis populations, the analyses were based on the first diagnostic ultrasonography and iuMRI scan unless otherwise stated. When groups were large enough, analyses were conducted on different subsets (i.e. analysis populations) of the data. Details of which population was assessed for each outcome are described in this section.

Within each of these populations, when appropriate, analyses were repeated on the following subgroups:

- first pregnancies (i.e. excluding any pregnancies subsequent to their first within the MERIDIAN study)
- singleton pregnancies
- first and singleton pregnancies
- repeat pregnancies
- repeat scans (i.e. second ultrasonography and iuMRI for the same pregnancy).

Finally, analyses of diagnostic accuracy were intended to be undertaken separately for pregnancies at < 22 weeks', 22–24 weeks' and > 24 weeks' gestational age at the time that the MRI scan was performed. However, these three categories were condensed to 18–24 weeks' and > 24 weeks' gestational age because of small numbers in the first two subgroups.

Summaries and analyses of diagnostic confidence were undertaken on the primary analysis population and on all fetuses undergoing scans. Summaries and analyses of changes in prognosis and management were undertaken on all mothers undergoing both scans. Summaries of safety and adverse outcomes were undertaken separately for all mothers undergoing a scan and all fetuses undergoing a scan.

Approach/consent

The number of mothers who were approached and consented is presented in the STARD flow diagram (Figure 1). When given, the reasons for not offering iuMRI and reasons for not consenting to join the study have been presented.

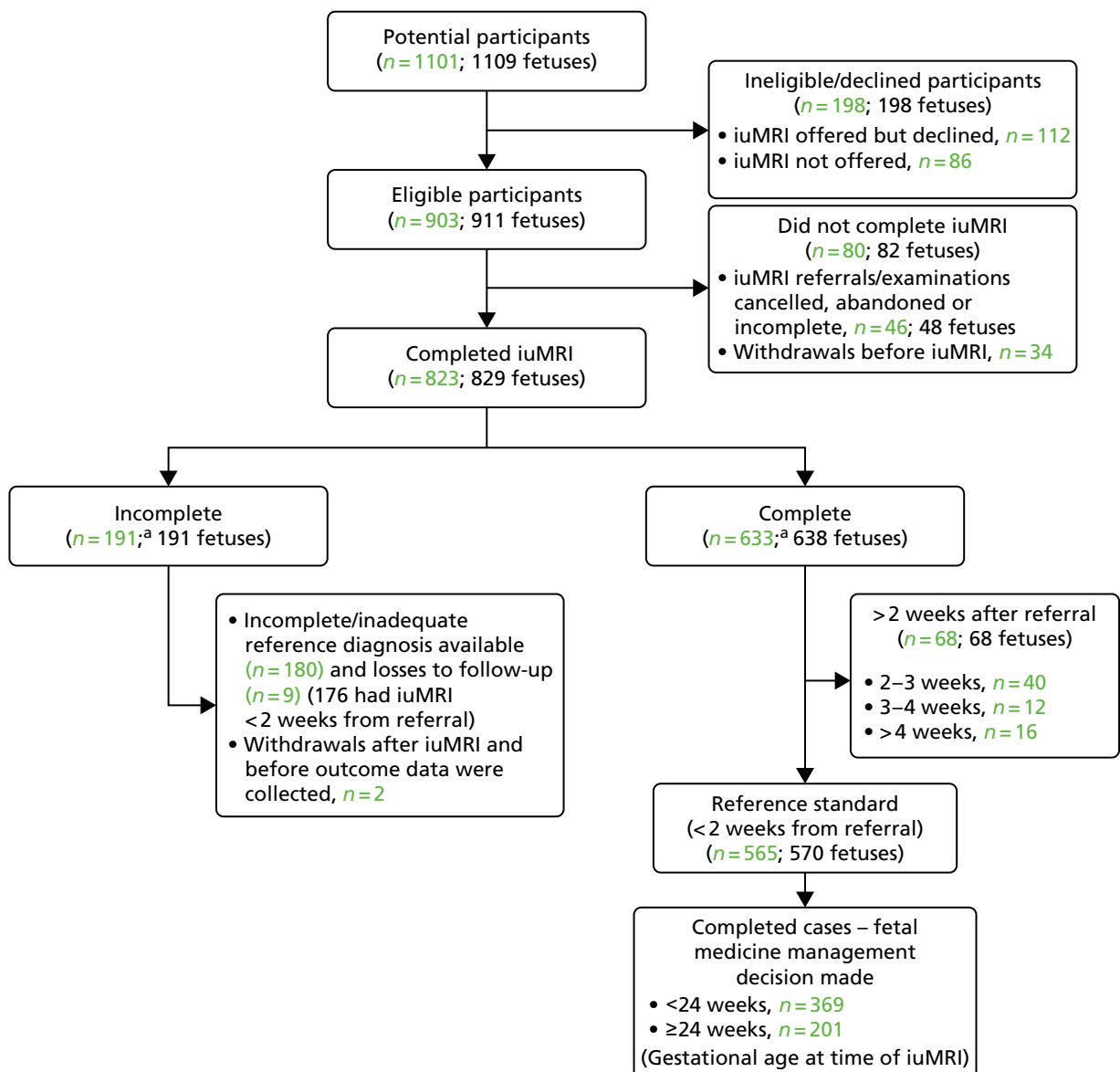


FIGURE 1 Flow diagram to show the pathway and recruitment numbers in the MERIDIAN study. a, One participant carried two fetuses, only one of which had an ORD. Therefore, the participant is counted in both the incomplete and complete participant groups.

Data completeness

The number of participants who completed data collection at each stage of the study is presented in the STARD flow diagram (see *Figure 1*). This included the number of participants who completed iuMRI, the number of those with a reference diagnosis available, the number of participants who withdrew or were lost to follow-up, the number of complete cases and at what gestational age a management decision was made.

Characteristics

The following summaries were presented for all mothers who entered the trial:

- iuMRI centre
- maternal age
- gestational age at initial approach
- whether the pregnancy was single or multiple
- dominant diagnosis of brain abnormality based on screening ultrasonography.

The number and percentage in each category were presented for categorical variables and appropriate summary statistics were presented for continuous variables. These figures were presented by population groups 1–3 (i.e. primary, ORD unavailable and excluded).

Diagnostic accuracy

Diagnostic accuracy was defined as the percentage of true-positive diagnoses for ultrasonography and the percentage of true-positive and true-negative diagnoses for iuMRI. This percentage is equivalent to the positive predictive value (PPV) for ultrasonography screening.

Primary analysis (agreement)

For the purposes of our trial, the ORD was defined as follows:

- For delivered neonates, the reference used was the neuroanatomical diagnosis recorded in the child's clinical notes at 6 months of age, based on all available follow-up imaging (i.e. postnatal MRI, computed tomography when performed for clinical purposes or postnatal transcranial ultrasonography). The exceptions were (1) cases of isolated VM, for which a resolved diagnosis made via third trimester ultrasonography was acceptable, and (2) isolated microcephaly, for which clinical examination was sufficient.
- For terminated fetuses, the reference diagnosis was based on postmortem data (i.e. autopsy and/or postmortem MRI when available).

Agreement between the PND and the reference diagnosis was determined by a two-level review process. The first-level review was carried out by one of two neuroradiologists, who were not associated with the MERIDIAN study. Their role was to determine whether or not a full review by the Multidisciplinary Independent Expert Panel (MIEP), described below, was required, based on the ORD. A full review by the MIEP was required unless the following criteria were met: (1) there was complete and unequivocal agreement between the anatomical findings on ultrasonography, iuMRI and the ORD, or (2) VM was the only finding described on both ultrasonography and iuMRI examinations, but the size of the ventricles had returned to normal as shown on ultrasonography later in pregnancy or on neonatal imaging. The latter was counted as agreement because the enlargement of ventricles can commonly resolve spontaneously during pregnancy.⁵⁵

The second-level review (by the MIEP) consisted of three NHS consultants (i.e. a neuroradiologist, a fetal and maternal medicine specialist and a paediatric neurologist) from a single centre, independent of the MERIDIAN study. The MIEP was provided with the diagnostic results for each fetus and members were blinded to knowing if it was an ultrasonography or iuMRI report. The MIEP was responsible for deciding whether or not each report agreed completely with the ORD (all listed diagnoses correct) and, when it was judged that there was disagreement between ultrasonography and iuMRI, which one indicated the more severe pathology. The MIEP could request additional information if required to assess a case and could be

given access to the full report. In the small number of cases when the full clinical report was required to assess a case, blinding was no longer possible. The research team at the Sheffield Clinical Trials Research Unit (CTRU) was able to unblind the results for the MIEP for analysis purposes.

The number of fetuses diagnosed, diagnostic accuracy (i.e. the proportion of fetuses correctly diagnosed) and its 95% CIs were presented for ultrasonography and iuMRI in the primary analysis population by gestational age (i.e. 18–23 weeks and ≥ 24 weeks). The differences in diagnostic accuracy between ultrasonography and iuMRI and the 95% CIs were presented along with associated p -values calculated using McNemar's test.

Missing data

To examine the effect of missing data on the primary analysis, two sensitivity analyses were undertaken on the subset of fetuses whose iuMRI was undertaken within 2 weeks of their ultrasonography. The first sensitivity analysis was to assume that outcome data are missing at random,⁵⁶ which means that the reason for missing data can be predicted based on measurable covariates. A bivariate Probit analysis was undertaken, in which the two outcomes (ultrasonography correct, iuMRI correct) were modelled with the same covariates for both outcomes but with their coefficients allowed to differ. The covariates were chosen based on their association with missingness and diagnostic accuracy. The predicted probabilities of ultrasonography and iuMRI being correct, conditional on their covariates, were then derived for fetuses both with and without ORD; these predictions are unbiased under the assumption of data missing at random.⁵⁷ The difference between ultrasonography and iuMRI was calculated as the difference between the two predicted probabilities, with a 95% CI derived using bootstrap methods with 10,000 simulations. The second sensitivity analysis estimated the impact of missing data worst-case assumptions about the diagnostic ability of iuMRI. The differences in diagnostic accuracy between ultrasonography and iuMRI were calculated under both assumptions.

Secondary analysis of agreement

Repeat in utero magnetic resonance imaging

The primary analysis was repeated using repeat iuMRI scans. When mothers had more than one repeat scan, the latest iuMRI report was used.

Anatomical subgroup analysis

The primary analysis was repeated on three subgroups of the primary analysis population. These subgroups were defined by the diagnosis at ultrasonography.

Definition of subgroups

Isolated ventriculomegaly

Cases were included in this group if they were diagnosed with isolated VM as the only finding on ultrasonography. No other brain abnormality was diagnosed on ultrasonography in this subgroup. These cases were further categorised into three types based on the largest trigone on ultrasonography: mild (10–12 mm), moderate (13–15 mm) and severe (≥ 16 mm).

Posterior fossa

Cases were included in this group if they were diagnosed with abnormalities confined to the posterior fossa (with or without associated VM) on ultrasonography. These cases were further divided into two specific diagnosis categories: those involving the brain stem or cerebellum (i.e. parenchymal abnormalities including Dandy–Walker spectrum malformations, Chiari II malformation and cerebellar hypoplasia) and those involving cerebrospinal fluid (CSF)-containing lesions (i.e. enlarged cisterna magna, Blake's pouch cyst and arachnoid cysts). In the first 'overall' group, correctness was judged by the presence of any diagnosis of isolated posterior fossa abnormality on ORD. For a diagnosis in the two subsequent groups to be deemed correct, an ORD included the same specific diagnosis described above.

Failed commissuration

Fetuses were included in this group if they were diagnosed with either agenesis or hypogenesis of the corpus callosum (CC) as the sole diagnosis or in conjunction with VM on ultrasonography only. These fetuses were further categorised into their specific diagnosis category of 'agenesis of the CC' or 'hypogenesis of the CC'. In the first 'overall' group, correctness was judged only by either the presence of CC abnormality or the absence of another brain lesion. For a diagnosis in the two subsequent groups to be deemed correct, an ORD included the same specific diagnosis (i.e. 'agenesis of the CC' or 'hypogenesis of the CC'), regardless of any other brain lesion.

Subgroup analysis

Fetuses were categorised into one of the three anatomical subgroups described above if they were diagnosed on ultrasonography. The number of fetuses diagnosed within these groups for whom the diagnoses were correct on ORD were used to calculate the diagnostic accuracy of ultrasonography. The diagnostic accuracy of iuMRI was calculated on the same subgroup as defined using ultrasonography diagnosis.

Diagnostic accuracy was also calculated for the subgroups within each of the three anatomical subgroups. The number of incorrectly diagnosed cases on ultrasonography and iuMRI were presented along with the difference in diagnostic accuracy and its associated 95% CI and *p*-value from McNemar's paired test.

Diagnostic confidence

Diagnostic confidence was defined for each diagnosis on a five-point scale (10% = very unsure to 90% = highly confident) on ultrasonography. Diagnostic confidence by iuMRI is similar but has one addition ('diagnosis excluded'). Ultrasonographers were also asked which method of ultrasonography was used and which factors contributed to low confidence in the diagnosis. High confidence was defined as a score of 70% or 90% and low confidence was defined as a score of 10%, 30% or 50%.

The numbers of missed diagnoses (false negatives) have been presented by confidence category following ultrasonography alone and ultrasonography plus MRI for the primary population. Changes in confidence were categorised into five categories: no change, ultrasonography with more confidence by one category, ultrasonography with more confidence by two or more categories, MRI with more confidence by one category and MRI with more confidence by two or more categories. The number and percentage of fetuses in each category were presented. A Wilcoxon signed-rank test was carried out to compare the difference in confidence on ultrasonography with the difference in confidence on MRI.

Diagnostic accuracy and diagnostic confidence

A score-based weighted-average analysis was conducted⁵⁸ that combined changes in diagnostic accuracy, diagnostic confidence and management to provide a summary measure of the clinical impact attributable to iuMRI. The scores were defined for each individual diagnosis by its presence/absence and the assigned diagnostic confidence. The scores for the overall impact (i.e. at the level of the fetus rather than the condition) were assigned by the expert panel on a case-by-case basis. The null hypothesis, that the impact scores are zero (no impact), was tested by a one-sample *t*-test in the primary analysis population. The mean, standard deviation (SD) and 95% CI of difference in confidence was presented overall and also by anatomical subgroup.

Clinical impact

Change in prognosis

The prognosis was defined at the pregnancy (mother) level, rather than the fetal level. Following ultrasonography, the prognosis was recorded by using a four-point scale of severity (i.e. poor, intermediate, favourable and normal) and if the ultrasonographer was unable to offer a prognosis on the basis of the scan this was recorded as not known. The same process was undertaken following iuMRI and the two were compared in a 5 × 5 contingency table, with the number and percentage of pregnancies in each group presented.

When a fetus had a known prognosis for both ultrasonography and ultrasonography plus iuMRI, the change in prognosis category was categorised as no change, more favourable with ultrasonography by one category, more favourable with ultrasonography by two or more categories, more favourable with iuMRI by one category or more favourable with iuMRI by two or more categories. The number and percentage of fetuses in each of these categories were presented and their distribution compared using a Wilcoxon signed-rank test.

When a prognosis was unknown on either ultrasonography or ultrasonography plus iuMRI, these fetuses were presented separately and further grouped into nine categories: no change, not known to known (i.e. poor, intermediate, favourable or normal) or known (poor, intermediate, favourable or normal) to not known.

Counselling and management of pregnancy

In 2012, the study opened to recruitment in Belfast. Under Northern Irish law, TOP is permissible only in very specific circumstances that are more limited than in the rest of the UK or Western Europe. The incidence of TOP and the grounds for it being performed are, therefore, likely to be very different between Belfast and other centres, and findings in Belfast would not be expected to generalise to other study centres. Therefore, the main summaries of counselling and management included all mothers who completed iuMRI but excluded patients recruited from Belfast.

Following both ultrasonography and MRI, the clinician was asked to record if they discussed TOP with the mother, offered TOP to the mother, or neither. If TOP was offered, the clinician recorded whether or not this was due to there being a substantial risk of disability. The clinician was also asked whether or not MRI changed their counselling from ultrasonography only to ultrasonography plus iuMRI. A cross-tabulation of counselling categories (i.e. TOP not discussed, TOP discussed but not offered, TOP discussed and offered, TOP offered based on substantial risk of disability) on ultrasonography and counselling on ultrasonography plus iuMRI has been presented in *Appendix 9*.

Following iuMRI, the chosen management plan was recorded (i.e. continued pregnancy, continued pregnancy with further imaging, fetal demise, termination, unable to decide). If further imaging was used, the type was recorded (i.e. ultrasonography, iuMRI or both) and a rating of iuMRI's contribution was recorded (i.e. none, minor, significant, major, decisive). The number and percentage of mothers in each category has been presented.

Magnetic resonance imaging errors

The number of iuMRI errors and the total number of scans were presented by central and non-central reporters for the primary analysis population. The overall error rate (number of errors/total number of scans) was calculated by central and non-central reporters along with the difference in error rate. Errors were also assessed in relation to their chronological order within the study (i.e. errors within the first 25 scans, errors in scans 26–50, errors in scans 51–75 and errors thereafter).

Patient acceptability

Participants were invited to complete two surveys to evaluate their satisfaction with the care received. The methods and results relating to this are outlined in *Chapter 4*.

Patient outcomes

Outcome of pregnancy

The outcome of pregnancy was categorised and the number and percentage of fetuses that were categorised as livebirth, perinatal, neonatal, infant death, stillbirth, spontaneous intrauterine fetal demise or terminated have been presented by each of the four analysis populations described in *Statistical analysis*.

The Hospital Anxiety and Depression Scale questionnaire

The Hospital Anxiety and Depression Scale (HADS) questionnaire was used to measure anxiety and depression, each being scored from 0 to 21, with higher scores indicating greater severity.⁵⁹ Prorating was used for partially answered questionnaires, provided that at least four of the seven questions within the domain had been answered. The change from survey 1 to survey 2 was calculated among mothers who completed both surveys and the difference was tested using the Wilcoxon signed-rank test.

Mothers were also classified separately for anxiety and depression as 'in the normal range' (0–7.5), 'cause for concern' (7.51–10.5) or 'probable clinical case' (> 10.5). The change in response categorisations from survey 1 to survey 2 was analysed using McNemar's test generalised to a three-point scale (also known as Bowker's test for symmetry).

Mean, SD, median, interquartile range (IQR) and minimum and maximum values were presented for anxiety and depression scores from each survey. The number and percentage of mothers in each category were also presented. When mothers completed both surveys, the mean, SD, median and IQR of the difference between scores was presented. The difference in the percentage by category for those completing both surveys was also presented.

Additional tests and resource use

When additional tests were used, the number and percentage of each type of test was presented for the whole cohort. The median and IQR of additional tests per mother were presented. The impact of these additional tests was presented by category.

A summary of clinic visits has not been presented because of inconsistent reporting; however, full resource use has been reported as part of the health economic evaluation.

Ultrasonography technical factors

The type of ultrasonography used was recorded (i.e. two or three dimensional) and the number and percentage of mothers were presented. Technical factors affecting the confidence of the sonographer (e.g. high body mass index of the patient, fetal position, oligohydramnios or other) were also reported.

Safety analysis

Adverse events (AEs) and serious adverse events (SAEs) were reported separately for all mothers and fetuses.

The number and percentage of mothers and fetuses experiencing an AE have been presented separately. Owing to the small number of AEs, they were reported descriptively by the participants. AEs related to the iuMRI procedure were reported separately. The AEs have also been presented by the type of AE and whether or not it was attributable to the procedure. The same summaries have been presented for SAEs. Descriptions of all AEs and SAEs have been presented in *Table 20*.

Missing data

When data have not been recorded, 'not reported' has been included in relevant summary tables throughout the report (see *Tables 15* and *16*).

Health economic methods

Background

Introducing iuMRI to diagnose fetal brain abnormalities during pregnancy will introduce additional direct costs to the NHS. However, iuMRI may also lead to cost savings in other areas and may improve outcomes. Therefore, there is a need for health economic analysis to assess the cost-effectiveness of iuMRI in the diagnosis of fetal brain abnormalities.

Overview

This section considers an economic evaluation based on data collected as part of the MERIDIAN study. Information on resource use and outcomes is analysed for participants in the MERIDIAN study, to estimate the total additional costs and the change in outcomes associated with iuMRI in the diagnosis of fetal brain abnormalities.

A health economic analysis involves the comparison of two or more alternative courses of action. One of the comparators should be usual care, so that it can be determined whether or not the new intervention represents good value for money compared with current practice. The MERIDIAN study is not a randomised controlled trial with an intervention and a comparator, but instead it reports the diagnostic accuracy of iuMRI. In the base case, the economic analysis compares the costs and outcomes associated with iuMRI with those of single ultrasonography. In scenario analyses, iuMRI is compared with repeat ultrasonography.

Costs

The base-case analysis considers the costs incurred until the end of pregnancy. This includes costs of delivery for continued pregnancies and costs of termination for terminated pregnancies. Scenario analyses consider the exclusion of costs of delivery and termination.

In the base-case analysis, resource use is calculated from the analysis of MERIDIAN study data. Costs in the analyses include costs of ultrasonography scans, iuMRI scans, consultations, further investigations and TOP. Resource use is multiplied by unit costs to calculate total costs. Unit costs considered in the base case are shown in *Appendix 1*. All costs are reported for 2016; when necessary, costs were inflated using the *Unit Costs of Health and Social Care 2016* by Curtis and Burns.⁶⁰

There are several NHS costs available for TOP.⁶¹ In almost all cases within the study (96.1%), gestational age at termination was > 20 weeks, so 'Medical or surgical termination of pregnancy, over 20 weeks' gestation' (MA20Z) appears to be the most appropriate classification. Clinicians advised that terminations would be planned and so elective inpatient costs are appropriate. Clinicians further advised that the majority of women would stay in hospital for 2 days, with 10% of women staying in hospital for longer and none staying for < 2 days. The elective inpatient cost of £1232 per patient corresponds to an average length of stay of 1.26 days, so additional excess bed-days were included (costed at £277 per day), such that the total average stay was 2.1 days, costing £1581 per patient. From the NHS costs it is unclear whether or not the total costs include feticide, which clinicians advised would be required for gestational ages of > 21⁺⁶ weeks.⁶¹ Clinicians advised that the cost of feticide would be similar to the cost of a specialised fetal diagnostic procedure, which is £366 on average per procedure (NZ72Z). The NHS reference cost for an elective inpatient medical termination at 14–20 weeks' gestation (MA18D) would not include feticide and is £322 less than the equivalent cost category for > 20 weeks' gestation.⁶¹ This suggests that the cost for termination at > 20 weeks' gestation already includes a feticide cost (at least for a proportion of admissions). Therefore, we used our cost for termination at 20 weeks' gestation for a 2.1-day stay for terminations that involved a feticide. For validation, clinicians advised that this cost should be similar to the cost of a normal delivery, which costs an average of £1643 per procedure (NZ30C0).⁶¹

For terminations that did not require feticide, we assumed that a 2.1-day stay cost for 14–20 weeks' gestation was appropriate, which was calculated as £1136.

Scenario analysis additionally considers the inclusion of costs borne by participants, such as travel (valued at £0.45 per mile as per UK Government expenses⁶²) and time (valued using earnings by age taken from the 2016 Annual Survey of Hours and Earnings⁶³). A further scenario analysis includes the long-term costs associated with caring for children born with disabilities.

Costs for the pathway with iuMRI and with ultrasonography alone are reported separately for participants who had and for those who did not have TOP. Resource use differs for participants who have TOP, there is the cost of TOP itself, and further imaging or consultation costs may differ. The proportion of participants

having a TOP is required to calculate the total cost of iuMRI and ultrasonography. The proportion of participants having a TOP following iuMRI was observed, but because the MERIDIAN study was not a comparative trial, the proportion of TOP for ultrasonography alone was not directly available and had to be estimated (see *Analyses*).

Outcomes

Health economic analysis from an NHS perspective typically uses the quality-adjusted life-year (QALY) as the measure of benefit. However, the relevance of the QALY in iuMRI is questionable as decision-makers may be uncomfortable with the use of QALYs to reflect the value of an unborn child. Furthermore, incorporating QALYs for unborn children may lead to the perverse situation whereby the intervention with the lowest rate of true-negative detections maximises QALY gain.

In utero magnetic resonance imaging is intended to provide information, additional to ultrasonography, that is of value to parents and treating clinicians. This additional information is of value only if it is both accurate and leads to changes in the management of pregnancy. Therefore, the primary outcome selected for the economic analysis is 'appropriate management decisions'. The decision for an individual mother is classified as appropriate if the revised decision is consistent with the presence or otherwise of neuro-developmental abnormalities at birth or postmortem. Therefore, appropriate management decisions are:

- TOP advised following ultrasonography, continuation of pregnancy advised following iuMRI. Continuation of pregnancy chosen and infant exhibits no abnormalities at birth.
- TOP advised following ultrasonography, continuation of pregnancy advised following iuMRI. TOP chosen and the fetus exhibits no abnormalities at postmortem.
- Continuation of pregnancy advised following ultrasonography, TOP advised following iuMRI. Continuation of pregnancy chosen and infant exhibits abnormalities at birth.
- Continuation of pregnancy advised following ultrasonography, TOP advised following iuMRI. TOP chosen and the fetus exhibits abnormalities at postmortem.

In the base-case and scenario analyses, the reported cost-effectiveness estimate is the cost per management decision appropriately revised. The proportion of management decisions appropriately revised is calculated by cross-tabulating the number of cases in which iuMRI was correct with the number of cases in which iuMRI changed the diagnosis. The proportion of management decisions appropriately revised is equal to the number of cases in which iuMRI was correct and iuMRI changed the diagnosis, divided by the total number of cases for which both outcomes were available.

Analyses

All analyses were conducted using Stata to calculate resource use and outcomes, and data were analysed in Microsoft Excel® (Microsoft Corporation, Redmond, WA, USA) to calculate total costs and cost-effectiveness. All analyses used completed cases with the exception of the imputed missing data, which used bootstrapping in Stata. The analyses conducted are summarised in *Table 1* and explained further in the sections that follow.

Base case, scenario 1 and scenario 2

The MERIDIAN study recorded if participants were offered a TOP following the ultrasonography, if participants were offered a TOP following the iuMRI and if participants had a TOP. In the base-case analysis, the proportion of participants who had a TOP is used for iuMRI and the proportion of participants who had a TOP following ultrasonography alone is assumed to be equal to the proportion of participants who were offered TOP following ultrasonography. This approach may overestimate the proportion of participants having a TOP following ultrasonography alone. After iuMRI, 189 out of 537 participants were offered a TOP but only 58 participants had a TOP. Therefore, two further scenarios are considered. In scenario 1, the proportion of participants having TOP following iuMRI is assumed to be the proportion of participants who were offered TOP following iuMRI. In scenario 2, the proportion of participants having TOP following ultrasonography is the proportion of participants offered TOP following ultrasonography adjusted by the proportion of those who were offered TOP following iuMRI who had a TOP.

TABLE 1 Summary of analyses

Scenario	Population	iuMRI: proportion of TOP	Ultrasonography: proportion of TOP	Effectiveness data	Costs included
Base case	All women with outcomes	Proportion of women who had TOP	Proportion of women offered TOP after ultrasonography	Proportion of cases for which iuMRI was correct and changed the prognosis	NHS costs only
Scenario 1	Same as base case	Proportion of women offered TOP after iuMRI	Same as base case	Same as base case	Same as base case
Scenario 2	Same as base case	Same as base case	Proportion of women offered TOP after ultrasonography, multiplied by the proportion of women offered TOP after iuMRI who had TOP	Same as base case	Same as base case
Inclusion of participant-borne costs 1	Same as base case	Same as base case	Same as base case	Same as base case	NHS costs plus costs for time and mileage
Inclusion of participant-borne costs 2	All women with outcomes whose total mileage was < 40 miles	Same as base case	Same as base case	Same as base case	NHS costs plus costs for time and mileage
Gestational age at iuMRI of < 22 weeks	Gestational age of < 22 weeks	Same as base case	Same as base case	Same as base case	Same as base case
Gestational age at iuMRI of 22–24 weeks	Gestational age of 22–24 weeks	Same as base case	Same as base case	Same as base case	Same as base case
Gestational age at iuMRI of > 24 weeks	Gestational age of > 24 weeks	Same as base case	Same as base case	Same as base case	Same as base case
Comparison with the second ultrasonography	All women with outcomes	Same as base case	Adjusted to account for a proportion of cases in which ultrasonography identified the same information as iuMRI, and a proportion of cases for which TOP was offered (after iuMRI) and occurred	Proportion of cases in which iuMRI was correct and changed the prognosis, multiplied by the proportion of cases in which additional iuMRI provided information that was not visible on follow-up ultrasonography	Ultrasonography arm patients: additional cost for one two-dimensional ultrasonography and one subsequent consultation
Delivery costs excluded	Base case, scenario 1 and scenario 2	Base case, scenario 1 and scenario 2	Base case, scenario 1 and scenario 2	Base case, scenario 1 and scenario 2	Same as base case but delivery costs excluded. Termination costs included
Delivery and termination costs excluded	Base case, scenario 1 and scenario 2	Base case, scenario 1 and scenario 2	Base case, scenario 1 and scenario 2	Base case, scenario 1 and scenario 2	Same as base case but delivery and termination costs excluded
Increased cost for termination	Base case, scenario 1 and scenario 2	Base case, scenario 1 and scenario 2	Base case, scenario 1 and scenario 2	Base case, scenario 1 and scenario 2	Same as base case, termination costs increased

Deterministic analysis and probabilistic sensitivity analysis were conducted for the base case and for the two scenarios. The deterministic analysis reports total costs, incremental cost and incremental cost per management decision appropriately revised. The probabilistic sensitivity analysis samples each uncertain value from its associated probability distribution (i.e. normal for mean costs, log-normal for resource use, beta for proportions; see *Appendix 2*) simultaneously and reports the probabilistic mean cost per management decision appropriately revised. Cost-effectiveness acceptability curves are presented for a range of willingness-to-pay thresholds for appropriate management decisions. The cost-effectiveness acceptability curves are calculated from the proportion of simulations for which the cost per management decision is below each willingness-to-pay threshold and summarise the uncertainty associated with the cost-effectiveness results. The expected value of perfect information (EVPI) analysis is conducted on a per-person basis for varying willingness-to-pay thresholds. The EVPI is the price we would be willing to pay for perfect information about all of the factors that determine which intervention is cost-effective. This is calculated from the difference in monetary benefit when the choice between interventions is made on current information and when the choice is made based on perfect information with no uncertainty. EVPI is calculated at each willingness-to-pay threshold (per management decision appropriately changed, from £0 to £10,000):

- Each uncertain parameter is sampled from its associated distribution – this is called ‘one run’.
- For each run, the following are calculated:
 - the net benefit of iuMRI (appropriate management decisions are valued using the willingness-to-pay threshold and the cost of the iuMRI strategy, for TOP and continued pregnancies)
 - the net benefit of ultrasonography
 - the net benefit with perfect information (the maximum of the net benefit of iuMRI and the net benefit of ultrasonography).
- The mean over all the runs is calculated for:
 - the net benefit of iuMRI
 - the net benefit of ultrasonography
 - the net benefit with perfect information.
- The expected outcome with current information is calculated as:
 - the maximum of the mean net benefit of iuMRI and the mean net benefit of ultrasonography.
- The EVPI is calculated as:
 - the difference between the mean net benefit with perfect information and the expected outcome with current information.

Some participant records were missing data on some elements of resource use. In order to include as many participants as possible in the analysis, imputed missing data were analysed. The imputation used multivariate normal regression with 75 replicate data sets with age (years), distance from iuMRI unit, the total number of fetal unit visits and gestational age at iuMRI as predictors of missing costs. Imputed missing data are analysed for the base case, scenario 1 and scenario 2.

Inclusion of participant-borne costs

Two scenarios considered participant-borne costs, including costs for mileage and costs for participant time to travel to and undergo an iuMRI procedure. The first analysis considered the whole population. The second analysis considered only patients whose mileage was < 40 miles to estimate the potential results of reconfiguring services so that iuMRI could be locally available. Outcome data were the same as in the base case.

Subgroups by gestational age

Subgroup analyses separately considered participants with a gestational age at iuMRI of < 22 weeks, 22–24 weeks and > 24 weeks. Total costs, the proportion of TOPs and the proportion of management decisions correctly revised following iuMRI are analysed separately for the subgroups.

Comparison with second ultrasonography

This analysis compares iuMRI with repeat ultrasonography, as recommended by the Royal College of Obstetricians and Gynaecologists (RCOG) pathway. In the RCOG pathway, if an abnormality is suspected, the woman should be referred for a second opinion as soon as possible, ideally within 3 working days.⁶⁴ As described in *Data collection and management*, after an iuMRI examination, participants attended a follow-up appointment that included a follow-up ultrasonography to determine whether or not the additional information from iuMRI was visible on follow-up ultrasound scans.

In this scenario, the proportion of management decisions appropriately revised following iuMRI is adjusted to account for the improved diagnostic accuracy of a second ultrasonography. This downgrades the effect of iuMRI to take into account the proportion of abnormalities that would have been detected by a repeat ultrasonography. The proportion of management decisions appropriately revised is multiplied by the proportion of cases in which the additional diagnostic information obtained from iuMRI was not visible on follow-up ultrasonography. The proportion of participants having a TOP following a second ultrasonography is calculated using the following approach:

1. Cross-tabulating the number of cases when TOP was offered following ultrasonography by the number of cases when TOP was offered following iuMRI to calculate the number of cases in which the management decisions agreed or disagreed.
2. Multiplying the number of cases when ultrasonography and iuMRI do not agree by the proportion of cases when the new information from iuMRI was not visible on follow-up ultrasonography, to calculate the proportion of cases when the second ultrasonography and iuMRI would agree or disagree.
3. Calculating the proportion of cases when TOP would be offered following the second ultrasonography.
4. Multiplying the proportion of cases when TOP would be offered following the second ultrasonography by the proportion of cases for which TOP was offered following iuMRI in which TOP happened.

In this scenario, the same costs for the iuMRI and ultrasonography pathway (whether or not the participant had a TOP) are used, but there is a cost for one additional two-dimensional ultrasonography and consultation for all participants in the second ultrasonography pathway.

Longer time horizon

This scenario considers costs over a longer time period, to include the cost of additional care for children born with disabilities. This analysis uses the proportions of children with abnormalities, borderline abnormalities and no abnormalities from the follow-up study to estimate the long-term costs associated with continued pregnancies. In the follow-up study, at 2–3 years, 31% of children had abnormalities, 14% of children had borderline abnormalities and 56% of children had no abnormalities (see *Table 38*).

A published study of the economic costs of childhood psychiatric disorders reports the mean public sector costs over the previous year of life for children with and without psychiatric disorders by level of cognitive impairment.⁶⁵ The costs were estimated by collecting resource use data from parents when the child was 11 years of age. The annual costs for moderate and severe impairment without psychiatric disorders were £5529 and £5842, respectively, inflated to 2016 prices.⁶⁰ Our analysis assumed that these costs would be applicable each year up until the age of 11 years, and so considered long-term cost for 11 years. Beyond the age of 11 years, it is anticipated that the costs could change substantially as children enter adolescence and their needs change. Children born with no abnormalities are assumed to incur no additional costs, children with abnormalities incur the costs associated with severe cognitive impairment and children with borderline abnormalities incur the costs associated with moderate cognitive impairment. Costs are discounted at 3.5% per annum.⁶⁶

Costing scenarios

The following three scenarios consider changes to the costs included in the analysis. These were applied to the base case, scenario 1 and scenario 2 to explore the sensitivity of the variation in results to the costing assumptions.

Delivery costs excluded

This scenario considers the same costs as the base case, except that it excludes the costs for the delivery of continued pregnancies.

Delivery and termination costs excluded

This scenario considers the same costs as the base case, except that it excludes the costs for TOP and the costs for the delivery of continued pregnancies.

Increased termination cost

This scenario considers a higher cost of termination, assuming that the cost for termination at 20 weeks' gestation for a 2.1-day stay did not involve feticide and that feticide would incur an additional cost of £366. In this scenario, the cost for a termination without feticide was £1581 and the cost for a termination with feticide was £1947.

Patient and public involvement

Representatives from Antenatal Results and Choices (ARC) and the Spina Bifida, Hydrocephalus, Information, Networking, Equality (SHINE) charities were involved in the study oversight throughout the project through the Trial Steering Committee (TSC). This included the review and development of the study protocol and patient documents, monitoring the study progress and review and discussion of the final results of the study. Feedback from the patient and public involvement (PPI) members informed our approach to potential participants and the content of the participant information sheets. The PPI members also had input in the content of the results summary/participant debrief letter and the method for disseminating results to participants.

Study governance and management

Study governance

The study conduct was governed by a number of oversight committees: the Trial Management Group (TMG), the TSC and the Data Monitoring and Ethics Committee (DMEC). The trial was conducted in accordance with Sheffield CTRU standard operating procedures (SOPs) and the committees convened at appropriate intervals as dictated by both study requirements and SOPs.

The TMG was composed of the chief investigator (chairperson), study collaborators and key staff within the CTRU and Newcastle University. The TMG met monthly or bimonthly via teleconference during trial recruitment and follow-up.

The TSC members were approved and appointed by the funder. The TSC was composed of a professor of fetal and maternal medicine as the independent chairperson, an independent consultant in neuroradiology, experts in antenatal screening and supporting families who provided expert advice and acted as PPI representatives. The DMEC consisted of an independent statistician, a professor of health-care research (chairperson) and a fetal medicine consultant. The TSC received formal recommendations from the DMEC.

Ethics arrangements and regulatory approvals

The study and subsequent amendments were approved by the Yorkshire and The Humber – South Yorkshire Research Ethics Committee (REC) (reference number 11/YH/0006). Each participating site gave UK NHS Research and Development approval. The study was conducted in accordance with the principles of good clinical practice.

Protocol amendments after study initiation

Details of substantial amendments submitted to the REC, which were important changes to trial methodology, are listed below.

Multidisciplinary independent expert panel

The original protocol stated that agreement between the PND and the outcome diagnosis would be judged independently by two fetal medicine experts and that a third expert would arbitrate if there was a discrepancy in opinion. In June 2011, the protocol was revised to state that an independent expert panel would be appointed to assess agreement rather than two independent experts. The independent expert panel was to consist of a fetal medicine clinician, a paediatric neuroradiologist and a paediatric neurologist or neurosurgeon.

Outcome data collection

A number of measures were introduced to assist with the collection of follow-up data about the outcome of the pregnancy. First, a data-collection sheet for handheld notes was developed; this sheet requested basic details of where and when the participant delivered. This meant that the research midwife could be notified of the delivery and could subsequently request the appropriate outcome data.

We also implemented the option of a postmortem MRI. When a woman opted for TOP, or when there was a neonatal death, we aimed to gather the results of any postmortem autopsies performed for clinical purposes. However, the offer of a postmortem is often declined; therefore, we introduced the option of a non-invasive postmortem MRI examination that could be performed in addition to, or instead of, a postmortem autopsy.

Following feedback from sites where, in some cases (e.g. mild VM that resolved during pregnancy), clinicians did not advocate postnatal imaging, we modified our ORD to include a third-trimester ultrasonography in cases of resolved VM.

Recruitment target

The recruitment target was changed on the advice of the DMEC and TSC from 750 participants to 'at least 750 participants' to ensure that the required 504 complete cases were obtained.

Chapter 3 Study results

Recruitment and participant flow

Between July 2011 and August 2014, 1101 women carrying 1109 fetuses were identified as potentially eligible. Of those, 198 women (198 fetuses) were deemed to be ineligible on further screening or declined participation, resulting in 903 participants (911 fetuses) being recruited. A total of 80 participants (82 fetuses) did not complete ultrasonography and iuMRI (see *Figure 1*). A total of 64% of all participants attended the Academic Unit of Radiology at the University of Sheffield for their iuMRI and the remaining 36% of participants attended one of the five collaborating centres. ORD was available in 638 out of 829 (77%) fetuses, of which 570 out of 638 (89%) had the iuMRI performed within 2 weeks of the referral ultrasonography. A total of 369 fetuses (65%) were in the 18–23 weeks' gestational age group (110% of required) and 201 fetuses (35%) were in the ≤ 24 weeks' gestational age group (120% of required). The three commonest ultrasonography diagnoses were isolated VM (306/570, 54%), an abnormality restricted to the contents of the posterior fossa (81/570, 14%) and failed commissuration (79/570, 14%).

Of the 1101 potential participants, 198 women were ineligible or declined participation. A total of 86 of these potential participants were not asked to be part of the study. The most common reason for this was that the consultant considered it to be clinically inappropriate or unhelpful ($n = 50$, 58%). A total of 112 potential participants were offered to participate but declined. There were several reasons for refusal that were all reported with a similar frequency [i.e. decision had already been made ($n = 18$), unwilling to travel ($n = 13$), does not want iuMRI ($n = 28$) and too upset or distressed ($n = 11$)]. A full summary of these reasons can be found in *Appendix 3*.

Characteristics

Participant characteristics are shown in *Table 2*. The majority of participants completed their iuMRI in the host institution (Sheffield, 67% of the 'ORD available cohort'). Belfast had a large number of participants who were excluded ($n = 14$, 17%) because of the time delay between the ultrasonography referral and the iuMRI. The median maternal age of the 'ORD available' cohort was 29 years (IQR 24–33 years). The median gestational age of the 'ORD available' cohort was 22.5 weeks (IQR 21.4–26.5 weeks). A total of 95% of the 'ORD available' cohort were singleton pregnancies. The majority (56%) of this cohort had a dominant diagnosis of VM.

Uptake of in utero magnetic resonance imaging

The iuMRI was successful in and diagnostic images were obtained from 823 participants (829 fetuses). Second iuMRI studies were performed on 97 participants and seven of these had a third iuMRI study. All follow-up studies were successful and useful diagnostic images were obtained.

Uptake of ultrasonography

A total of 681 mothers had two-dimensional ultrasonography and 140 mothers had both two-dimensional and three-dimensional ultrasonography. Technical factors experienced during ultrasonography are summarised in *Appendix 4*.

TABLE 2 Characteristics of mothers and fetuses who were recruited to the MERIDIAN study and included in the primary analysis

Characteristic	ORD available (N = 570)	ORD unavailable (N = 175)	Excluded ^a (N = 84)
iuMRI site, n (%)			
Sheffield	380 (67.0)	120 (69.0)	31 (37.0)
Birmingham	75 (13.0)	34 (19.0)	15 (18.0)
Newcastle	66 (12.0)	6 (3.4)	9 (11.0)
Leeds	34 (6.0)	12 (6.9)	11 (13.0)
Nottingham	12 (2.1)	3 (1.7)	1 (1.2)
Belfast	3 (0.5)	0 (0.0)	17 (20.0)
Maternal age at approach (years), median (IQR)	29 (24–33)	28 (24–33)	30 (25–35)
Gestational age at iuMRI (weeks), median (IQR)	22.5 (21.4–26.5)	22.0 (21.2–24.7)	24.8 (23.1–28.5)
Singleton pregnancy, n (%)	539 (94.6)	165 (94.3)	80 (96.4)
Dominant diagnosis on ultrasonography, n (%)			
Isolated VM	321 (56.0)	33 (19.0)	51 (63.0)
ACC	89 (16.0)	51 (29.0)	9 (11.0)
Posterior fossa	81 (14.0)	44 (25.0)	9 (11.0)
Other	79 (14.0)	47 (27.0)	12 (15.0)

ACC, agenesis of the corpus callosum.

^a Includes participants who completed iuMRI > 2 weeks after referral and participants who were withdrawn after iuMRI.

Outcome of pregnancy

Of the 829 fetuses undergoing both ultrasonography and iuMRI, 642 (77%) were livebirths, 124 (15%) underwent TOP and 53 (6.4%) were categorised as stillbirths, miscarriages, intrauterine demise or deaths within the postnatal follow-up (*Table 3*). Data were not available for 10 (1.2%) fetuses.

Diagnostic source for outcome reference diagnosis

The most frequently used diagnostic source for ORD in the primary cohort was transcranial ultrasonography ($n = 261$, 46%; *Table 4*). Postnatal MRI was used in one-quarter of cases ($n = 144$, 25%).

Agreement

Overall

The overall diagnostic accuracies of ultrasonography and iuMRI were 68.1% and 93.0%, respectively, with a difference of 24.9% (95% CI 21.1% to 28.7%; *Table 5*). The difference between ultrasonography and iuMRI increased with gestational age: in the 18–23 week group, the diagnostic accuracies were 69.9% for ultrasonography and 92.4% for iuMRI (difference of 22.5%, 95% CI 17.8% to 27.2%); in the ≥ 24 week group, the diagnostic accuracies were 64.7% for ultrasonography and 94.0% for iuMRI (difference of 29.3%, 95% CI 22.6% to 36.1%).

TABLE 3 Outcome of pregnancy in 827 cases that completed iuMRI

Outcome of pregnancy	ORD available, <i>n</i> (%) (<i>N</i> = 570)	ORD unavailable, <i>n</i> (%) (<i>N</i> = 175)	Excluded, <i>n</i> (%) (<i>N</i> = 84)	Overall, <i>n</i> (%) (<i>N</i> = 829)
Livebirth	482 (84.6)	90 (51.4)	70 (83.3)	642 (77.4)
Stillbirth or miscarriage/fetal demise	20 (3.5)	27 (15.4)	6 (7.1)	53 (6.4)
Terminated	68 (11.9)	51 (29.2)	5 (6.0)	124 (15.0)
Not available	0 (0.0)	7 (4.0)	3 (3.6)	10 (1.2)

TABLE 4 Diagnostic source used for ORD in 570 cases in the primary analysis

Diagnostic source for ORD	<i>n</i> (%)
Transcranial ultrasonography	261 (46.0)
Postnatal MRI	144 (25.0)
Postmortem autopsy	70 (12.0)
Third-trimester ultrasonography	63 (11.0)
Postnatal computed tomography	19 (3.3)
Abnormality visible on postnatal inspection	11 (1.9)
Postmortem MRI	2 (0.4)

TABLE 5 Diagnostic accuracy of ultrasonography and iuMRI by gestational age in the 570 cases in the primary analysis population

Gestational age	Ultrasonography correct, <i>n</i> (%)	iuMRI correct, <i>n</i> (%)	Percentage difference (95% CI)	<i>p</i> -value ^a
18–23 weeks (<i>N</i> = 369)	258 (69.9)	341 (92.4)	22.5 (17.8 to 27.2)	< 0.0001
≥ 24 weeks (<i>N</i> = 201)	130 (64.7)	189 (94.0)	29.3 (22.6 to 36.1)	< 0.0001
Combined (<i>N</i> = 570)	388 (68.1)	530 (93.0)	24.9 (21.1 to 28.7)	< 0.0001

^a McNemar's test between ultrasonography and iuMRI.

In 386 out of 570 cases (67.7%), both the ultrasonography and iuMRI reports were correct, and in 144 out of 570 cases (25.3%), the ultrasonography report was incorrect but the iuMRI report was correct. There were two fetuses (0.4%) for whom the ultrasonography was correct and the iuMRI was incorrect and 38 (6.7%) fetuses for which both the ultrasonography and the iuMRI were incorrect.

Missing data

Of the 175 fetuses with no ORD, ultrasonography and iuMRI agreed in 86 cases and disagreed in 89 cases. In the most conservative assumption possible, ultrasonography would be correct in all 175 cases, with iuMRI incorrect in all cases of disagreement, resulting in a diagnostic accuracy for ultrasonography of 76.6% and for iuMRI of 82.7%. The difference of 7.1% (95% CI 3.0% to 11.2%), although vastly reduced, remains highly statistically significant ($p = 0.0007$). The improved diagnostic accuracy of iuMRI over ultrasonography was, therefore, not attributable to missing data.

A less extreme method of handling a missing ORD was to assume that it was missing at random and then to predict what the outcome would have been based on other, similar fetuses for which ORDs were available. The prediction model took into account the relationship between diagnostic accuracy, the probability of missing ORD data and the following characteristics: maternal age, fetal age (weeks) at time of iuMRI, single/multiple pregnancy, outcome (TOP/no TOP) and prognosis.

A missing ORD was more commonly encountered among individuals with a normal, poor or unknown prognosis, but was similar in relation to whether or not iuMRI changed the prognosis. The association of missing data with a prognosis was similar regardless of whether the prognosis was based on ultrasonography or iuMRI. As there were no missing data for the ultrasonography prognosis, it was used in the prediction model. Missing data were also more commonly observed among individuals who underwent TOP, although this was highly associated with the prognosis. Missing data were slightly more common among younger fetuses but showed little association with maternal age or single/multiple pregnancy. Because gestational age was associated with diagnostic accuracy (the diagnostic accuracy of ultrasonography was reduced in those of older gestational age), this was retained in the prediction model along with the prognosis based on ultrasonography.

Using the prediction model to predict missing outcomes (under the assumption of missing at random), the diagnostic accuracy of ultrasonography was 66.9% and iuMRI was 92.8%, resulting in a slightly greater difference of 25.9% (95% CI 22.3% to 29.7%).

All participants with outcome reference diagnosis

The primary analysis was repeated on all participants who had scans and an ORD (Table 6). Similar results to the primary analysis were observed and the overall diagnostic accuracies of ultrasonography and iuMRI, at 68.8% and 92.3%, respectively, resulted in a difference of 23.5% (95% CI 20.0% to 27.1%; $p < 0.0001$).

Repeat in utero magnetic resonance imaging

There were 65 fetuses for which repeat ultrasonography and iuMRI were undertaken and an ORD was obtained. As with the primary analysis, the majority (46/65) underwent iuMRI at the host institution. The median gestational age was 31 weeks with just five fetuses under 24 weeks. The overall diagnostic accuracies of ultrasonography and iuMRI were 67.7% and 87.7%, respectively, indicating a difference of 20.0% (95% CI 7.8% to 32.2%; $p < 0.0001$; Table 7).

A summary of the analysis repeated on first, single and repeat pregnancies can be found in Appendix 5.

TABLE 6 Diagnostic accuracy of ultrasonography and iuMRI by gestational age in the 638 cases who had scans and ORDs

Gestational age	Ultrasonography correct, <i>n</i> (%)	iuMRI correct, <i>n</i> (%)	Percentage difference (95% CI)	<i>p</i> -value ^a
18–23 weeks (<i>N</i> = 399)	279 (69.9)	365 (91.5)	21.6 (17.1 to 26.0)	< 0.0001
≥ 24 weeks (<i>N</i> = 239)	160 (66.9)	224 (93.7)	26.8 (20.6 to 32.9)	< 0.0001
Combined (<i>N</i> = 638)	439 (68.8)	589 (92.3)	23.5 (20.0 to 27.1)	< 0.0001

^a McNemar's test between ultrasonography and iuMRI.

TABLE 7 Diagnostic accuracy of ultrasonography and iuMRI in 65 cases with complete outcome data for which repeat iuMRI was used (the latest scan was used if more than two were completed)

Gestational age	Ultrasonography correct, n (%)	iuMRI correct, n (%)	Percentage difference (95% CI)	p-value ^a
18–23 weeks (N = 5)	4.0 (80.0)	4.0 (80.0)	0.0 (20.0 to 20.0)	N/A
≥ 24 weeks (N = 60)	40.0 (66.7)	53.0 (88.3)	21.6 (8.6 to 34.7)	< 0.0001
Repeat MRI (N = 65)	44.0 (67.7)	57.0 (87.7)	20.0 (7.8 to 32.2)	< 0.0001

N/A, not available.
^a McNemar's test between ultrasonography and iuMRI.

Anatomical subgroup analysis

Isolated ventriculomegaly

Ventriculomegaly was the most common brain abnormality for referral to the MERIDIAN study.⁶⁷ In 421 out of 570 (74%) fetuses, VM formed part of the ultrasonography diagnosis; in 306 out of 570 (54%) fetuses, VM was the only brain abnormality diagnosed on ultrasonography. This analysis is based on the 306 cases with isolated VM on ultrasonography.

A total of 199 out of 306 (65%) cases were in the 18–23 weeks group at the time of iuMRI and 107 out of 306 (35%) cases were in the ≥ 24 weeks group. The category of VM was determined by the largest trigone measurement and were divided into mild (10–12 mm), moderate (13–15 mm) and severe (≥ 16 mm) categories. Ultrasonography diagnosed mild VM in 244 out of 306 (80%) fetuses, moderate VM in 36 out of 306 (12%) fetuses and severe VM in 26 out of 306 (8%) fetuses.

The ORD found additional brain abnormalities other than VM in 31 out of 306 fetuses (10.1%); 17 out of 199 (8.5%) of those were in the 18–23 weeks group and 14 out of 107 (13.1%) of those were in the ≥ 24 weeks group. Of those fetuses, 5.7% (14/244) of those referred with mild VM had other brain abnormalities on ORD, 19.4% (7/36) of those referred with moderate VM had other brain abnormalities and 38.5% (10/26) of those referred with severe VM on ultrasonography showed other brain abnormalities on ORD (*Table 8*). The other brain abnormalities identified were agenesis of the CC (ACC) or hypogenesis of the CC in 17 out of 31 (55%) fetuses, encephalomalacia and/or intracranial haemorrhage in 8 out of 31 (26%) fetuses, cortical formation abnormalities in 4 out of 31 (13%) fetuses and absent cavum septum pellucidum (CSP) caused by septo-optic dysplasia in 2 out of 31 (6%) fetuses.

TABLE 8 Diagnostic accuracy of ultrasonography and iuMRI in cases diagnosed with isolated VM on ultrasonography

VM diagnosis	N	Other brain abnormality on ORD (n)	Diagnostic accuracy (%)	Other brain abnormality correct on iuMRI (n)	Diagnostic accuracy (%)	Difference (95% CI)	p-value ^a
All cases of isolated VM	306	31	89.9	27	98.7	8.8 (4.9 to 12.1)	< 0.0001
Mild	244	14	94.3	12	99.2	4.9 (1.3 to 7.8)	0.0034
Moderate	36	7	80.6	7	100.0	19.4 (3.7 to 35.2)	0.0156
Severe	26	10	61.5	8	92.3	30.8 (9.2 to 52.4)	0.0078

^a McNemar's test between ultrasonography and iuMRI.

The iuMRI was correct in 302 out of 306 cases when compared with the ORD. Of these 302 cases, 275 were correctly diagnosed with isolated VM and in 27 cases the iuMRI correctly identified additional brain abnormalities. The iuMRI provided an incorrect diagnosis in 4 out of 306 fetuses (three cases of hypogenesis of the CC and one case of absent CSP).

There were complete clinical impact data available for 295 out of 306 (96%) fetuses; of these, 29 fetuses had additional brain abnormalities. Fetal medicine clinicians reported that iuMRI provided additional information compared with ultrasonography in 106 out of 295 (35.9%) cases. Overall, there was a change in the prognosis category after iuMRI in 69 out of 295 (23.4%) cases. This was a change to a worse prognosis in 22 cases and to an improved prognosis in 47 cases.

Forty-two (14.2%) women were offered TOP based on the ultrasonography diagnosis alone and in eight of those cases the offer of TOP was reversed after iuMRI. After iuMRI, an additional 14 women were offered TOP based on the diagnosis of brain abnormalities other than VM. The decision to offer TOP, therefore, was changed in 22 out of 295 (7.5%) fetuses after iuMRI and TOP was performed in 14 out of 295 (4.7%) cases. The contribution of iuMRI to the final choice of management was rated by the fetal medicine expert to be 'of no value' in 41 out of 295 (14%) cases, minor in 177 out of 295 (60%) cases, significant in 64 out of 295 (22%) cases, major in 12 out of 295 (4%) and decisive in one (0.3%) case.

For the 29 fetuses that had brain pathology other than VM confirmed on ORD, the prognosis was considered worse in 13 out of 29 (44.8%) cases after the iuMRI and the effect of iuMRI on overall clinical management was rated as 'significant', 'major' or 'decisive' in 15 out of 29 (51.7%) cases.

Failed commissuration

Ultrasonography reported some form of 'failed commissuration' abnormality in 92 out of 570 (16.1%) referrals into the MERIDIAN study.⁶⁸ Thirteen of those referrals also had additional brain pathology other than VM identified on ultrasonography and were, therefore, not included in this subgroup analysis. The remaining 79 cases form the basis of this analysis, 54 of which had associated VM. The iuMRI was performed at 18–23 weeks of gestational age in 51 out of 79 (65%) fetuses and at ≥ 24 weeks of gestational age in 28 out of 79 (35%) fetuses. In total, 24 out of 79 cases were diagnosed with hypogenesis of the CC on ultrasonography and 55 out of 79 cases were diagnosed with ACC.

The diagnostic accuracy for 'isolated failed commissuration' as a whole group was 34.2% for ultrasonography, compared with 94.9% for iuMRI (difference of 60.7%, 95% CI 47.6% to 73.9%; $p < 0.0001$) (Table 9). Of the 52 fetuses whose diagnosis was incorrect on ultrasonography, the ORD recorded a normal CC in 46 fetuses and additional diagnoses not detected on ultrasonography in 12 fetuses (five fetuses had a normal CC but additional diagnoses).

TABLE 9 Diagnostic accuracy of ultrasonography and iuMRI in cases diagnosed with either ACC or hypogenesis of the CC on ultrasonography

Diagnosis	N	Ultrasonography		iuMRI		Difference (95% CI)	p-value ^a
		Incorrect (n)	Diagnostic accuracy (%)	Incorrect (n)	Diagnostic accuracy (%)		
Any malformation of the CC	79	52	34.2	4	94.9	60.7 (47.6 to 73.9)	< 0.0001
Hypogenesis of the CC	24	22	8.3	3	87.5	79.2 (55.1 to 103.3)	< 0.0001
ACC	55	33	40.0	4	92.7	52.7 (37.7 to 67.7)	< 0.0001

a McNemar's test between ultrasonography and iuMRI.

The iuMRI had an error rate of 4 out of 79 fetuses; all four of these errors were in relation to the CC diagnosis. All 12 other brain abnormalities shown on ORD were diagnosed correctly on iuMRI.

Further analysis looked at the diagnostic accuracy of ultrasonography and iuMRI in specifically diagnosing hypogenesis of the CC or ACC as discrete abnormalities, irrespective of the presence of other brain abnormalities (see *Table 9*).

Ultrasonography was shown to have a diagnostic accuracy of 8.3% for detecting hypogenesis of the CC. In comparison, iuMRI demonstrated 87.5% accuracy (difference of 79.2%, 95% CI 55.1% to 100%; $p < 0.0001$). The diagnostic accuracy for detecting ACC specifically was 40.0% for ultrasonography and 92.7% for iuMRI (difference of 52.7%, 95% CI 37.7% to 67.7%; $p < 0.0001$).

Ultrasonography correctly diagnosed hypogenesis of the CC in 2 out of 24 fetuses (8.3%). The majority of ultrasonography errors (20/24) for hypogenesis of the CC occurred when the ORD showed a normal CC, and in 2 out of 24 cases the ORD reported ACC.

Ultrasonography correctly diagnosed ACC in 22 out of 55 cases (40.0%). Of the remaining 33 cases, which ultrasonography diagnosed with ACC, the ORD reported hypogenesis of the CC in seven cases and a normal CC in 26 cases.

The iuMRI reported a normal CC in 45 out of 79 fetuses, 43 (95.5%) of which were subsequently confirmed as normal on ORD; the other two fetuses were incorrectly diagnosed by iuMRI and were found to have hypogenesis of the CC on ORD. In 8 out of 79 cases, the iuMRI reported hypogenesis of the CC, of which 6 (75%) were found to be correct. Finally, iuMRI diagnosed 26 cases of ACC, of which 23 (88.5%) were subsequently found to be correct.

Posterior fossa abnormalities

Ultrasonography reported some form of abnormality in the posterior fossa in 81 out of 570 (14.2%) fetuses.⁶⁹ The iuMRI was performed at 18–23 weeks of gestational age in 57 out of 81 (70%) fetuses and at ≥ 24 weeks of gestational age in 24 out of 81 (30%) fetuses. A total of 67 out of 81 cases were diagnosed with a parenchymal abnormality confined to the posterior fossa on ultrasonography and 14 out of 81 cases were diagnosed with a CSF-containing abnormality. Associated VM was found on ultrasonography in 25 out of 81 (31%) cases overall and in 3 out of 15 (20%) fetuses with cerebellar hypoplasia, 6 out of 21 (29%) fetuses with Dandy–Walker spectrum abnormalities and 11 out of 21 (52%) fetuses with Chiari type II malformations.

The overall diagnostic accuracy of ultrasonography was 65.4%. A normal brain was reported on ORD in 18 out of 28 of the ultrasonography errors and other supratentorial brain abnormalities were reported on in 10 out of 28 of the ultrasonography errors (the other brain abnormalities consisted of six ACC or hypogenesis of the CC cases, two cortical formation abnormalities and two acquired pathologies). The iuMRI gave a diagnostic accuracy of 87.7%, showing an improvement in diagnostic accuracy of 22.3% (95% CI 14.0% to 30.5%; $p < 0.0001$) over ultrasonography.

In the 10 iuMRI errors, eight were cases in which the ORD was normal and two were cases in which iuMRI failed to detect an abnormality of the CC shown on the ORD.

The specific diagnoses of parenchymal abnormalities on ultrasonography were:

- cerebellar hypoplasia in 25 out of 67 (38%) fetuses – vermian hypoplasia in 15 fetuses and hypoplasia of the cerebellar hemispheres in 10 fetuses
- Dandy–Walker spectrum abnormalities in 21 out of 67 (31%) fetuses
- Chiari type II malformation in 21 out of 67 (31%) fetuses.

The label 'parenchymal abnormality' in *Table 10* requires the exact pathological diagnosis of the posterior fossa abnormality to be correct when compared with the ORD and ignores the presence or absence of other brain abnormalities. The correct specific diagnosis was made by ultrasonography in 54 out of 67 fetuses, demonstrating a diagnostic accuracy of 80.6%. The iuMRI correctly diagnosed 64 out of 67 fetuses, demonstrating a diagnostic accuracy of 95.5%. The difference in diagnostic accuracy was 14.9% (95% CI 5.7% to 24.1%; $p = 0.002$) in favour of iuMRI.

The CSF-containing abnormalities made up 14 out of 81 of the diagnoses on ultrasonography. Eight fetuses had an enlarged CM and six fetuses had abnormal cystic structures (four arachnoid cysts and two unspecified). In total, 3 out of 14 cases with an ultrasonography diagnosis of a CSF abnormality were correct (diagnostic accuracy of 21.4%). In the 11 out of 14 cases of incorrect diagnoses on ultrasonography, the ORD consisted of 'no brain abnormality' in eight fetuses, failed commissuration in two fetuses and dural venous thrombosis with ectasia⁷⁰ in one fetus. The iuMRI was correct in 8 out of 14 of those cases (diagnostic accuracy of 57.1%), with a difference of 35.7% (95% CI -1.6% to 73.0%; $p = 0.0625$) in favour of iuMRI. All of the iuMRI errors were cases in which iuMRI reported a CSF-containing abnormality but this was not confirmed on ORD, although in one case a CC abnormality was also missed.

Diagnostic confidence

The assessment of the effect of iuMRI on diagnostic confidence is based on the confidence of which the dominant diagnosis was reported. The dominant diagnosis is defined as the one most likely to influence the prognosis and for each case was decided on by the independent panels. Overall, iuMRI reported the dominant diagnosis with high confidence (70% or 90% confidence) in 13% more cases than ultrasonography, with iuMRI reporting 544 out of 570 (95%) cases with high confidence compared with ultrasonography reporting 465 out of 570 (82%) cases (*Table 11*). There were 122 cases that were diagnosed with high confidence on ultrasonography but were subsequently found to be incorrect on ORD (22% of the overall population), compared with 32 (6%) cases on iuMRI. Diagnoses were reported with low confidence ($\leq 50\%$ confidence) in 105 out of 570 (18%) cases on ultrasonography and in 26 out of 570 (5%) cases on iuMRI. Eighteen (3% of all cases) cases reported with low confidence on iuMRI were found to be correct on ORD, compared with 45 (8%) correct on ultrasonography.⁴⁴

The majority of diagnoses showed no change in confidence from ultrasonography to iuMRI ($n = 331$, 58%; *Table 12*). When confidence did change, most diagnoses were more confident with iuMRI by one ($n = 91$, 16%) or two categories ($n = 93$, 16%). This paired difference between confidence on ultrasonography and iuMRI was found to be statistically significant ($p < 0.0001$).

TABLE 10 Diagnostic accuracy of ultrasonography and iuMRI in cases diagnosed with abnormalities of the posterior fossa on ultrasonography

Diagnosis	N	Ultrasonography		iuMRI		Difference (95% CI)	p-value ^a
		Incorrect (n)	Diagnostic accuracy (%)	Incorrect (n)	Diagnostic accuracy (%)		
Any abnormality of the posterior fossa	81	28	65.4	10	87.7	22.3 (14.0 to 30.5)	< 0.0001
CSF abnormality	67	13	80.6	3	95.5	14.9 (5.7 to 24.1)	0.002
Parenchymal abnormality	14	11	21.4	6	57.1	35.7 (-1.6 to 73.0)	0.0625

a McNemar's test between ultrasonography and iuMRI.

TABLE 11 Diagnostic confidence categories on ultrasonography and iuMRI in the primary analysis population ($n = 570$)

Diagnostic confidence (%)	Ultrasonography, n (%)		iuMRI, n (%)	
	Overall ($N = 570$)	Incorrect diagnoses ($N = 182$)	Overall ($N = 570$)	Incorrect diagnoses ($N = 40$)
10	18 (3.2)	12 (6.6)	6 (1.1)	3 (7.5)
30	25 (4.4)	18 (9.9)	5 (0.9)	2 (5.0)
50	62 (11.0)	30 (17.0)	15 (2.6)	3 (7.5)
70	102 (18.0)	27 (15.0)	51 (8.9)	7 (18.0)
90	363 (64.0)	95 (52.0)	493 (87.0)	25 (62.0)

TABLE 12 Difference in diagnostic confidence on ultrasonography alone and ultrasonography with iuMRI in the primary analysis population ($n = 570$)

Difference between ultrasonography alone and ultrasonography with iuMRI	ORD available, n (%) ($N = 570$)
No change	331 (58.0)
Ultrasonography more confident, one category	38 (6.7)
Ultrasonography more confident, two or more categories	17 (3.0)
iuMRI more confident, one category	91 (16.0)
iuMRI more confident, two or more categories	93 (16.0)
p -value ^a	< 0.0001

^a Wilcoxon signed-rank test between ultrasonography alone and ultrasonography with iuMRI.

Table 13 shows three approaches to assessing the changes brought about in the diagnostic confidence using iuMRI. The simple (conventional without correction) approach is to subtract the ultrasonography confidence from the iuMRI confidence; the estimated difference of 0.44 points indicates an average increase of almost half a point on the 5-point scale of confidence. Although intuitive, this difference takes no account of whether or not the increase in confidence is appropriate. The corrected conventional method penalises ultrasonography (but not iuMRI) confidence in cases when the two diagnoses differ but the ultrasonography diagnosis is made with high confidence. The Ng and Palmer⁵⁸ method also combines diagnostic confidence and accuracy but uses the ORD as the benchmark for the latter. The iuMRI was found to be preferable to ultrasonography, with an average improvement of 0.75 points (95% CI 0.63 to 0.83 points; $p < 0.0001$) on the ± 4 -point scale.

TABLE 13 Difference in diagnostic confidence on ultrasonography and iuMRI in the primary analysis population ($n = 570$)

Method	Difference in confidence, mean (SD)	95% CI	Test statistic (t)	p -value
Score-based method	0.75 (1.5)	0.63 to 0.88	12.0	< 0.0001
Conventional analysis without correction	0.44 (1.2)	0.35 to 0.54	8.9	< 0.0001
Conventional analysis with correction	1.1 (1.57)	0.97 to 1.23	16.8	< 0.0001

Clinical impact

Prognosis

Clinical feedback was available for 783 out of 823 pregnancies. Of the remaining 40 pregnancies, there were 26 participants for whom no follow-up took place, 13 participants for whom the clinical feedback form was not fully completed and one participant who withdrew from the study after the iuMRI but before the clinical feedback visit. The ORD was available for 545 pregnancies but was not a requisite for this analysis, meaning that the analysis was based on 783 mothers. A full summary of prognosis categories on ultrasonography and iuMRI can be found in *Appendices 6 and 7*.

Fetal medicine clinicians reported that the iuMRI provided additional diagnostic information in 387 out of 783 (49%) cases. In 201 out of 387 (52%) cases, the 'new' fetal brain abnormalities identified by iuMRI were visible on follow-up ultrasonography.

In 189 out of 783 (24%) cases, referring clinicians answered 'yes' to the question 'Did iuMR change your prognosis?'. In 138 out of 189 cases, additional non-neuroimaging investigations, such as karyotyping, fetal cardiac echo and infection screening, were performed in conjunction with the iuMRI examination. The results of such tests were considered to have a major influence on the prognosis in 32 of those cases. Therefore, we estimate that iuMRI per se changed the prognostic information in at least 157 out of 783 cases (20%). Although 24% of clinicians reported that iuMRI changed their prognosis, the actual prognostic information was modified in 342 out of 783 (44%) participants. The changes were primarily shifting away from 'intermediate' and 'unknown' to 'normal', 'favourable' or 'poor'. The most frequent change in prognosis was from 'unknown' on ultrasonography to a quotable risk category after iuMRI (113/783, 14% of cases). Conversely, on 33 occasions (4%) after iuMRI the prognosis was changed from a quotable known risk to 'unknown'. A full summary of these changes can be found in *Appendix 8*.

The prognosis was quotable after both scanning methods in 77% of cases; of these, the prognosis was improved after iuMRI in 102 out of 783 (13%) cases and worsened in 94 (12%) cases (*Table 14*). Most participants showed no change in prognosis ($n = 409$, 68%).

Counselling and management

Termination of pregnancy was discussed in 51% of consultations before iuMRI and in 49% of consultations after iuMRI. However, the rate of TOP being offered increased from 25% before iuMRI to 36% after iuMRI, with an additional 84 (11%) cases in which TOP was offered after iuMRI (*Table 15*). A higher proportion of cases had TOP offered based on substantial risk of disability following iuMRI than ultrasonography (33% following iuMRI and 23% following ultrasonography alone). There were also fewer cases when TOP was discussed but not offered following iuMRI (12% following iuMRI and 25% following ultrasonography alone).

TABLE 14 Change in prognosis from ultrasonography to iuMRI, by category, for mothers who have a known prognosis on both modalities

Change in prognosis	<i>n</i> (%)
Number of fetuses	605
No change	409 (68.0)
More favourable with ultrasonography, one category	83 (14.0)
More favourable with ultrasonography, two or more categories	11 (1.8)
More favourable with iuMRI, one category	89 (15.0)
More favourable with iuMRI, two or more categories	13 (2.1)
<i>p</i> -value ^a	0.543

a Wilcoxon signed-rank test.

TABLE 15 Counselling based on ultrasonography alone and iuMRI for all mothers who completed iuMRI

Counselling	Overall, n (%) (N = 810)
Ultrasonography alone	
TOP not discussed	400 (49.0)
TOP discussed but not offered	203 (25.0)
TOP offered based on substantial disability risk	184 (23.0)
TOP discussed and offered, other reason	23 (2.8)
iuMRI	
TOP not discussed	393 (49.0)
TOP discussed but not offered	98 (12.0)
TOP offered based on substantial disability risk	270 (33.0)
TOP discussed and offered, other reason	14 (1.7)
Not reported	35 (4.3)

Note

Belfast participants (n = 17) have been excluded from this table. A paired comparison of ultrasonography and iuMRI counselling can be found in *Appendix 9*.

In response to the question 'Did iuMR change your counselling in this case?' the referring clinician reported no influence on counselling in 172 out of 783 (22%) cases, minor influence in 496 out of 783 (63%) cases and major influence in 115 out of 783 (15%) cases. The contribution of iuMRI to the final choice of management was felt to be of 'no value' in 95 out of 783 (12%) cases, a 'minor influence' in 419 out of 783 (53%) cases, 'significant' in 201 out of 783 (26%) cases, a 'major influence' in 49 out of 783 (6%) cases and 'decisive' in 19 out of 783 (3%) cases (*Table 16*).⁴⁴ The contribution of iuMRI was rated particularly highly among those participants whose chosen management plan was TOP (*Table 17*). A total of 31 out of 103 (30%) participants said that it was 'significant', 9 out of 103 (8.8%) participants said that it was 'major' and 12 out of 103 (12%) participants said that it was 'decisive'.

TABLE 16 Management of pregnancy following iuMRI in 783 cases with complete clinical feedback data

Rating of iuMRI contribution	Overall, n (%) (N = 783)
Chosen management plan	
Continued pregnancy	213 (27.0)
Continued pregnancy, further imaging	465 (59.0)
TOP	102 (13.0)
Unable to decide	3 (0.4)
Further imaging	
Ultrasonography	368 (47.0)
iuMRI	16 (2.0)
Both	81 (10.0)
Not reported	318 (41.0)
Rating of iuMRI contribution	
None	95 (12.0)
Minor	419 (54.0)
Significant	201 (26)
Major	49 (6.3)
Decisive	19 (2.4)

TABLE 17 Management of pregnancy by iuMRI contribution in 783 cases with complete data

Rating of iuMRI contribution	Overall, <i>n</i> (%) (<i>N</i> = 783)	Management, <i>n</i> (%)			
		Continued with pregnancy (<i>N</i> = 213)	Continued with pregnancy with further imaging (<i>N</i> = 465)	Termination (<i>N</i> = 102)	Unable to decide (<i>N</i> = 3)
None	95 (12.0)	35 (16.0)	51 (11.0)	8 (7.8)	1 (33.0)
Minor	419 (54.0)	118 (55.0)	259 (56.0)	42 (41.0)	0 (0.0)
Significant	201 (26.0)	48 (23.0)	121 (26.0)	31 (30.0)	1 (33.0)
Major	49 (6.3)	10 (4.7)	29 (6.2)	9 (8.8)	1 (33.0)
Decisive	19 (2.4)	2 (0.9)	5 (1.1)	12 (12.0)	0 (0.0)

Additional tests/resource use

In the 821 mothers who had complete scan data, 924 additional tests were carried out with 559 mothers having at least one additional investigation (*Table 18*). The most common of these was 'infection screen' (*n* = 395, 48%). The median number of additional investigations was one (IQR 0–1). *Table 19* shows that three-dimensional ultrasonography and karyotyping were more likely to have an influence on prognosis and counselling (27% and 25% of referring clinicians said that the test changed prognosis and counselling).

TABLE 18 Summary of additional investigations performed in the 821 mothers who completed iuMRI

Investigation and clinical impact	Overall
Mothers with at least one additional test performed	559
Investigation, <i>n</i> (%)	
Karyotype	235 (29.0)
Echocardiography	128 (16.0)
Infection screen	395 (48.0)
Three-dimensional ultrasonography	145 (18.0)
Other NAIT	17 (2.1)
Other	4 (0.5)
Median number of additional investigations (IQR)	1 (0–1)
Additional investigation changed prognosis and counselling, <i>n</i> (%)	190 (21.0)
Major factor in choice of management, <i>n</i> (%)	144 (16.0)
Abnormal finding, <i>n</i> (%)	45 (6.0)
Three-dimensional ultrasonography, <i>n</i> (%)	
Prior to iuMRI	120 (85.0)
Volume saved offline and interpreted by expert	56 (39.0)
Modified existing diagnosis	37 (26.0)
Provided same diagnostic information as iuMRI	82 (58.0)
Modified confidence from two-dimensional ultrasonography, <i>n</i> (%)	
No change	85 (60.0)
Increased	56 (39.0)
Decreased	1 (0.7)

NAIT, neonatal alloimmune thrombocytopenia.

TABLE 19 Additional investigative tests and their impact on the prognosis in 559 mothers who received them

Investigative test	N	Impact, n (%)		
		Did the test change the prognosis and counselling?	Was the test a major factor in the choice of management?	Abnormal finding
Karyotype	235	59 (25.0)	50 (21.0)	10 (4.3)
Echo	128	28 (22.0)	17 (13.0)	23 (18.0)
Infection screen	395	64 (16.0)	43 (11.0)	12 (3.1)
Three-dimensional ultrasonography	145	39 (27.0)	33 (23.0)	0 (0.0)
Other NAIT	17	0 (0.0)	1 (6.2)	0 (0.0)
Other	4	0 (0.0)	0 (0.0)	0 (0.0)

NAIT, neonatal alloimmune thrombocytopenia.

There was a similar result when the clinicians were asked about whether or not the test was a major factor in the choice of management (three-dimensional ultrasonography at 22.8% and karyotyping at 21%). Abnormal findings were not shown in many of these additional tests (45/924 tests).

Safety

There were 214 SAEs in total, all of which were defined as expected in the study protocol. There were 138 preterm deliveries (defined as a delivery before 37 weeks' gestational age), 33 cases of spontaneous intrauterine fetal demise or stillbirth, 29 neonatal or infant deaths and 14 cases in which the iuMRI diagnosis differed from the ORD and a TOP had been performed.

All of the above were reviewed by a nominated fetal medicine specialist and the DMEC, who did not have any concerns regarding the rate of events.

A full case review was instigated for the 14 participants for whom the iuMRI diagnosis differed from the ORD and a TOP was performed. In all cases, the TOP was judged to have been performed because of the severity of a brain abnormality confirmed on iuMRI, another abnormality not related to the study diagnosis (e.g. a cardiac abnormality or a diagnosis of a chromosomal disorder) or the case underwent a full multidisciplinary team (MDT) review during which it was concluded that the ORD was incomplete (e.g. no autopsy) or incorrect and an addendum to the ORD was issued.

There were 130 AEs reported during iuMRI scanning, with 123 (15%) participants experiencing at least one AE (Table 20). The most common reason associated with this was persistent fetal movement ($n = 104$ AEs, 80%).

In utero magnetic resonance imaging errors

In the primary analysis population ($n = 570$ participants), there were 40 iuMRI errors overall, resulting in an overall error rate of 7% (Table 21). When summarising by central and non-central reporters, 12 out of 40 (30%) errors were attributed to the central reporter and 28 out of 40 (70%) errors were attributed to the non-central reporters. The overall error rate was higher in non-central reporters (11%) than in the central reporter (3.8%), with a difference of 7.2% (95% CI 2.5% to 12.0%).⁷¹

TABLE 20 Adverse events during iuMRI that occurred during the MERIDIAN study

Adverse events	Mothers in the MERIDIAN cohort (N = 823), n (%)	Mothers with incomplete scans (not part of the MERIDIAN cohort), n (%)
Total iuMRI-related AEs	130	7
Mothers experiencing at least one AE	123 (15.0)	6 (13.0)
Type of AE		
Persistent fetal movement	104 (80.0)	1 (14.0)
Anxiety/claustrophobia	12 (9.0)	4 (57.0)
MRI equipment or software failure	5 (3.8)	0 (0.0)
Skin heating/other sensory effects	1 (0.8)	0 (0.0)
Other AE	17 (13.0)	2 (29.0)

TABLE 21 Error rates in relation to the experience category of the reporting radiologist⁷¹

Experience category (number of radiologists)	Number of cases reported in the MERIDIAN study (overall error rate %)	Error rate (%) in the radiologists' reports			
		First 25 reports	Reports 26–50	Reports 51–75	Reports ≥ 76
Non-central reporter; 0–50 iuMRI studies (3)	95 (12.0)	13.0 (of 54 cases)	16.0 (of 25 cases)	0.0 (of 16 cases)	Not performed
Non-central reporter; 50–150 iuMRI studies (4)	138 (12.0)	14.0 (of 87 cases)	11.0 (of 37 cases)	7.1 (of 14 cases)	Not performed
Non-central reporter; 150–300 iuMRI studies (4)	71 (8.5)	9.4 (of 64 cases)	0.0 (of 7 cases)	Not performed	Not performed
Central reporter; > 1000 iuMRI studies (1)	316 (3.8)	4.0 (of 25 cases)	0.0 (of 25 cases)	0.0 (of 25 cases)	4.6 (of 241 cases)

When further categorised by previous experience, radiologists with less previous experience had a higher error rate (12% and 12%) than radiologists with more previous experience (8.5% and 3.8%). Radiologists with > 150 previous scans had a consistently low error rate (< 10%) and the central reporter had an error rate that was consistently < 5%. However, the data are limited by the small number of radiologists in each category.

Hospital Anxiety and Depression Scale

The HADS questionnaire was used to assess anxiety and depression among study participants, both during their pregnancy (survey 1) and postnatally (survey 2). In total, 414 participants completed survey 1 and 274 participants completed survey 2, with 194 participants completing both. High rates of anxiety were apparent during pregnancy, with over one-third of participants reporting scores in the range of 'probable case'; a total of 15% of study participants were in the 'probable case' range for depression (Table 22). The mean anxiety score was lower in survey 2 than in survey 1 (paired difference of –3 points; $p < 0.0001$) for those participants who responded to both surveys. This trend was also observed in depression scores (paired difference of –2.2 points; $p < 0.0001$). There were also more participants in the 'normal' anxiety category in survey 2 than in survey 1 (39% of participants in survey 1, 64% of participants in survey 2), with fewer participants in the 'cause for concern' and 'probable clinical case' categories in survey 2. Although not as strong, a similar trend was observed in the depression categories.

TABLE 22 Summary statistics of the HADS questionnaire completed as part of surveys 1 and 2

HADS category	Survey 1	Survey 2	Difference
HADS: anxiety			
<i>N</i>	414	274	194
Mean (SD)	9 (4.7)	6.5 (4.7)	-3 (4.5)
Median (IQR)	9 (5 to 12)	6 (3 to 9)	-3 (-6 to 0)
Minimum to maximum	0 to 21	0 to 19	-15 to 14
<i>p</i> -value			≤ 0.0001 ^a
Anxiety category, <i>n</i> (%)			
Normal range (≤ 7.5)	160 (39)	176 (64)	-25%
Cause for concern (> 7.5–10.5)	102 (25)	42 (15)	10%
Probable clinical case (> 10.5)	152 (37)	56 (20)	17%
<i>p</i> -value			≤ 0.0001 ^b
HADS: depression			
<i>N</i>	414	274	194
Mean (SD)	5.7 (4.4)	3.7 (3.8)	-2.2 (3.9)
Median (IQR)	5 (2 to 9)	3 (1 to 5)	-1 (-4 to 0)
Minimum to maximum	0 to 21	0 to 21	-18 to 7
<i>p</i> -value			≤ 0.0001 ^a
Depression category, <i>n</i> (%)			
Normal range (≤ 7.5)	289 (70.0)	229 (84.0)	-14%
Cause for concern (> 7.5–10.5)	65 (16.0)	29 (11.0)	5%
Probable clinical case (> 10.5)	60 (15.0)	16 (5.8)	9.2%
<i>p</i> -value			≤ 0.0001 ^b

^a Paired Wilcoxon signed-rank test between survey 1 and survey 2.

^b McNemar's test between survey 1 and survey 2.

Health economic results

Outcomes

A total of 570 women were part of the primary analysis set, had outcomes recorded and had iuMRI within 2 weeks of the ultrasonography. Information on whether or not iuMRI changed the prognosis was missing for 19 women. Table 23 shows whether or not iuMRI was correct and if it changed the prognosis for the remaining 551 women.

TABLE 23 'Was iuMR correct?' by 'Did iuMR change the prognosis?'

Was iuMRI correct?	Did iuMRI change the prognosis?		Total
	No	Yes	
No	30	10	40
Yes	391	120	511
Total	421	130	551

Of the 570 women in the primary analysis set, information on whether or not TOP was offered after ultrasonography was missing in 23 women and information on whether or not TOP was offered after iuMRI was missing in 33 women.

Of the 570 women with information on whether or not TOP happened, 68 (11.93%) women had a TOP.

Of the 547 women with information on whether or not TOP was offered after ultrasonography, 135 (24.68%) women were offered TOP after ultrasonography.

Of the 537 women with information on whether or not TOP was offered after iuMRI, 189 (35.20%) women were offered TOP. A total of 58 of the 189 (30.69%) women offered TOP after iuMRI had a TOP (*Table 24*).

Costs

Disaggregated and total costs for iuMRI resulting in TOP, iuMRI not resulting in TOP, ultrasonography resulting in TOP and ultrasonography not resulting in TOP for complete cases are shown in *Table 25*.

Base case, scenario 1 and scenario 2

Results for the base case, scenario 1 and scenario 2 analyses are shown in *Table 26*.

Cost-effectiveness

In the base case, 21.78% (120/551) of decisions were appropriately revised following iuMRI. The incremental cost per management decision appropriately revised is therefore £2203.

TABLE 24 'Did TOP happen?' by 'Was TOP offered after iuMRI?'

Did TOP happen?	Was TOP offered after iuMRI?		Total
	No	Yes	
No	348	131	479
Yes	0	58	58
Total	348	189	537

TABLE 25 Disaggregated and total costs

Costs (£)	iuMRI no TOP	iuMRI followed by TOP	Ultrasonography without TOP	Ultrasonography followed by TOP
Mileage	26	37		
Patient-reported personal costs	51	69		
Ultrasonography	271	269	271	269
iuMRI	264	264		
Investigation	206	243	206	243
Further imaging	336	285	336	285
Further consultations	140	12	140	12
Cost of TOP	0	1098	0	1098
Cost of delivery	2647	0	2647	0
Total cost (NHS)	3864	2170	3600	1907
Total cost (all)	3941	2277	3600	1907

TABLE 26 Deterministic results, imputed missing cost data results and probabilistic results

Cost type	iuMRI		Ultrasonography		Incremental cost (£)	Proportion of management decisions appropriately revised following iuMRI (%)	Cost-effectiveness (per management decision appropriately revised following iuMRI) (£)	
	No TOP	TOP	No TOP	TOP				
Deterministic	Base case. iuMRI: proportion of patients who had TOP; ultrasonography: proportion of patients offered TOP after ultrasonography							
	Proportion (%)	88.1	11.9	75.3	24.7			
	Cost (£)	3864	2170	3600	1907			
	Total cost (£)	3662		3182		480	21.8	2203
	Scenario 1. iuMRI: proportion of patients offered TOP after iuMRI; ultrasonography: proportion of patients offered TOP after ultrasonography							
	Proportion (%)	64.8	35.2	75.3	24.7			
	Cost (£)	3864	2170	3600	1907			
	Total cost (£)	3268		3182		86	21.8	394
	Scenario 2. iuMRI: proportion of patients who had TOP; ultrasonography: proportion of patients offered TOP after ultrasonography, adjusted by the proportion of patients offered TOP after iuMRI who had TOP							
Proportion (%)	88.1	11.9	92.4	7.6				
Cost (£)	3864	2170	3600	1907				
Total cost (£)	3662		3472		190	21.8	873	
Imputed missing cost data	Base case. iuMRI: proportion of patients who had TOP; ultrasonography: proportion of patients who were offered TOP after ultrasonography							
	Proportion (%)	88.1	11.9	75.3	24.7			
	Cost (£)	3848	2143	3584	1879			
	Total cost (£)	3645		3163		481	21.8	2210
	Scenario 1. iuMRI: proportion of patients offered TOP after iuMRI; ultrasonography: proportion of patients offered TOP after ultrasonography							
	Proportion (%)	64.8	35.2	75.3	24.7			
	Cost (£)	3848	2143	3584	1879			
	Total cost (£)	3248		3163		85	21.8	389

continued

TABLE 26 Deterministic results, imputed missing cost data results and probabilistic results (*continued*)

Cost type	iuMRI		Ultrasonography		Incremental cost (£)	Proportion of management decisions appropriately revised following iuMRI (%)	Cost-effectiveness (per management decision appropriately revised following iuMRI) (£)	
	No TOP	TOP	No TOP	TOP				
Probabilistic sensitivity analysis	Scenario 2. iuMRI: proportion of patients who had TOP; ultrasonography: proportion of patients offered TOP after ultrasonography, adjusted by the proportion of patients offered TOP after iuMRI who had TOP							
	Proportion (%)	88.1	11.9	92.4	7.6			
	Cost (£)	3848	2143	3584	1879			
	Total cost (£)	3645		3455		190	21.8	871
	Base case. iuMRI: proportion of patients who had TOP; ultrasonography: proportion of patients offered TOP after ultrasonography							
	Proportion (%)	88.1	11.9	75.3	24.8			
	Cost (£)	3876	2329	3612	2065			
	Total cost (£)	3690		3232		458	21.8	2102
	Scenario 1. iuMRI: proportion of patients offered TOP after iuMRI; ultrasonography: proportion of patients offered TOP after ultrasonography							
	Proportion (%)	64.8	35.3	75.3	24.6			
	Cost (£)	3881	2318	3617	2054			
	Total cost (£)	3333		3230		103	21.8	470
	Scenario 2. iuMRI: proportion of patients who had TOP; ultrasonography: proportion of patients offered TOP after ultrasonography, adjusted by the proportion of patients offered TOP after iuMRI who had TOP							
Proportion (%)	88.1	11.9	92.4	7.6				
Cost (£)	3878	2337	3614	2073				
Total cost (£)	3693		3497		195	21.8	899	

For scenario 1, if we assumed that each woman offered TOP after iuMRI had a TOP, then the average cost to the NHS of iuMRI would be £3268. The cost of ultrasonography is unchanged from the base case. The incremental cost is therefore £396 and the incremental cost per management decision appropriately revised is £394.

For scenario 2, if we assumed that 30.69% of women offered TOP after ultrasonography had a TOP, then 7.57% of women who underwent ultrasonography would have had a TOP. The average cost to the NHS of ultrasonography would be £3472. The cost of iuMRI is unchanged from the base case. The incremental cost is therefore £301 and the incremental cost per management decision appropriately revised is £873.

The cost of a delivery is higher than the cost of a termination. In the base case, iuMRI leads to fewer TOPs than ultrasonography; together with the additional cost for iuMRI, this means that the total cost is higher for iuMRI than for ultrasonography. In scenarios 1 and 2, the proportion of TOPs for iuMRI and ultrasonography is more similar than in the base case. In scenarios 1 and 2, iuMRI leads to more TOPs than ultrasonography, so the costs are more comparable, although the cost of iuMRI means that the total cost is higher for iuMRI than for ultrasonography but the incremental cost is lower than the base case.

Imputed missing data

When costs were analysed to impute missing data, the total cost to the NHS for iuMRI was £3848 without TOP and £2143 with TOP. The total cost for ultrasonography was £3584 without TOP and £1879 with TOP. Results for the base case, scenario 1 and scenario 2 using imputed missing data for costs are shown in *Table 26*.

Probabilistic sensitivity analysis

Mean results from 1000 probabilistic simulations are shown in *Table 26*.

Cost-effectiveness acceptability curves are presented for the base case, scenario 1 and scenario 2 in *Figure 2*. For the base-case analysis, if we are willing to pay \geq £4000 for an appropriate management decision, iuMRI has a probability of 100% of being the most cost-effective option. For scenarios 2 and 3, we can be certain that iuMRI imaging is the most cost-effective option.

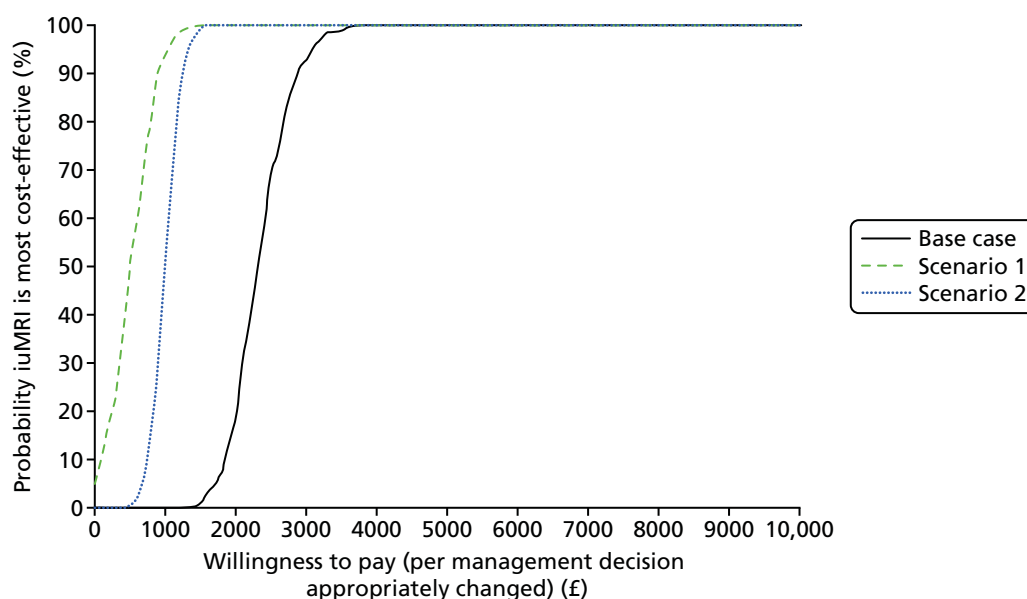


FIGURE 2 Cost-effectiveness acceptability curve.

Graphs for the per-person EVPI at varying willingness-to-pay thresholds are presented in *Figure 3*. The EVPI peaks at the willingness-to-pay threshold at which there is most uncertainty about whether iuMRI or ultrasonography is the most cost-effective technique. The maximum EVPI is relatively low (£14–30), indicating that there is little value in conducting further research on the parameters in this short-term analysis.

Scenario analyses

Inclusion of participant-borne costs

If we include personal costs (i.e. mileage costs and time to travel to and undergo iuMRI) as well as NHS costs, the average cost of iuMRI increases and the cost of ultrasonography stays the same. This reflects the additional costs incurred by the participants to travel to and undergo a iuMRI scan. If we limit the distance travelled to 40 miles and include personal costs (i.e. mileage costs and time to travel to and undergo iuMRI) as well as NHS costs, all of the costs change slightly as different data are being analysed. Results including personal costs are shown in *Table 27*.

Subgroup analyses

Results are presented for gestational ages of < 22 weeks, 22–24 weeks and > 24 weeks at iuMRI scan in *Table 28*.

Comparison with repeat ultrasonography

A comparison of iuMRI with repeat ultrasonography is shown in *Table 29*. In this analysis, the incremental cost is negative, meaning that iuMRI is cost-saving compared with repeat ultrasonography. The incremental effectiveness of iuMRI over repeat ultrasonography is reduced.

Different costing scenarios

Results excluding delivery costs only, excluding delivery and termination costs, and increasing termination costs are shown in *Table 30*.

When only delivery costs are excluded, TOP becomes more costly than continuation of pregnancy. In the base case, there are more TOPs for ultrasonography than for iuMRI, but these costs are offset somewhat by the cost of iuMRI itself. However, the incremental cost is much less than when delivery costs are included (deterministic results in *Table 26*) and, therefore, the cost per management decision appropriately revised decreases. In scenarios 1 and 2, the incremental costs have increased (compared with deterministic results in *Table 26*) because there are more TOPs for iuMRI than for ultrasonography, which lead the cost per management decision appropriately revised to increase. The estimates of cost-effectiveness for the

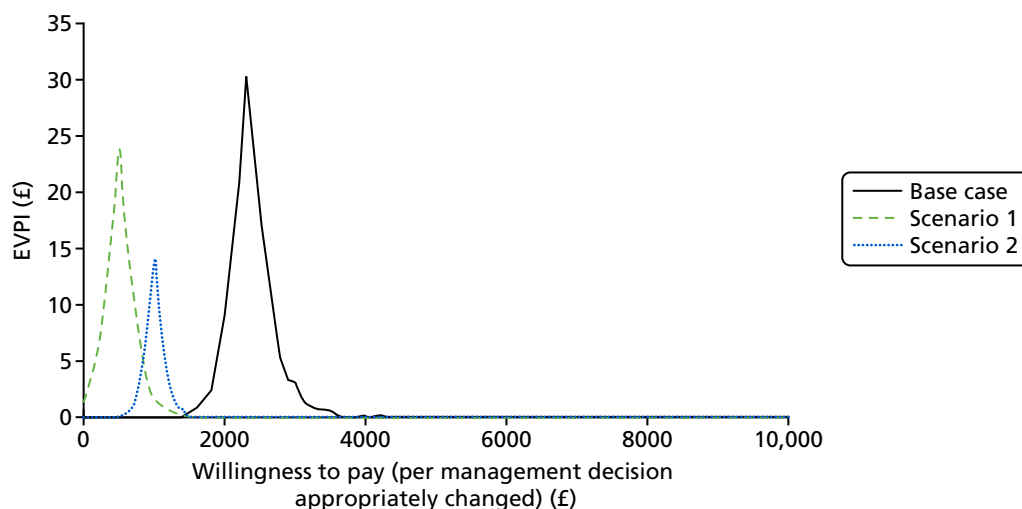


FIGURE 3 The EVPI.

TABLE 27 Scenario analyses including personal costs

Scenario	iuMRI		Ultrasonography		Incremental cost (£)	Proportion of management decisions appropriately revised following iuMRI (%)	Cost-effectiveness (per management decision appropriately revised following iuMRI) (£)
	No TOP	TOP	No TOP	TOP			
All participants							
Proportion (%)	88.1	11.9	75.3	24.7			
Cost (£)	3941	2277	3600	1907			
Total cost (£)	3743		3182		561	21.8	2574
Mileage (< 40 miles)							
Proportion (%)	88.9	11.2	73.5	26.5			
Cost (£)	3880	1987	3570	1682			
Total cost (£)	3669		3070		599	21.8	2753

TABLE 28 Subgroup analyses by gestational age at iuMRI

Gestational age (weeks)	iuMRI		Ultrasonography		Incremental cost (£)	Proportion of management decisions appropriately revised following iuMRI (%)	Cost-effectiveness (per management decision appropriately revised following iuMRI) (£)
	No TOP	TOP	No TOP	TOP			
< 22							
Proportion (%)	79.8	20.2	71.0	29.0			
Cost (£)	3851	2213	3587	1949			
Total cost (£)	3520		3112		408	20.5	1993
22–24							
Proportion (%)	88.6	11.4	75.5	24.5			
Cost (£)	3883	1815	3619	1551			
Total cost (£)	3646		3112		535	18.8	2851
> 24							
Proportion (%)	96.0	4.0	79.5	20.5			
Cost (£)	3857	2284	3593	2020			
Total cost (£)	3794		3270		524	25.7	2043

TABLE 29 Scenario analysis comparing iuMRI with repeat ultrasonography

Scenario	iuMRI		Repeat ultrasonography		Incremental cost (£)	Proportion of management decisions appropriately revised following iuMRI (%)	Cost-effectiveness (per management decision appropriately revised following iuMRI) (£)
	No TOP	TOP	No TOP	TOP			
Proportion (%)	88.1	11.9	90.7	9.3			
Cost (£)	3864	2170	3827	2134			
Total cost (£)	3662		3670		–8	10.6	Cost saving

TABLE 30 Different costing scenarios

Scenario	iuMRI		Ultrasonography		Incremental cost (£)	Proportion of management decisions appropriately revised following iuMRI (%)	Cost-effectiveness (per management decision appropriately revised following iuMRI) (£)
	No TOP	TOP	No TOP	TOP			
Excluding delivery costs							
<i>Base case. iuMRI: proportion of patients who had TOP; ultrasonography: proportion of patients offered TOP after ultrasonography</i>							
Proportion (%)	88.1	11.9	75.3	24.7			
Cost (£)	1217	2170	953	1907			
Total cost (£)	1331		1188		142	21.8	654
<i>Scenario 1. iuMRI: proportion of patients offered TOP after iuMRI; ultrasonography: proportion of patients offered TOP after ultrasonography</i>							
Proportion (%)	64.8	35.2	75.3	24.7			
Cost (£)	1217	2170	953	1907			
Total cost (£)	1553		1188		364	21.8	1672
<i>Scenario 2. iuMRI: proportion of patients who had TOP; ultrasonography: proportion of patients offered TOP after ultrasonography, adjusted by the proportion of patients offered TOP after iuMRI who had TOP</i>							
Proportion (%)	88.1	11.9	92.4	7.6			
Cost (£)	1217	2170	953	1907			
Total cost (£)	1331		1025		305	21.8	1403
Excluding delivery and termination costs							
<i>Base case. iuMRI: proportion of patients who had TOP; ultrasonography: proportion of patients who were offered TOP after ultrasonography</i>							
Proportion (%)	88.1	11.9	75.3	24.7			
Cost (£)	1217	1072	953	809			
Total cost (£)	1200		917		282	21.8	1297
<i>Scenario 1. iuMRI: proportion of patients offered TOP after iuMRI; ultrasonography: proportion of patients offered TOP after ultrasonography</i>							
Proportion (%)	64.8	35.2	75.3	24.7			
Cost (£)	1217	1072	953	809			
Total cost (£)	1166		917		249	21.8	1142

continued

TABLE 30 Different costing scenarios (continued)

Scenario	iuMRI		Ultrasonography		Incremental cost (£)	Proportion of management decisions appropriately revised following iuMRI (%)	Cost-effectiveness (per management decision appropriately revised following iuMRI) (£)
	No TOP	TOP	No TOP	TOP			
<i>Scenario 2. iuMRI: proportion of patients who had TOP; ultrasonography: proportion of patients offered TOP after ultrasonography, adjusted by the proportion of patients offered TOP after iuMRI who had TOP</i>							
Proportion (%)	88.1	11.9	92.4	7.6			
Cost (£)	1217	1072	953	809			
Total cost (£)	1200		942		258	21.8	1183
Increased termination costs							
<i>Base case. iuMRI: proportion of patients who had TOP; ultrasonography: proportion of patients who were offered TOP after ultrasonography</i>							
Proportion (%)	88.1	11.9	75.3	24.7			
Cost (£)	3864	2459	3600	2196			
Total cost (£)	3696		3253		443	21.8	2034
<i>Scenario 1. iuMRI: proportion of patients offered TOP after iuMRI; ultrasonography: proportion of patients offered TOP after ultrasonography</i>							
Proportion (%)	64.8	35.2	75.3	24.7			
Cost (£)	3864	2459	3600	2196			
Total cost (£)	3370		3253		116	21.8	534
<i>Scenario 2. iuMRI: proportion of patients who had TOP; ultrasonography: proportion of patients offered TOP after ultrasonography, adjusted by the proportion of patients offered TOP after iuMRI who had TOP</i>							
Proportion (%)	88.1	11.9	92.4	7.6			
Cost (£)	3864	2459	3600	2196			
Total cost (£)	3696		3494		203	21.8	931

base case, scenario 1 and scenario 2 are closer together, because the incremental cost of iuMRI itself, which is the same between the scenarios, is a bigger driver of costs.

When delivery and termination costs are excluded, the incremental costs for the base case, scenario 1 and scenario 2 are very similar to each other, as the only differences in the costs are the cost of the iuMRI and the differences in investigations and further imaging.

When the cost of termination increases, it becomes closer to the cost of delivery. Therefore, there is less difference in the costs for people with and without TOP. The incremental costs for the base case, scenario 1 and scenario 2 become closer to each other than they were in the deterministic results in *Table 26*, but the difference between them is still larger than when termination and delivery costs are excluded.

Longer-term horizon

In this scenario (*Table 31*), the cost of a continued pregnancy becomes much higher than the cost of a termination, and iuMRI becomes much more costly than ultrasonography when it results in fewer TOPs (as in the base case) and becomes cost-saving when it results in more TOPs (as in scenarios 1 and 2). This analysis demonstrates the uncertainty in the longer-term economic evaluation. Longer-term follow-up studies of children diagnosed with fetal brain abnormalities would be required to address this uncertainty.

TABLE 31 Long-term scenario

Scenario	iuMRI		Ultrasonography		Incremental cost (£)	Proportion of management decisions appropriately revised following iuMRI (%)	Cost-effectiveness (per management decision appropriately revised following iuMRI) (£)
	No TOP	TOP	No TOP	TOP			
Base case. iuMRI: proportion of patients who had TOP; ultrasonography: proportion of patients offered TOP after ultrasonography							
Proportion (%)	88.1	11.9	75.3	24.7			
Cost (£)	27,527	2170	27,264	1907			
Total cost (£)	24,502		21,005		3497	21.8	16,057
Scenario 1. iuMRI: proportion of patients offered TOP after iuMRI; ultrasonography: proportion of patients offered TOP after ultrasonography							
Proportion (%)	64.8	35.2	75.3	24.7			
Cost (£)	27,527	2170	27,264	1907			
Total cost (£)	18,603		21,005		-2402	21.8	Cost saving
Scenario 2. iuMRI: proportion of patients who had TOP; ultrasonography: proportion of patients offered TOP after ultrasonography, adjusted by the proportion of patients offered TOP after iuMRI who had TOP							
Proportion (%)	88.1	11.9	92.4	7.6			
Cost (£)	27,527	2170	27,264	1907			
Total cost (£)	24,502		25,343		-841	21.8	Cost saving

Chapter 4 Qualitative substudy

Introduction

The aim of the patient and health professional perspectives substudy was to assess the acceptability of the clinical care package including iuMRI using qualitative and quantitative methods. This aim encompassed three objectives: (1) describing, exploring and understanding women's views about the acceptability of iuMRI in service user perspectives of the PND for brain abnormality; (2) describing, exploring and understanding the acceptability of iuMRI in the PND for brain abnormality from a service provider perspective; and (3) drawing on health professionals' views about potential barriers to and facilitators of developing policy and provision should the study outcomes suggest that there is a case for enhancing UK service provision.

Existing research on women's experiences of iuMRI in the PND pathway is relatively limited. Evidence suggests that women are accepting of the procedure, but that including iuMRI can be associated with higher levels of psychological distress.⁷² Existing sociological work suggests that including iuMRI can help manage the uncertainty associated with the PND.⁷³ Understanding women's perceptions of the PND for brain abnormality, including iuMRI, is important in two ways. First, MRI techniques are poorly understood in non-professional domains.^{74,75} Second, social scientists have noted that applying innovative health technologies can have unintended social effects, so early assessment of service user perspectives could help understand the impact of introducing new technologies in particular social and cultural contexts.³¹ Research on health professionals' views of iuMRI is limited but suggests that iuMRI can function as a bridge for clinicians working across specialties.⁷⁶ Diagnosis in particular can be conceptualised as a 'powerful social tool',⁷⁷ therefore understanding health professionals' perspectives on innovation within the PND process is an important part of understanding its impact, and can offer insights to facilitate future developments. Well-informed policy development is crucial to enhance experiences of the PND for brain abnormality (often with uncertain outcomes).

This substudy drew on quantitative data from women and qualitative data from women and health professionals to assess the acceptability of iuMRI in the PND for fetal brain abnormality. This mixed-method approach provided an in-depth analysis of key issues and experiences that matter to those with first-hand knowledge of iuMRI in practice.

Recruitment, sampling and data-collection methods

The three data sets (i.e. women's survey responses, women's interviews and health professionals' interviews) were used to explore service users' and health professionals' experiences of engaging with the clinical study. The three elements are outlined as separate activities but are considered together in informing the discussion of the findings, providing triangulation in data interpretation. The surveys also collected qualitative data to enhance the interpretation of the main data sets.

Questionnaire surveys

Women who were approached about the sociological study (after recruitment into the study) were provided with a site-specific information pack. Survey 1 was administered after the fetal medicine consultation during which the iuMRI findings were discussed and a provisional management plan was made. Survey 2 was administered 3–6 months post pregnancy outcome (either birth or TOP). Both surveys included questions to evaluate (1) the overall satisfaction with care, (2) the practical utility of results from iuMRI to inform understanding and decision-making and (3) to what extent participants' experiences were influenced by their situation (e.g. psychological state) at the time of iuMRI. This last objective utilised the HADS questionnaire data at the end of each survey.

Survey 1 differed slightly from survey 2. Both surveys comprised questions on age, marital status, education and questions on the impact of practical issues on the acceptability of the care package related to referral pathways (e.g. travel times). Survey 2 included (1) an open-text question that allowed participants to raise issues they felt were important to evaluating their health-care experience and (2) a filter question with an expression of interest (EOI) form for an in-depth interview. Follow-up letters and survey forms were sent to participants who had not responded after a certain time (Figure 4), improving returns by 4% (survey 1) and 6% (survey 2).

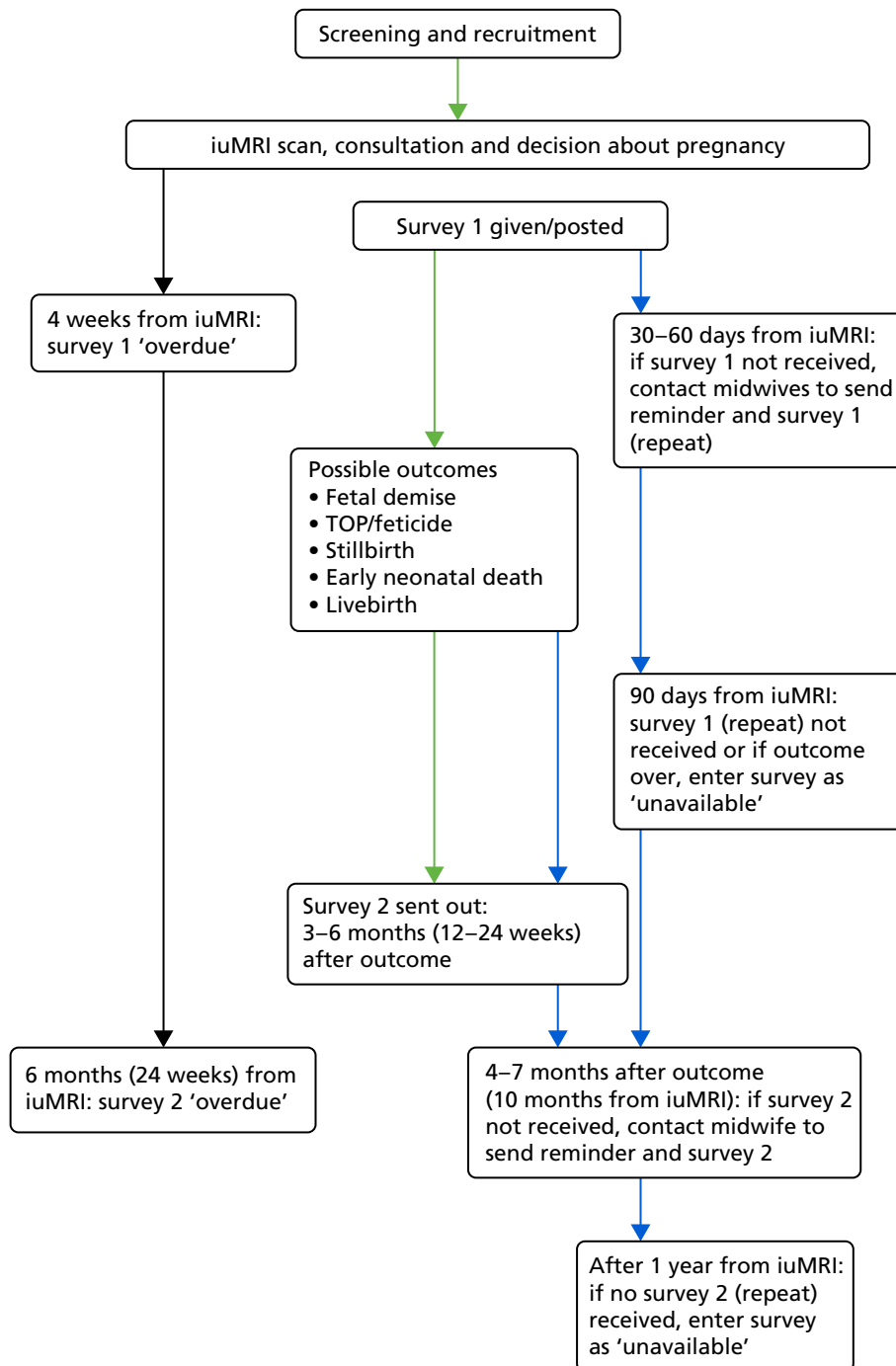


FIGURE 4 Substudy survey flow chart. Black, MERIDIAN database; green, site fetal medicine unit clinicians; blue, substudy researchers.

The primary focus of the survey was the quantitative data collected, but there were also a small number of open-text responses in survey 2 and marginalia. Content analysis using NVivo 10 (QSR International, Warrington, UK) was employed to code these textual data that were then used, when appropriate, to inform the analytic process.

Qualitative interviews with women

Survey respondents who completed an EOI form in survey 2 formed the sampling frame for the qualitative interviews with women. This approach facilitated the use of background characteristics, diagnostic category and pregnancy outcome alongside fetal medicine unit and iuMRI centre in the purposive sampling for maximum variation.⁷⁸ Women were contacted using information provided in EOI forms. Following the verification of the agreement to be interviewed, they were sent a study pack with information about the interview phase (including a reply slip and information on how a partner/relative/friend could participate in the woman's interview if she desired).

Following written consent, interviews were conducted in participants' homes and recorded on an encrypted digital voice recorder. A narrative approach to interviewing allowed participants to develop their own account of their 'story'. The topic guide prompted participants to describe their feelings on learning about the abnormality, undergoing the iuMRI and decision-making about pregnancy management. The topic guide also explored perceived differences between ultrasonography and iuMRI, information-seeking behaviour and the formal and informal support received.

Qualitative interviews with health professionals

Health professional participants were recruited from an identified pool of 42 health professionals associated with the clinical study, including 10 sites that were recruited in a suitable time frame for inclusion. Purposive sampling ensured that the participant group was sufficiently diverse (1) to cover a range of clinical sites, (2) to cover a range of professionals (e.g. fetal medicine specialists, radiology specialists), (3) in gender and (4) in length of service in the field. Potential participants were contacted by post with a study information pack, including an information sheet and a reply slip. On receipt of a reply agreeing to take part, the interviewer made contact using the details provided by the potential participant.

Interviews were carried out at two time points. The first round of interviews took place during the first year of clinical recruitment at the participant's site, and the second round of interviews took place during the third year of recruitment at that site (originally the final year of recruitment, prior to the extension of the clinical study). Whenever possible, participants interviewed in the first round were reinterviewed in the second round, but this was not possible in all cases. In the second round of interviews, participants were offered a telephone or face-to-face interview and additional participants were included. Face-to-face interviews were held at the participants' place of work. The topic guide included prompts on key issues, including experiences of the clinical project, perceived value of iuMRI in the PND and potential barriers to and facilitators of future policy implementation.

Qualitative data management

Interviews continued until data saturation was reached (i.e. we achieved a sufficiently diverse sample with no new themes emerging in the ongoing thematic analysis, leading to a stable descriptive coding framework). All digital sound files were stored on a password-secure server and deleted from the digital recorders. The interviews were transcribed, edited for accuracy, anonymised and entered into a qualitative data-management software for indexing and retrieval, ATLAS.ti version 7 (ATLAS.ti Scientific Software Development GmbH, Berlin, Germany).

Data analysis

Data analysis approach for survey data

Responses to the four key questions were investigated using multiple logistic regression. For the purpose of these analyses, responses were recategorised into 'positive' and 'negative' binary variables as follows:

- For Likert scale questions – 1 if the response was 'strongly agree' or 'agree', 0 for everything else.
- For rating scale questions – 1 if the response was 'excellent' or 'very good', 0 for everything else.

A variable selection approach was used with variables that were selected for use in the final model if they had a p -value of < 0.1 . The following models were calculated for each of the four outcomes in questionnaires 1 and 2:

- background – maternal age, gestational age, level of education
- service – Sheffield versus non-Sheffield MRI centres, journey length
- prognosis on iuMRI
- diagnosis on iuMRI
- change from ultrasonography after iuMRI – change in counselling, change in prognosis.

A further model was produced for questionnaire 2:

- outcome – outcome for pregnancy, agreement between ultrasonography and iuMRI.

Odds ratios (ORs) with CIs and corresponding p -values were calculated. Only complete cases were included (i.e. those mothers who had a value for the outcome and each of the covariates were included in the model). Model fit was assessed for each of the models by the Hosmer–Lemeshow goodness-of-fit test. All 49 questionnaire 2 responders who underwent TOP strongly agreed or agreed, meaning that the model could not be estimated owing to separation; therefore, TOP was removed.

Data analysis for qualitative interview data

For the service user and the health professional interview data analyses, anonymised interview transcripts (primary documents) constituted the formal data. Service users' transcripts were coded separately from health professionals' transcripts. The health professionals' transcripts were initially subdivided by round. When a high degree of similarity in the data from the two data-collection periods became clear, the health professionals' transcripts were analysed as one set of primary documents. In both the service user and the health professional interview data analyses, a generative thematic approach was used, incorporating a multistage process similar to that described by Braun and Clarke^{79,80} (Table 32). Our approach to generative thematic analysis is informed by Silverman^{81,82} and shares some principles with grounded theory approaches,⁸³ for example the identification of themes that emerge from the data.⁸⁴ However, we did not follow any specific grounded theory protocol. Initial coding frames were developed by each lead analyst and with coding samples reviewed by the co-analyst, providing a qualitative equivalent to inter-rater reliability, enhancing consistency and theme refinement. The final theme content was then condensed and further refined by the lead analyst, with input from the other two members of the substudy team.

The sensitive topic area meant that interviews were sometimes emotional for service user participants, but still generated rich, in-depth narratives. In most cases, the presence of a partner or relative was supportive and their participation gave an added dimension to the interviews. Field notes were written for unusual cases (e.g. when family circumstances were difficult).

To guard against the fragmentation of data through the 'code and retrieve' method of analysis, synopses of each woman's interview were composed by the research associate, providing an overview of all interview cases and read by the qualitative lead. A composite case study was produced to highlight the difficulties faced by patients (see Appendix 10).

TABLE 32 Summary of stages of thematic analysis. Stages are drawn from Braun and Clarke's⁷⁹ sixfold typology of the phases of qualitative thematic analysis

Stages of thematic analysis	Project activities	Clinician data	Patient data
1. Familiarisation with the data	Initial reading of transcripts to identify possible key themes	RG all, ML sample	ML all, RG sample
	Data meeting to discuss provisional themes to guide coding	RG, ML	ML, RG
2. Coding the data	Coding of transcripts	RG	ML
	Refining of codes to (1) clarify meaning, (2) remove overlaps and (3) integrate very small codes	ML, RG	ML, RG
3. Searching for themes	Mapping of codes to themes	RG with detailed feedback from ML	ML
4. Reviewing themes	Initial refinement of themes in relation to coded data	RG	ML
	Team review of themes emerging from the data	RG, ML, SCR	ML, RG, SCR
5. Defining and naming themes	Data meetings to consolidate the review of the thematic framework	RG, ML, SCR Includes some feedback from dissemination workshops (various participants)	ML, RG, SCR Includes data meeting feedback (ML, RG and SCR with MERIDIAN study members, including the statisticians)
	Further reduction of themed data and adjustment of subthemes	RG, ML	ML, RG
	Final content review of themed data, refine when necessary	Initially RG, ML and SCR, then the full MERIDIAN study team, particularly GM and CM	ML
	Contrast findings with existing literature to draw out and develop discussion points	RG, with input from ML and SCR	ML, RG
6. Reporting the data	Extract key themes for the Health Technology Assessment report	RG	ML
	Finalise report chapter and discussion of themes across data sets	RG in collaboration with ML, SCR and the wider MERIDIAN team	ML in collaboration with RG, SCR and the MERIDIAN study statisticians

Interviews with health professionals were, on occasion, challenging for pragmatic reasons. Participants were usually interviewed at work, where they were busy and often interrupted. However, even the most fragmented interviews retained a coherent narrative structure and detailed discussion, producing a rich, complex data set. The interviews included transparent discussion of the context leading to the interview and field notes were produced only as needed (e.g. when disruption had been experienced).

Results

Service user substudy participants

The questionnaire survey was conducted between September 2011 and August 2014 and included the first 681 participants recruited to the MERIDIAN study (questionnaire 1) and 630 participants who were eligible for the follow-up questionnaire (questionnaire 2). The response rate was higher for questionnaire 1 (428/681, 62.8%) than for questionnaire 2 (284/630, 45.1%) ($p = 0.005$). A total of 206 participants completed both questionnaires. Respondents' characteristics are shown in *Table 33*. The proportion of respondents to questionnaire 1 who opted for TOP was 18%, in contrast to the 11% of questionnaire 2 respondents who opted for TOP.

In total, 104 EOI forms were returned with survey 2. Of these, 10 respondents were not contactable and 49 respondents were withdrawn because their forms were returned after recruitment was completed. Those eligible for the qualitative interview phase were contacted by telephone. Three respondents subsequently declined and data from one interview was prematurely deleted, leaving data from 41 participants available for analysis (39% of EOIs returned, 14% of surveys for questionnaire 2 completed).

TABLE 33 Characteristics of substudy patient respondents

Characteristics	Questionnaire 1 (<i>N</i> = 428)	Questionnaire 2 (<i>N</i> = 284)	Questionnaire 1 + questionnaire 2 (<i>N</i> = 206)	Interview (<i>N</i> = 41)
Age (years), mean (SD)	29.3 (6.0)	29.8 (6.0)	30.1 (5.9)	31.3 (6.0)
Education (highest qualification), <i>n</i> (%)				
GCSE	172 (40.2)	121 (42.6)	91 (44.2)	4 (9.8)
A level/NVQ	167 (39.0)	119 (41.9)	84 (40.8)	17 (41.5)
Degree/N/A	42 (9.8)	32 (11.3)	25 (12.1)	19 (46.3)
Gestation at iuMRI (weeks), mean (SD)	24.2 (4.1)	24.5 (4.4)	24.1 (4.2)	23.4 (4.6)
MRI centre, <i>n</i> (%)				
Host institution	280 (65.4)	192 (67.6)	139 (67.5)	27 (65.9)
Other sites	148 (34.6)	92 (32.4)	67 (32.5)	14 (34.1)
Diagnosis, ^a <i>n</i> (%)				
Isolated VM	195 (45.8)	120 (42.4)	98 (47.8)	20 (49.0)
VM and other pathology	93 (21.8)	75 (26.5)	42 (20.5)	13 (32.0)
Pathology not including VM	138 (32.4)	88 (31.1)	65 (31.7)	8 (20.0)
Pregnancy outcome, ^b <i>n</i> (%)				
Livebirth	359 (84.5)	220 (77.7)	175 (85.5)	30 (73.0)
Death	20 (4.7)	12 (4.2)	6 (2.9)	4 (10.0)
TOP	44 (10.4)	50 (17.7)	23 (11.2)	7 (17.0)

A level, advanced level; GCSE, General Certificate of Secondary Education; N/A, not available; NVQ, National Vocational Qualification.

^a Missing data: Q1 (*n* = 2), Q2 (*n* = 1), Q1 + Q2 (*n* = 1).

^b Missing data: Q1 (*n* = 5), Q2 (*n* = 2), Q1 + Q2 (*n* = 1).

Patient survey findings

A total of 97% of questionnaire 1 respondents and 95% of questionnaire 2 respondents would choose to have iuMRI as part of their care if they were ever in a similar situation again. In response to the question about whether or not iuMRI was undertaken at an acceptable time and place, 73% of questionnaire 1 respondents and 68% of questionnaire 2 respondents strongly agreed or agreed. Only 3% of respondents in both groups strongly disagreed and 9% of questionnaire 1 respondents and 11% of questionnaire 2 respondents disagreed. The following open-text responses in the survey are illustrative:

We didn't mind travelling a long distance there, but the return journey was extremely difficult after being told such heart-breaking news.

ID239

... we had to stay overnight in a hotel with a 7-year-old, hours away from home. It would be very helpful if this test was more widely available.

ID125

Responses to questions about the impact of iuMRI on understanding and decision-making are shown in Figure 5. Overall, 83% of respondents to questionnaire 1 and 80% of respondents to questionnaire 2 gave positive responses to the statement 'The information from the iuMRI scan helped me to understand the baby's problem'. Positive responses were more likely if the iuMRI was performed at Sheffield compared with non-Sheffield sites [questionnaire 1 (OR 2.22, 95% CI 1.28 to 3.86; $p = 0.005$) and questionnaire 2 (OR 3.41, 95% CI 1.82 to 6.48; $p < 0.001$)] and less likely if the iuMRI had been performed after 24 weeks of pregnancy [questionnaire 1 (OR 0.52, 95% CI 0.29 to 0.91; $p = 0.022$) and questionnaire 2 (OR 0.39, 95% CI 0.20 to 0.750; $p = 0.005$)].

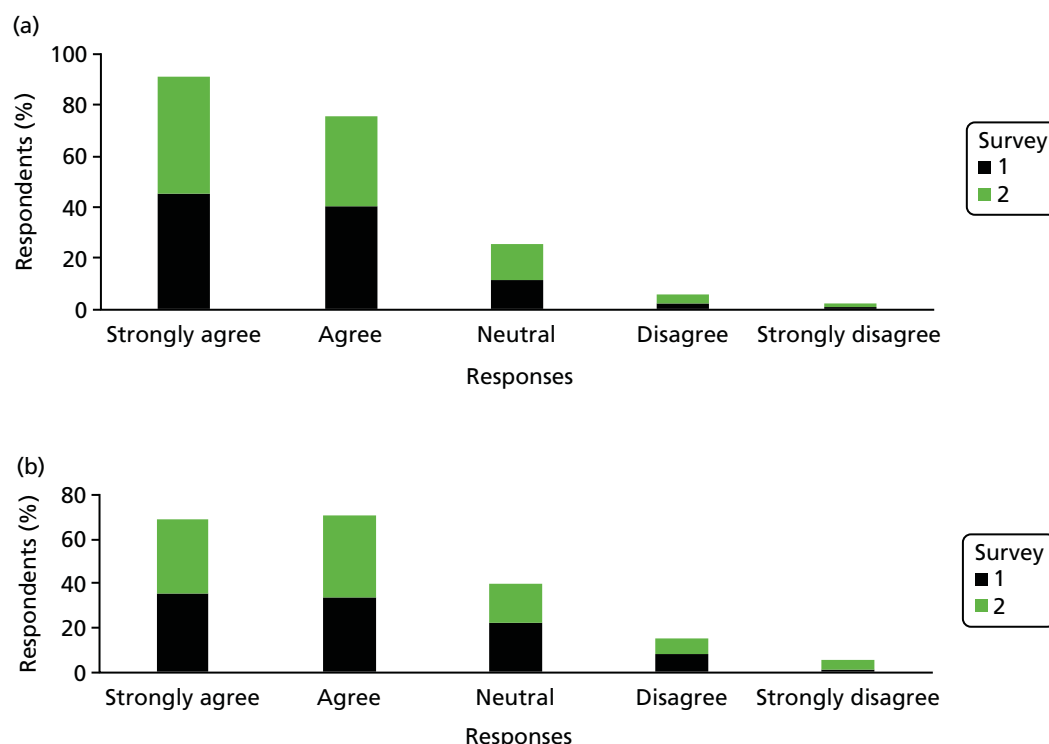


FIGURE 5 Responses to questions about the impact of iuMRI on understanding and decision-making in questionnaire 1 (black) and questionnaire 2 (green). (a) Responses to the question 'The information from the MRI helped me to understand the baby's problem'; (b) responses to the question 'The information from the MRI helped me to understand how the baby's problem could affect his/her future quality of life'; and (c) responses to the question 'The information on the MRI scan was useful to me when I made the decision about whether or not to continue with the pregnancy'. Of the 206 women who responded to both surveys, the proportion of those who either strongly agreed or agreed to (a) were 73% (Q1) and 68% (Q2), to (b) were 73% (Q1) and 68% (Q2), and to (c) were 76% (Q1) and 72% (Q2). (continued)

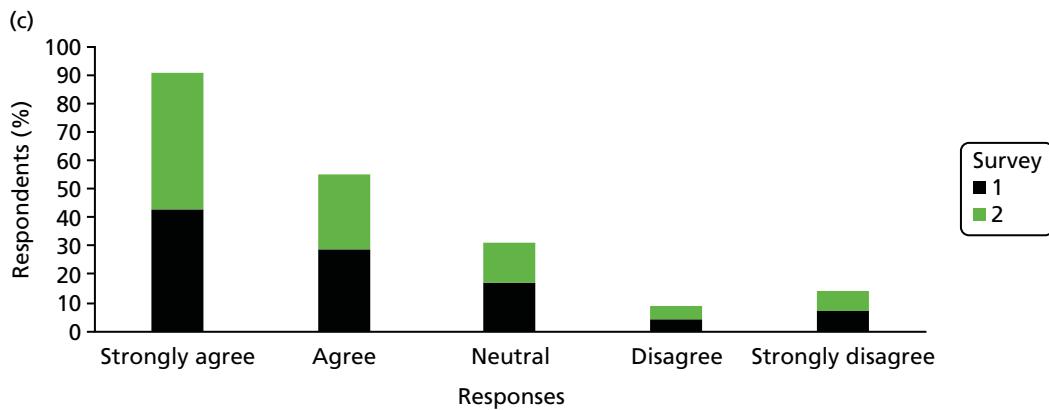


FIGURE 5 Responses to questions about the impact of iuMRI on understanding and decision-making in questionnaire 1 (black) and questionnaire 2 (green). (a) Responses to the question ‘The information from the MRI helped me to understand the baby’s problem’; (b) responses to the question ‘The information from the MRI helped me to understand how the baby’s problem could affect his/her future quality of life’; and (c) responses to the question ‘The information on the MRI scan was useful to me when I made the decision about whether or not to continue with the pregnancy’. Of the 206 women who responded to both surveys, the proportion of those who either strongly agreed or agreed to (a) were 73% (Q1) and 68% (Q2), to (b) were 73% (Q1) and 68% (Q2), and to (c) were 76% (Q1) and 72% (Q2).

A total of 69% of respondents to questionnaire 1 and 71% of respondents to questionnaire 2 gave positive responses to the statement ‘The information from the iuMRI scan helped me to understand how the baby’s problem could affect his/her future quality of life’. Positive responses to questionnaire 1 were more likely when the prognosis was graded as poor (i.e. < 50% chance of a normal outcome) (OR 4.20, 95% CI 1.64 to 11.10; $p = 0.003$). Positive responses to questionnaire 2 were more likely if iuMRI was performed at the Sheffield site (OR 1.96, 95% CI 1.15 to 3.35; $p = 0.013$).

Of the questionnaire 1 respondents who answered either one of the previous two questions (see *Figure 5a* and *5b*), 37% of them did not respond to the statement ‘The information from the iuMRI scan was useful to me when I made the decision about whether or not to continue with the pregnancy’ (see *Figure 5c*). In total, 71% of the remaining respondents gave positive responses to the statement. None of the factors included in the model influenced the response to questionnaire 1. Of the questionnaire 2 respondents who answered either one of the two previous questions, 27% did not respond to the same statement but 74% of the remaining questionnaire 2 respondents responded positively.

Women who were not faced with the dilemma of whether to have a termination because the condition was ‘mild’ or were going to continue regardless of what the iuMRI scan revealed might have opted to leave the question blank because it was not ‘relevant or appropriate’. The only factor that influenced the response to the statement about whether or not the iuMRI scan was useful in decision-making was gestational age at iuMRI; fewer mothers who were scanned at ≥ 34 weeks of pregnancy gave positive responses (OR 0.11, 95% CI 0.02 to 0.47; $p = 0.007$).

Overall, 90% of questionnaire 1 respondents rated their care as ‘excellent’ or ‘very good’, 9% of respondents rated their care as ‘good’ and only 1% of respondents rated their care as ‘fair’. Responses to questionnaire 1 did not appear to be influenced by any of the factors examined and responses were consistently positive across groups. Overall, 88% of questionnaire 2 respondents rated their care as ‘excellent’ or ‘very good’, 9% of respondents rated their care as ‘good’ and 3% of respondents rated their care as ‘fair’. Only one respondent to questionnaire 2 rated their care as ‘poor’. Responses to questionnaire 2 were only influenced by journey length; positive responses were less likely if the journey length was 2–4 hours (OR 0.38, 95% CI 0.13 to 0.99; $p = 0.05$).

Patient qualitative interview findings

The sample of 41 women interviewed included experiences from three of the six iuMRI sites. A total of 18 family members were included: 16 family members were in joint interviews and two were individually

interviewed after their partners. Four of the women's relatives were their mothers. Interviews were carried out between July 2012 and December 2013, 4–12 months post pregnancy outcome and with the majority > 7 months post outcome. Interviews lasted between 21 minutes (with a family member only) and 146 minutes (with the participant). The average length of the interviews was 69 minutes. Four key themes were identified from the data.

Theme 1: diagnosis of a fetal brain abnormality

The most common fetal brain abnormality was VM. The diagnosis had an impact on patient experience because fetuses with VM often had more frequent ultrasonography monitoring of ventricle size (some patients had an additional iuMRI scan). For most parents, the iuMRI scan was only one appointment among many others along the care pathway:

... when it was diagnosed it was 10 and a half and 12 and a half or something like that [ML: that's not too] you know one was mild and one was moderate [ML: 'yes, yes'] erm so the next time one was moderate and one was heading towards severe you know it was [ML: 'uhuh'] and when, we went once and it was they were both over 15 [both left and right] both left and right were both over 15 and I was, I cried at that one because I knew that, that was severe and it wasn't decreasing it was getting worse . . .

P9

In cases in which the fetus had other pathologies, there were referrals to other specialists to clarify the infant's future quality of life with the condition. Many such cases were serious brain abnormalities involving structural defects that often led to a TOP or a continuation of pregnancy expecting a non-surviving infant or an infant with disability. The data suggested that iuMRI could help parents adjust and prepare for the probable outcome, but uncertainties about prognoses could be disappointing for those who had hoped that the new technology could provide more definite answers:

When we went . . . to the MRI scanner they still weren't sure what they [cysts] meant [Family member: 'no'] so they found the corpus callosum on the, on the MRI which just confirmed that it really hadn't disappeared which was good . . .

P25

The interview accounts show that participants benefited from clear explanations from consultants who drew on different resources in communicating the diagnosis (e.g. analogies, drawings, textbooks, models and scan images). The iuMRI images were seen as a useful source of information and explanation, but much depended on how image interpretation was communicated to the women and families. At the Sheffield site, parents could view the iuMRI images with the radiologist after the scan and found this process to be beneficial. Other patients who were unable to view the images did not express much dissatisfaction; a small number of patients were concerned about the waiting time or results.

Theme 2: experiences of in utero magnetic resonance imaging technology

A majority of iuMRI scans were conducted at the Sheffield site (27/41), but experiences of the technology varied between and within sites. Scans took between 30 minutes and > 1 hour, depending on how well the individual woman tolerated the scan and on fetal activity. Women understood the point of undergoing iuMRI and for many it was an especially anxious time. In most women's accounts, their discomfort seemed to stem from being pregnant more than from the noise, heat and the confined space in the machine. Those who described negative experiences in the machine were mostly those patients in advanced stages of pregnancy or with a high body mass index:

I mean I was like 30-odd weeks pregnant and I was massive and I got, I got crushed in there so when I come out I was crying . . .

P23

Larger patient size led to more discomfort and could result in panic attacks. Most women tried their best to keep as still as possible to try to settle the fetus for the scan to be completed. Many described resisting urges to come out and tolerated any discomfort or pain to avoid delaying the process. The noise from the machine was experienced as particularly loud and some women worried about their baby's hearing. In general, women were willing to undergo discomfort because of their concern for their baby.

Women appreciated being made comfortable and properly briefed about the procedure before and during the scan via headphones. Several spoke about the support from their partner's presence during the scan. Those women and family members who were able to view the iuMRI images after the scan reported positive reactions and some experienced surprise and, in one case, shock. Most women gave accounts that expressed amazement and fascination at the image clarity compared with ultrasonography images. Participants felt that the iuMRI images facilitated detailed explanations about the fetal brain abnormality at follow-up. Those who had the opportunity to view the iuMRI images as a video clip found it incredible and some described it as 'surreal'. The possibility of seeing the video clip was also a turning point for those parents facing an unexpected 'reality':

... like seeing that video when they put the frames together [ML: yeah] was just so amazing it really made you feel like actually there is a real little baby in there ... it's not just ... the ventriculomegaly ...

P38

Theme 3: decision-making about a pregnancy at risk

Before considering the iuMRI, many participants experienced amniocentesis as a major crossroads, and some participants decided to terminate the pregnancy after a trisomy 21 diagnosis. For other participants continuing with pregnancies affected by trisomy 21, iuMRI images provided more detail of the brain and assisted in the decision-making. Women who had had the amniocentesis and had normal results and those who refused seemed more reliant on the iuMRI imaging than other patients:

... it would be MRI we got told it was all, he was all fully formed sort of and everything seemed to be developing [F: yeah], sort of as much as it should be for that stage of pregnancy, we decided not to [F: so we decided it (amniocentesis) wasn't worth the risk] ...

P11

Following the identification of a developmental brain abnormality, most participants experienced an anxiety-ridden journey of abnormality monitoring. In VM cases, definitions of abnormality hinged on ventricle measurements; iuMRI could be quite significant in parents' decision-making in relation to what the measurements meant. The clarity and detail provided in iuMRI images were particularly significant for parents facing complex abnormalities that they could not quite comprehend. Viewing the iuMRI images helped parents feel that they understood the abnormality and its possible consequences and helped them come to terms with their baby's condition. This feeling of enhanced understanding seemed to have a strong impact on parental decision-making about their 'at-risk' pregnancy.

The timing of abnormality identification coinciding with fetal movements made it particularly difficult for some mothers, heightening their desire for swift iuMRI results. Parents contemplating termination worried about their baby's future quality of life. In cases when disability would likely result from the brain abnormality, the degree of severity was important. In these instances, iuMRI appeared important in influencing the tipping point of parental decision-making. Some accounts suggested that iuMRI added to parents' understanding of the structural features of the brain:

... the clarity that, that the MRI scans can show is just and the detail, the level of detail they can see is amazing it really is ... I think it really hit it home for us because you could literally you know he showed us the pictures of the brain and you could just see this big section missing, you know and he showed us where it should be and what it should look like.

P41

Theme 4: assessing the added value of in utero magnetic resonance imaging

In functional terms, participants understood that iuMRI technology could capture hundreds of images in one scan session from various angles regardless of fetal position. The technology was also understood as capable of scanning women with a higher body mass index. Participants saw iuMRI as producing clearer images, representing comprehensive visualisations of the abnormality (compared with ultrasonography). For this reason, many parents felt that iuMRI could provide additional information to ultrasonography. Because parents recognised the importance of minute change in measurements, especially in VM, they appreciated what they perceived as the greater accuracy and reliability of the information provided by iuMRI. Nevertheless, some participants acknowledged the value of ultrasonography in identifying initial problems and viewed these different technologies as being complementary.

For some parents, the perceived clarity and precision of iuMRI images reinforced their trust in the patient–doctor relationship. Seeing structural defects with such clarity and establishing severity prepared (or helped) them to think through either ending the pregnancy or facing the continuing pregnancy and the possibilities at and beyond birth. However, parents who felt that the iuMRI had insufficiently prepared them spoke of the shock of discovering an abnormality (or a more severe abnormality). This mother had not felt prepared when her baby was born with trisomy 21:

ML: And do you think it's all worth it going through all that?

P12: No not really cos we didn't get anything found out. We didn't, we don't, we didn't find what's wrong with him . . . it were a bit like a false sense really that everything's going to be all right, cos I, by this point I was like, he's going to be all right.

The following response from the open-text section of a survey response also expresses a more negative view of iuMRI:

The MRI scan showed additional problems that the ultrasound did not show – problems with the cerebellum. I went for the MRI scan hoping for reassurance, but all I got was bad news. Problems with the brain cannot be 'fixed' – therefore it should have been made clear that if we weren't intending to terminate then we shouldn't have the scan, as nothing could be done except us worry.

ID100

Some parents felt that iuMRI merely confirmed/disconfirmed the ultrasonography findings, but some did report relief and reassurance from the iuMRI findings. Those expecting a baby with a disability were reassured when the extent of the disability was less serious than expected; others found that the iuMRI images confirmed severity and so had greater confidence in their decision to end the pregnancy. Women's accounts reported that clinicians made it clear that expectations of the technology should be tempered with caution.

Health professional participant characteristics

From 30 health professional recruitment packs, 23 (77%) returned a reply slip, of which 21 (70%) health professionals were willing to be interviewed (50% of the originally identified sample). In round 1, 11 were interviewed between March and August 2013. In round 2, 19 were interviewed [six face to face, 12 over the telephone and one via Skype™ (Microsoft Corporation, Redmond, WA, USA)] between July 2014 and March 2015. In round 1, interviews lasted between 41 and 89 minutes (average length of 63 minutes). In round 2, interviews lasted between 26 and 62 minutes (average length of 44 minutes).

A total of 21 health professionals were recruited and interviewed at least once. Two participants were interviewed in round 1 only, seven participants in round 2 only and nine participants were interviewed in both. Participants' characteristics are described in *Table 34*. Sites have not been included to protect

TABLE 34 Characteristics of substudy health professional participants

Study ID	Gender	Specialism	Site type	Round 1	Round 2
HP01	Male	Fetal medicine	Tertiary	Yes	Yes
HP02	Male	Rad	Tertiary	Yes	Yes
HP03	Male	Fetal medicine	Tertiary	Yes	Yes
HP04	Male	Fetal medicine	Tertiary	Yes	Yes
HP05	Male	Fetal medicine	Tertiary	Yes	Yes
HP06	Male	Fetal medicine	Tertiary	Yes	Yes
HP07	Male	Rad	Tertiary	Yes	Yes
HP08	Male	Rad	Tertiary	Yes	Yes
HP09	Female	Obstetrics/fetal medicine	Non-tertiary	Yes	Yes
HP10	Female	Rad	Non-tertiary	Yes	Dropped out
HP11	Female	Fetal medicine	Non-tertiary	Yes	Not approached
HP12	Male	Fetal medicine	Tertiary	Not approached	Yes
HP13	Male	Obstetrics/fetal medicine	Non-tertiary	Not approached	Yes
HP14	Female	Obstetrics/fetal medicine	Non-tertiary	Not approached	Yes
HP15	Female	Other	Non-tertiary	Not approached	Yes
HP18	Male	Rad	Tertiary	Not approached	Yes
HP20	Female	Rad	Tertiary	Not approached	Yes
HP21	Female	Fetal medicine	Tertiary	Not approached	Yes
HP22	Male	Other	Tertiary	Not approached	Yes
HP23	Female	Other	Tertiary	Not approached	Yes
HP24	Female	Rad	Tertiary	Not approached	Yes
<i>n</i> = 21	12 male, nine female	Eight fetal medicine only, seven Rad only, six other and obstetrics/fetal medicine	15 tertiary and six non-tertiary	<i>n</i> = 11	<i>n</i> = 19

Rad, radiography/radiology.

participant anonymity in this close-knit professional field. However, participant recruitment covered 10 clinical sites and included sites with diverse provision structure (tertiary and non-tertiary care, those providing iuMRI on site and those referring elsewhere).

Most participants fell within the fetal medicine and radiology specialties, but a number of participants were in other categories (e.g. midwifery, radiography) or did not fit neatly into one category. To ensure the confidentiality of participants in smaller subgroups, some specific categories are not identified.

Health professional qualitative interview findings

Theme 1: in utero magnetic resonance imaging in the health-care context

Health professionals using imaging techniques often enjoyed the intellectual challenge of imaging, and the threshold competency for image interpretation is high. Producing and interpreting fetal images was described as intuitive, especially in ultrasonography. Understanding the skill and experience involved in

imaging was seen as important in highlighting the consequences of lack of expertise for both individuals and newly emerged fields. Lack of expertise was considered problematic:

The harm is the misinterpretation both ways, so misinterpreting abnormal as normal but misinterpreting normal as abnormal.

HP10, R1

Participants drew distinctions between radiologists' interest in the clarity of anatomical structure compared with fetal medicine specialists' interest in images to understand fetal development. Radiologists in our sample had regular patient contact but not all radiologists do. In other contexts, participants felt that this could cause problems in MDT working when a radiology clinician 'says no' to a MRI request (to the treating doctor, not to the patient). Radiology services face consistently increasing demand for MRI from other medical specialties. Although there has been some increase in capacity, it has not kept pace with growing demand.

Fetal medicine provision is primarily within specialist centres staffed by subspecialists, but some obstetricians with specialist ultrasonography training provide high-risk work beyond specialist centres. Clinical practice is therefore dependent on effective referral networks. Referral practices had implications for how (and where) iuMRI was integrated in clinical pathways. Further layers of specialisation were apparent within established fetal medicine centres. Views on the level of specialist knowledge needed for PND work varied, often reflecting participants' current role; those participants with subspecialty training felt that fetal medicine was poorly understood by those outside the subspecialism.

Clinicians saw iuMRI as useful in PND, partly because of what could be 'seen' (e.g. symmetry of iuMRI images compared with fetal ultrasonography). However, clinicians mainly noted the impact on counselling. Both fetal medicine and radiology clinicians felt that iuMRI affirmed (rather than changed) provisional ultrasonography diagnoses. Caveats to iuMRI utility included movement artefact and moving body parts (e.g. the heart).

Interviewees compared past developments in fetal ultrasonography with more recent developments in iuMRI. The expansion of fetal ultrasonography (and the migration of fetal imaging expertise from radiology to fetal medicine) was described as a period of expansion and transition:

... very rapidly during [the late 1980s] it became apparent that obstetricians were learning obstetric ultrasound and radiologists were beginning not to ...

HP07, R1

Participants' views about contemporary fetal ultrasonography provision drew on understandings of fetal ultrasonography as an established but dynamic practice. Some clinicians felt that ultrasonography quality had improved because of iuMRI feedback, whereas others worried that the 'value added' of iuMRI reflected, in part, decreasing expertise in ultrasonography.

Most clinicians described incremental change in the context of iuMRI, including the more commonplace lay experience of MRI, increasing data capture speed (shortening scan times) and image generation developments. The significance of technological change in data capture and image processing was less clear for clinical practice and patient experience in PND.

Developments in iuMRI practice reflected two key influences: (1) evidence-based policy and (2) managing uncertainty. Formal research evidence was crucial, with clinical conferences providing a conduit for sharing up-to-date information. Participants' comments on evidence referred to concerns about uncertainty in developing new practices given the impact that findings may have on women's decision-making about pregnancy. More material concerns included the funding of iuMRI provision and the timing of provision for service users. Service development was conceptualised as incremental. In geographical terms, iuMRI

provision had developed where there was sufficient demand, goodwill, professional interest and collaborative working. Trust developed in teams providing care for parents making 'high stakes' decisions about a pregnancy. This trust-building process centred on the reliability of interpretation:

... I don't know of any case where [name] has made a diagnosis that's been subsequently proved to be wrong.

HP21, R2

Theme 2: in utero magnetic resonance imaging as added value in the prenatal diagnosis for brain abnormality

Clinicians responsible for patient counselling saw specific advantages in iuMRI (e.g. improved imaging of structures like the CC and for patients with a higher body mass index). Most comments related to the perceived impact on counselling patients, sometimes because iuMRI confirmed the absence of additional problems, improving confidence in the ultrasonography diagnosis. Alternatively, iuMRI information differed from ultrasonography findings, clarifying either fetal normality or abnormality. Clinicians also saw significant limitations, some of which were practical (e.g. inability to fit in the scanner, additional appointments, more travelling). There were concerns that the possibility of iuMRI might create demand even when clinically unnecessary. Participants consistently noted that iuMRI was useful, but as an adjunct to, not a replacement for, ultrasonography.

Clinical decisions about management options were considered more straightforward for severe abnormalities but more difficult when fetal abnormality lay close to the bounds of normality. Clinicians finalising the diagnosis and the counselling of parents took reassurance from iuMRI as a second opinion (either from the radiologist reporting the images or the modality). Perceived benefits included increased patient confidence, increased clinician confidence/learning and reduced clinician isolation:

... I think if you're making decisions based on life or death ... you want to make sure you're making, you feel much, much, much, much, much happier making those recommendations or those choices knowing that, that, that you're not in isolation. Making those decisions before MRI I think er, was a quite an isolating and lonely place.

HP4, R2

The uncertain risk associated with some anomalies made the lack of follow-up data on prognosis problematic, but most clinicians saw iuMRI as helpful in refining diagnoses (if not prognoses). Limiting possible outcomes was seen as helpful.

Clinicians compared the views of the brain offered by the two modalities and the physical processing underpinning the imaging. Some felt that iuMRI was better than ultrasonography for viewing some specific structural brain abnormalities. Participant HP06 (R1) offered a threefold typology of utility: (1) ultrasonography suggests an abnormality but iuMRI finds none, (2) iuMRI gives additional information about an ultrasonography-identified abnormality and (3) iuMRI identifies an abnormality not seen on ultrasonography (e.g. polymicrogyria).

Imager–interpreter relationships were also compared. Ultrasonography interpreters look at the image while producing it; in iuMRI, the imager–interpreter relationship can be distant. This had an impact on participants' conceptualisations of quality and clinical applicability. Some felt that ultrasonography's interpreter-as-imager model improved quality, whereas others felt that double reporting (common in MRI) was important. Modality also had an impact on clinician–patient relationships; some clinicians felt that iuMRI images were more understandable to parents and others felt that ultrasonography's 'real-time' image was better than an in utero magnetic resonance image 'snapshot'. Participants discussed advantages and disadvantages of ultrasonography's perceived role in building rapport with pregnant women and perceptions of safety and cost. Physical safety was seen as relatively clear for iuMRI, but cost was more problematic. Economic viability required ongoing dominance of ultrasonography, with iuMRI as an adjunct.

Fetal medicine specialists utilised iuMRI in different ways. Digital magnetic resonance images are transportable, but many clinicians focus on only the textual report. Those with MDT experience generally found MDT discussion helpful. Some clinicians felt that iuMRI reports could be more transparent about normality:

So just as you would have an ultrasound scan of a baby with a checklist of things that you say look normal or don't look normal, to have that for MRI would be very useful. So an MRI report that says 'brain appearance is normal' you think OK fine but that wouldn't cut it with an ultrasound . . .

HP03, R2

Some clinicians providing fetal medicine counselling did not use the full details, instead looking for the 'nitty gritty towards the end of the report' (HP09, R2), whereas most clinicians saw value in structured reporting to raise standards and as the basis of effective dialogue between clinicians:

. . . [the iuMRI team] have a very open approach in terms of getting hold of people that could explain things better before I can explain it to the patient.

HP09, R1

Some clinician–patient encounters involved viewing images, but most fetal medicine specialists focused on the iuMRI report to inform counselling. The iuMRI images were seen as potentially useful but not essential and potentially problematic owing to lack of training in image interpretation. Participants identified the ongoing and developing dialogue between radiology and fetal medicine as crucial for good-quality care. Similar discussions were developing in other areas of fetal imaging (e.g. placenta) where cross-specialty discussion of images was considered productive.

Theme 3: acceptable practice in providing in utero magnetic resonance imaging

Participants identified accessibility of services and perceived patient acceptability as key issues in acceptable professional practice.

One aspect of accessibility is the location of the iuMRI scanner relative to the counselling site (same site, local or regional/national travel). Participants reported that most women would travel but some would not because of cost, inconvenience or wanting to be near home. Distance was balanced against service quality and added value. Local provision was often considered more 'ideal', but some clinicians noted that this might increase referral rates.

Accessibility of expertise was important. Time pressures in PND require swift access to specialised provision (both image production and interpretation). Avoidable delay was considered highly problematic:

. . . pregnancy is very time sensitive and there's decisions that parents make they need to have all the information to make the decisions . . . some patients had to wait for, in my opinion, for too long to get the MRI.

HP23, R2

Key indicators of acceptability for iuMRI included lay demand and ease of recruitment to the study. Increased demand for iuMRI was seen as reflecting an increased demand for MRI in general. Expectations sometimes required management, especially when women requested iuMRI when it was considered not clinically justified. For some, the social characteristics of the population, or remote location, meant that acceptability was less straightforward. Most professionals felt that their patients were accepting of MRI technology, but some patients had reservations about the scanning process because of claustrophobia or being pregnant and uncomfortable. Clinicians felt that patients appreciated the additional information, with hunger for information driving patient demand.

High-quality provision featured strongly in clinicians' accounts of iuMRI in the PND for brain abnormality. Acceptability was seen as underpinned by pragmatic and standards issues. Pragmatic issues often included (in)adequate funding structures and (in)sufficient work volume to develop image interpretation expertise. Clinicians sharing experience and standardised reporting were seen as important in raising technical standards. Similarly, double reporting was considered beneficial in managing the uncertainties that might stem from imaging or from variation in clinicians' reporting experience. Lone provision was problematic (e.g. annual leave, illness) and collaborative working was strongly endorsed:

... the worst thing that can happen is that the patient comes along, has the scan, you just reporting it, it disappears off in to the nowhere [RG: yeah] with no knowledge of who's receiving that whether anyone's bothered to read it.

HP17, R2

Standardisation was important, but clinicians explained that some variation in practice is inevitable and acceptable in good-quality care at the individual level (e.g. relative levels of consultant–patient contact) or at the unit level (e.g. referral practices). Variations in ultrasonography practice were framed around skill levels (e.g. routine vs. specialist ultrasonography). Most comments about iuMRI variation centred on how reports were written:

... the reporting length we get say from [specific site] even if it's just telling us lots of things are normal just feels much more reassuring to us as clinicians than the very brief summary we sometimes get from our own units ...

HP03, R2

Theme 4: enablers for future service development

Magnetic resonance imaging is just one option for radiology specialists, and iuMRI is one small element of MRI activity. In a similarly niche area of fetal imaging, some participants noted that the low volume of placental iuMRI did not justify funding a specific post (HP18, R1). Specialisation recurred in professionals' accounts, suggesting a complex map of perceived authority to speak: 'I think you should always stick to what you're good at and not dabble' (HP10, R1). Clinicians understood their work in terms of specialist training, highlighting the advantages of team working. Notions of expertise featured in fetal medicine and radiology accounts, but as expertise in fetal imaging is more peripheral in radiology, regulated pathways to expertise in fetal imaging were less developed. The lack of clear accreditation pathways to train the next cohort of highly specialised fetal imagers was considered problematic. More informally, networks featured strongly in clinicians' accounts of iuMRI developments. Professional reputation really counted in this network; one clinician noted that '*... you learn which people often don't get it right*' (HP21, R2). Trust and reputation were crucial in professionals' accounts; knowing who you were referring your patients to (or receiving information from) was vital to good-quality provision:

... most clinicians the first thing they'll do in a radiology is, who's that, written the report ... you can automate a lot of things can't you but you can't automate that professionalism.

HP24, R2

Future iuMRI provision was discussed in terms of barriers to growth and influences on policy implementation, with appropriate expertise and resourcing as common concerns. Participants saw a future for iuMRI provision, but barriers to growth centred on difficulties securing sufficient staff with sufficient expertise, adequate resourcing and infrastructure development. Several interviewees commented on recruitment difficulties for radiology consultant posts.

Most clinicians linked problems associated with expanding provision to competing pressures for NHS resource, and the possible outsourcing of iuMRI image interpretation was a significant concern. Three key influences on policy change were identifiable in the data: (1) will professionals comply with new

expectations, (2) how can collaborative working be promoted and (3) what is the best balance between centralisation, geographical accessibility and value for money?

For changes in policy and provision to work, this demands engaged and willing clinicians. Guidelines and funding mechanisms were seen as a way of facilitating and encouraging clinicians to work with the evidence. Credible guidelines were seen as requiring meaningful consultation with clinician groups.

Fetal imaging in the PND for brain abnormality draws from both fetal medicine and radiology. Some clinicians saw it as inevitable that radiologists would learn more fetal medicine, others that fetal medicine specialists would learn iuMRI skills. However, training time for iuMRI is time not available for other aspects of subspecialist training. MDT working was commonly seen as a positive development, but much depended on getting the 'right' people in the room and confidence in the information presented. Working across specialty boundaries was evident but challenging; sharing information between organisations remained problematic (e.g. information transfer between NHS trusts). Most participants described a growing gap between radiology and fetal medicine. Some saw iuMRI as signalling the need to reinvigorate this interface. Collaborative working was seen as beneficial, but views differed on how this could be done. Some found communication via the iuMRI report effective on its own, others suggested that collaborative working needed more dialogue:

... it would be helpful if all the imagers could sit down together and learn exactly where we measure say the ventricles ... their atlas of normality is slightly different from our atlas of normality ...

HP24, R2

Service user demand for imaging technology was seen as a key component of future policy developments. Clinicians, aware of the need for cost-effective use of limited NHS funds, were mindful that cost-effectiveness is balanced with quality control and professional acceptability. Provision of iuMRI on a larger scale (e.g. available routinely for fetal brain abnormality diagnosis) required local interest and goodwill. Participants suggested that resolving key capacity issues would be critical if the MERIDIAN study prompted expectations of UK-wide provision. Capacity issues are not solely concerned with funds to pay for equipment and staff, but also the time lag in training sufficient additional specialised clinicians. Generating additional posts (and trained specialists to fill them) was a key concern for those thinking about how iuMRI should be developed in the UK.

The envisaged clinical groups needed to provide a good-quality service varied. Some felt that the primary responsibility for the provision of iuMRI would remain with the radiology community, whereas others felt that the expertise might migrate into fetal medicine. Future planning and investment for provision was seen as important. The bottom line for participants was avoiding the dangers of adopting a new technique in PND with a steep learning curve:

... there's gonna be the people around the country that are doing it ... or there's going to be lots of people trying to do it from ground zero and making all the mistakes that you make in your first 100, 150 cases erm. And that's how you get a new technique a bad name.

P08, R2

Chapter 5 The 2- to 3-year follow-up study

Introduction

Background

Although the main study demonstrated improved diagnostic accuracy and clinical impact of iuMRI, further considerations must be addressed to fully answer whether or not iuMRI should be embedded in routine clinical practice. First, the number of cases without an ORD was noted; a likely consequence of this is that some diagnoses will not have been picked up and will only become clinically apparent as the child develops. It is also possible that some incorrect diagnoses have been made. Thus, the 6-month follow-up used in the original study may include diagnoses that differ from the confirmed diagnosis later in life.

Perhaps more importantly, the main study focused purely on the diagnosis, whereas to parents the functional significance (i.e. motor, cognitive and behavioural) of the brain abnormality is of a greater interest. The increased diagnostic accuracy only partially addresses the question of whether or not MRI improves outcomes for the child and the family; outcomes in this context are very wide-ranging and include more timely medical interventions, such as appropriate specialist referrals or the early involvement of therapists, as well as improved understanding/knowledge, impacts on quality of life and parental anxiety. Following the diagnosis of a fetal abnormality, parents often want to hear all the possible outcomes⁸⁵ in order to build a picture of what those difficulties would be like to live with. Aspects such as the emotional impact of caring for a child with potentially serious disabilities and the implications for family life (financial and logistical) affect their perceived ability to be able to provide for the child's possible needs.⁸⁶ The follow-up study was intended to address these concerns by examining the predictive validity of antenatal diagnosis of a brain abnormality for longer-term functional outcomes over the course of life. Specifically, we evaluated whether or not a prognosis based on iuMRI offers an improvement on those prognoses made by ultrasonography alone through a longer-term follow-up of our cohort.

Study aims and objectives

There were three distinct projects within the follow-up study, the aims and objectives of which are detailed below.

Project 1

To estimate the diagnostic accuracy of ultrasonography and iuMRI using diagnoses collected over a 2- to 3-year time frame.

Project 2

To estimate the extent to which prognoses based on ultrasonography and iuMRI predict a neurodevelopmental outcome at 2–3 years.

Project 3

To assess the clinical significance of isolated, mild VM and whether or not prognoses that incorporate iuMRI are superior to those based on ultrasonography.

Methods

Participants and eligibility criteria

Study participants had undergone ultrasonography and iuMRI as part of the original MERIDIAN study during their pregnancy. During the original consent process, women were asked whether or not they were

happy to be contacted by the research team for future studies about their child's development. Participants were eligible if they:

- participated in the MERIDIAN study and had a surviving child aged ≥ 2 years
- underwent both ultrasonography and iuMRI during pregnancy as part of the MERIDIAN study
- consented to be contacted by the research team about future studies related to their child's development.

For children who were no longer alive, the date of death and the cause of death were obtained when possible. No contact was made with the family.

Children who were aged > 42 months (term corrected) were not eligible for an assessment with the Bayley Scales of Infant and Toddler Development, Third Edition (BSID3), because there is a ceiling for the upper age at which a child can be tested using this tool, but they did complete an age-appropriate Ages and Stages Questionnaire-3 (ASQ-3); the medical case notes were reviewed for additional diagnoses up until the child was 42 months of age.

Participants were excluded from the study if any of the following criteria were met:

- if the child born to a participant in the MERIDIAN study was no longer alive
- if the child was no longer in the care of the biological mother who consented to participate in the original MERIDIAN study
- the parent or guardian was unable to give informed consent
- the parent or guardian was unable to understand English (except in situations when another parent/guardian of the child could translate and provide consent)
- if the participant was withdrawn at any stage of the MERIDIAN study.

Participant tracking and recruitment

To assess eligibility prior to contacting the participant for the follow-up study, the research team completed a consent form audit to identify those participants who had consented to be approached about future studies regarding their child's development (optional question 7 in the original consent form). Of those who had consented, medical notes and available NHS systems were checked to assess eligibility and suitability for the study.

Section 251 Confidentiality Advisory Group approval (reference 15/CAG/0155) was obtained to allow the central research team to check that the child was still alive using, when applicable, the Health and Social Care Information Centre (now NHS Digital) Patient Tracking system.

Once participants were screened and deemed eligible, they were sent a letter of invitation and a consent form/reply form to the most recent known address. The ASQ-3 was posted out along with the consent form.

Outcome measures and assessments

Project 1

Once written informed consent was received, a review of the participant's medical case notes was completed. We collected information about any further aetiological tests performed and diagnosis made, as well as any follow-up care of the child and life events that may have had an impact on the child's development.

Project 2

For project 2, the MERIDIAN study children were invited to participate in a developmental assessment; the BSID3 comprises cognitive, language and motor components. This is a standardised developmental assessment tool validated in large numbers of children in both the USA and UK, and for which population-

normative values are described. It provides details on cognitive, motor and communication skills and is considered the best tool for assessing development in young children. The assessments were completed either in the local recruiting hospital or, when possible, in the child's home.

In addition to the BSID3, the Strengths and Difficulties Questionnaire (SDQ) and the MERIDIAN Gross Motor Skills Questionnaire, which is based on the Gross Motor Function Classification Score (GMFCS),⁸⁷ were completed. If a face-to-face appointment was not possible, then these two questionnaires were completed via post or telephone. In some cases, the case review undertaken within project 1 provided information about the child's development that could be used in place of an assessment.

Two paediatricians independently reviewed the data. The first step involved looking at individual assessments and their developmental domains. Each domain was categorised as 'normal', 'borderline' or 'abnormal'. For the BSID3, we considered that a composite score in any domain of > 85 (within 1 SD of the population mean) was normal, 70–85 (between 1 and 2 SDs below the population mean) was borderline and < 70 (2 SDs below the population mean) was abnormal. For the ASQ-3, a score in any domain in the 'above the cut-off point' category (within 1 SD of the population mean) was normal, 'near the cut-off point' (between 1 and 2 SDs below the population mean) was borderline and 'below the cut-off point' (2 SDs below the population mean) was abnormal. For the SDQ, normal was 0–12, borderline was 12–15 and abnormal was ≥ 16 .

The paediatricians independently reviewed all of the data from the developmental assessment and categorised each child's overall developmental outcome as 'normal', 'at risk' or 'abnormal'. Because the ASQ-3 is a screening tool with a well-described false-positive rate, whereas the BSID3 is a formal hands-on assessment, more weight was given to the BSID3. Therefore, a child who was below the cut-off line on the ASQ-3 but had abilities in the normal range in the BSID3 was considered normal. The data from the SDQ were used primarily when children were considered borderline between classifications. Therefore, when a child showed a normal outcome on ASQ-3 and BSID3 but the SDQ was raised, the child was categorised as having a normal outcome.

A child was classified as having an abnormal outcome when:

- the child's development was abnormal in one or more domain on the BSID3
- the child's development was abnormal in two or more domains on the ASQ-3 and no BSID3 data were present
- the child's development was abnormal in one domain and borderline in two or more other developmental domains on the ASQ-3 and no additional BSID3 data were available.

A child with cerebral palsy was not automatically classified as being abnormal because a spectrum of functional abilities and quality of life is possible for individuals with this diagnosis. When BSID3 or ASQ-3 scores for motor abilities were ≥ 2 SDs below the mean, or the GMFCS was ≥ 2 , the child was categorised as 'abnormal'. When the child's motor skills were within 2 SDs of the population mean, GMFCS was 1 and there were no concerns on the SDQ, the child was categorised as 'at risk'. This decision was made because during antenatal counselling families want clinicians to build a picture of their child's likely functional abilities, and a diagnostic label associated with a broad spectrum of outcomes is not helpful in clinical practice.

A child was categorised as 'at risk' if the child:

- was borderline in any developmental domain on the BSID3 (79–85)
- was borderline in two or more domains on the ASQ-3
- had a diagnosis of cerebral palsy but motor abilities on the BSID3 and/or ASQ-3 were normal and the GMFCS was 0 or 1.

The choice of words was important; it is possible that a child in the 'at-risk' group in our study could be in either the normal or the abnormal group later in life. The implications of developmental outcome data in antenatal counselling are that clinicians and women will use the data to make decisions on whether or not to continue the pregnancy. In this situation, it was inappropriate to label children as 'abnormal' when their scores fell into the normal population range and when they may have normal development in the future.

For the purpose of this study, participants who were 'normal' or 'at risk' were combined into a single group.

All reviews were undertaken blind to diagnoses as obtained from ultrasonography, iuMRI or ORD. Any cases for which there was disagreement were discussed by the clinicians at a panel review meeting and a final decision was made by consensus.

The following were assessed using the outcome classifications:

- The prognosis as reported by antenatal ultrasonography alone (i.e. prior to iuMRI) relative to the neurological development classifications calculated above.
- The prognosis as reported by iuMRI (following antenatal ultrasonography) relative to the neurological development classifications calculated above.
- The prognosis as reported by antenatal ultrasonography alone (i.e. prior to iuMRI) relative to individual components of neurological development.
- The prognosis as reported by iuMRI (following antenatal ultrasonography) relative to individual components of neurological development.

Project 3

The primary outcome of project 3 was the developmental outcome at 2–3 years (i.e. 'normal', 'at risk' or 'abnormal') as defined in project 2 for participants diagnosed on iuMRI with an isolated mild VM.

Sample size

The sample size was constrained by the number of MERIDIAN study participants consenting to further study. This did not affect project 1 for which the 6-month gestational age-corrected diagnosis was used if no follow-up was available, meaning that the original sample size calculation remained applicable. The sample size for project 2 was informed by the fact that there were almost 200 instances when the prognosis changed as a result of iuMRI, of which 38 are now classified as the poorest prognosis. It was also assumed that there would be approximately 400 surviving infants. Scaling these prevalences down, 400 cases would have a 90% power to detect a 20% increase in sensitivity and a 10% increase in specificity using the tests outlined above at a two-sided significance level of 5%.

Project 3

With approximately 140 cases and assuming the prevalence of poor outcome is indeed less than 10%, we will be able to estimate the prevalence to within a standard error of 2.5%.

Statistical methods

In project 1, the diagnostic accuracy was analysed as per the main MERIDIAN study, with the most recent diagnosis (original MERIDIAN study diagnosis or subsequent) used as the ORD.

For project 2, the sensitivity, specificity, PPVs and negative predictive values (NPVs) of iuMRI and of ultrasonography were reported. The sensitivity of each test was calculated for surviving children and also for non-surviving children who had died because of reasons other than TOP. For surviving children, the sensitivity was defined as the:

- number of participants with abnormal development and a prognosis of 'poor'
- number of participants with abnormal development and a non-missing prognosis.

The ability of ultrasonography and iuMRI to detect abnormalities associated with non-survival was defined as the:

- number of non-surviving infants and a prognosis of 'poor'
- number of non-surviving infants and a non-missing prognosis.

In order to provide an overall sensitivity averaged across both surviving and non-surviving infants, an adjustment was required to ensure that surviving and deceased infants were represented proportionately to their incidence. By definition, all deceased infants may be considered to have an 'abnormal' outcome and, therefore, a response rate of 1. Defining the overall uptake of project 2 among surviving infants by p , the sensitivity was therefore:

$$n_s + p \times n_d, \quad (1)$$

$$N_s + p \times N_d, \quad (2)$$

in which N_s infants were classified as having an abnormal prognosis and of whom n_s had received a poor prognosis; the total number of deceased infants at 2–3 years was N_d , of whom n_d were classified as poor prognosis.

The specificity was calculated for surviving children and was defined as the:

- number of infants with normal development and a prognosis of either 'favourable' or 'normal'
- number of infants with normal development and a non-missing prognosis.

The relative prognostic accuracy of ultrasonography and iuMRI was compared by calculating the difference in their respective sensitivities and specificities using McNemar's paired-sample chi-squared test. CIs were calculated using the Wilson method for simple proportions.⁸⁸

The analyses were repeated with children classified as 'abnormal' and 'at-risk' neurodevelopment considered as poor outcomes. Developmental scores (BSID3, GMFCS, SDQ and ASQ-3) were summarised using box plots.

To further understand the characteristics of those recruited to the study compared with the original cohort, specifically in relation to the prognosis on ultrasonography and iuMRI, simple bar charts were produced. The number of infants who were not approached, who had died within the original study, who were approached but did not consent and those who consented were presented by ultrasonography and iuMRI prognosis category. Waffle charts comparing the proportion within each prognosis category in the original cohort with those recruited into each project were produced. The number and type of deaths were also presented by prognosis in a bar chart.

Owing to the small numbers who had completed developmental assessment, no further subgroup analysis was conducted.

Analysis of project 3

Using the categories detailed in project 2 (i.e. 'normal', 'at risk' or 'abnormal'), the prevalence of severe and non-severe impairments was calculated together with exact binomial CIs. When the prognosis changed as a result of iuMRI, we assessed whether or not this was attributable to the iuMRI identifying further abnormalities.

Additional analyses

Ages and Stages Questionnaire versus Bayley Scales of Infant and Toddler Development, Third Edition

As an exploratory analysis, we assessed the diagnostic accuracy of the ASQ-3 in diagnosing neurodevelopmental impairment. The PPV, NPV, sensitivity and specificity were calculated using the BSID3 assessment as the 'reference standard' and 'abnormal' development or not as the binary classification. Sensitivity and specificity were presented with 95% exact binomial CIs.

Protocol amendments after study initiation

Details of substantial amendments submitted to the REC, which were important changes to trial methodology, are listed below.

Contacting participants

A barrier to recruitment was being able to make contact with potential participants; therefore, a number of amendments were implemented after reviewing recruitment throughout the study. First, the protocol was updated to allow the research team to attempt to make contact through text messages when a research mobile phone was available and to make a face-to-face approach if it was identified that the potential participant was attending a clinic appointment.

A further amendment was then implemented to allow the research team to include the consent forms in the initial invitation pack. The invitation letter was updated to clearly explain that the participant could contact the research team if they wished to ask any questions prior to completing the consent form. This approach was advised by the TSC to help improve the recruitment rate and reduce the burden on participants.

Child age ranges for outcome measures

During the study, the initial cut-off age for offering a participant a BSID3 assessment was increased from 38 to 42 months. The age for being eligible for an ASQ-3 was also increased from 42 months to 60 months to capture data on as many of the cohort as possible.

Results

Participant flow

Of the 829 participants from the MERIDIAN study who had both ultrasonography and iuMRI, 214 could not be approached (*Figure 6*). A total of 615 participants were screened and 39 were found to be ineligible at screening owing to the following: not wanting to be contacted ($n = 12$), being unable to confirm if the child was alive ($n = 3$), the child was no longer in the care of the mother ($n = 12$) and it was deemed to be inappropriate to contact ($n = 12$). A further 22 participants were excluded because a correct address could not be found to which an invitation letter could be sent. A total of 554 invitation letters were sent; 196 participants did not respond and 41 declined to enter the study. A total of 317 participants agreed to initial consent. Of these, 79 participants did not agree to formal consent.

Eight participants were screened in error as they did not complete iuMRI; two of these participants completed consent and a developmental assessment despite being ineligible. For simplicity, these cases have been excluded from the flow diagrams.

A total of 238 participants formally consented to project 1 (*Figure 7*); of these participants, medical case note review could not be undertaken for 28 children. Medical case note review was undertaken in 210 children and updated imaging was found in 81 cases. Of those with updated imaging, five cases had imaging that changed the original ORD and five had an ORD when they had not had one previously.

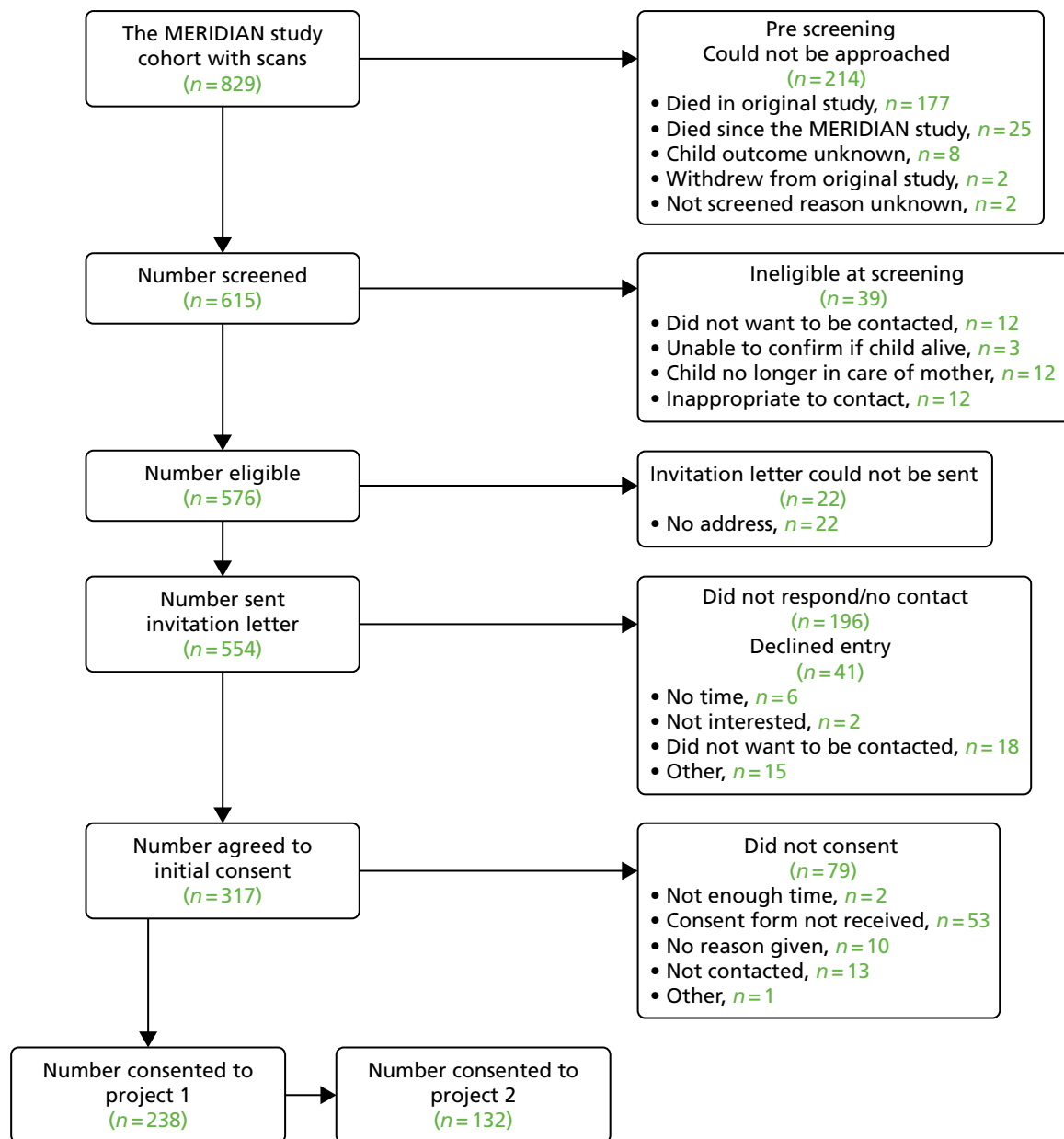


FIGURE 6 Participant flow in the MERIDIAN follow-on study.

When combined with the original MERIDIAN study data, 574 cases were included in the primary analysis and 643 cases were included in the secondary analysis.

A total of 132 participants formally consented to project 2 (Figure 8); 127 of these completed at least one type of developmental assessment. A further 52 participants completed the ASQ-3 as part of project 1. When combined, a total of 179 participants completed at least one developmental assessment. Eleven participants had relevant information in their medical case note review as part of project 1 that could be used by clinicians to provide an overall assessment of development. A total of 190 participants completed the clinicians' review. The primary analysis was carried out on these 190 cases and on 78 children who had died. Fifty-three children had died during the original MERIDIAN study and a further 25 children were known to have died since the end of the MERIDIAN study.

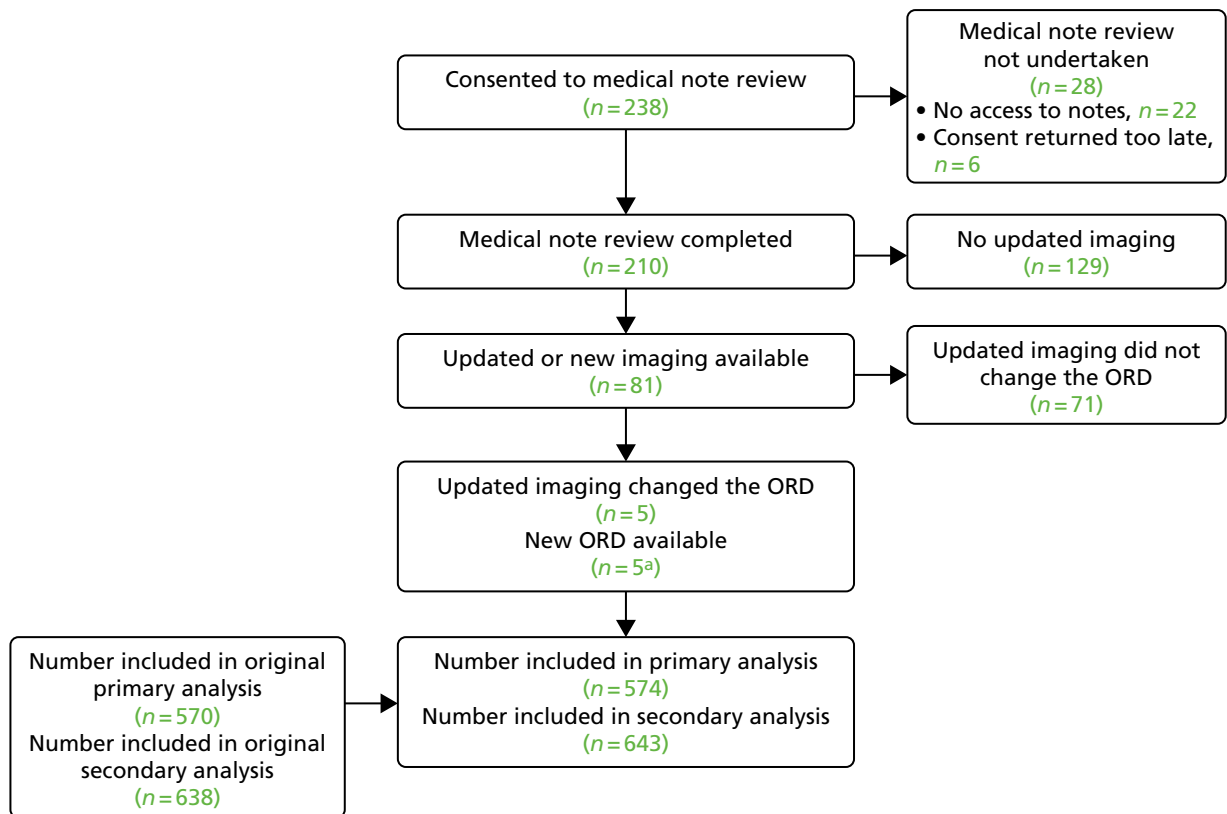


FIGURE 7 Participant flow in the project 1 component of the follow-on study.

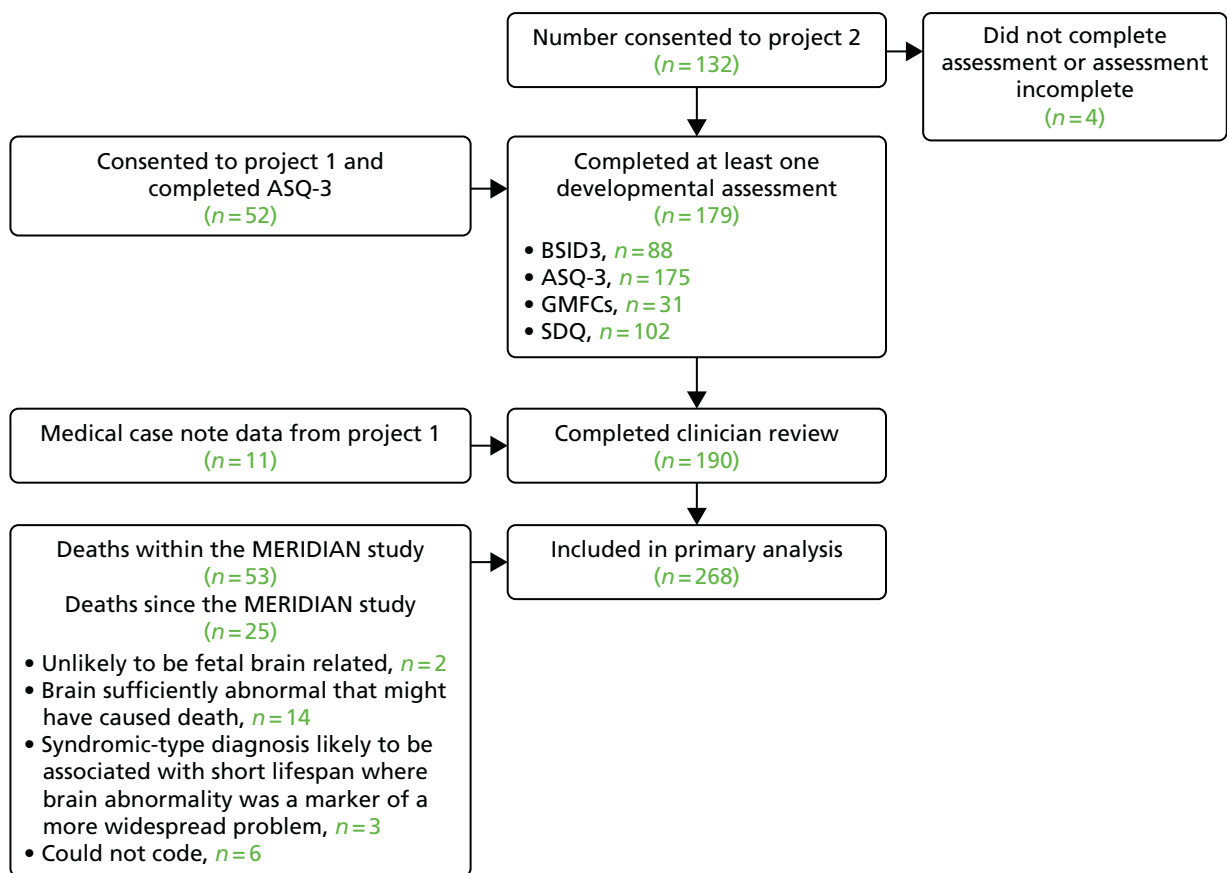


FIGURE 8 Participant flow in the project 2 component of the follow-on study.

Participant characteristics

Participants in projects 1 and 2 appear to be fairly representative of the overall MERIDIAN study cohort (Table 35). Compared with the original cohort, a lower proportion of participants with a poor prognosis on ultrasonography or iuMRI were recruited into project 1 or 2. A high proportion (55% on ultrasonography and 62% on iuMRI) of participants with a poor prognosis died within the original MERIDIAN study follow-up period (Figure 9). This can also be seen in Figure 10, demonstrating that a high proportion of non-surviving infants had a poor or intermediate prognosis. Additional tables and figures summarising the participants recruited in the study in comparison with the original cohort can be found in Appendix 11.

TABLE 35 Participant characteristics at entry to the MERIDIAN study

Characteristics	Overall (N = 829)	Project 1 and project 2 (N = 194)	Project 1 only (N = 45)	Not approached (N = 274)	Approached, not consented (N = 316)
iuMRI site, n (%)					
Sheffield	531 (64.0)	138 (71.0)	25 (56.0)	177 (65.0)	191 (60.0)
Birmingham	124 (15.0)	27 (14.0)	17 (38.0)	40 (15.0)	40 (13.0)
Newcastle	81 (10.0)	12 (6.2)	2 (4.4)	24 (8.8)	43 (14.0)
Leeds	57 (6.9)	5 (2.6)	1 (2.2)	21 (7.7)	30 (9.5)
Nottingham	16 (1.9)	2 (1.0)	0 (0.0)	5 (1.8)	9 (2.8)
Belfast	20 (2.4)	10 (5.2)	0 (0.0)	7 (2.6)	3 (0.9)
Maternal age (years) at approach, median (IQR)	29 (24–33)	3 (26–35)	32 (25–34)	29 (24–33)	28 (24–32)
Single/multiple pregnancy, n (%)					
Multiple	44 (5.3)	10 (5.2)	7 (16.0)	10 (3.7)	17 (5.4)
Single	784 (95.0)	184 (95.0)	38 (84.0)	263 (96.0)	299 (95.0)
Gestational age (weeks) at iuMRI, median (IQR)	23 (21–26)	23 (22–29)	22 (22–24)	22 (21–25)	23 (22–28.4)
Gestational age at iuMRI (weeks), n (%)					
≤ 24	531 (64.0)	116 (60.0)	34 (76.0)	195 (71.0)	186 (59.0)
> 24	298 (36.0)	78 (40.0)	11 (24.0)	79 (29.0)	130 (41.0)
Diagnosis on iuMRI, n (%)					
VM	405 (49.0)	106 (55.0)	30 (67.0)	89 (33.0)	180 (57.0)
CC	149 (18.0)	34 (18.0)	7 (16.0)	60 (22.0)	48 (15.0)
Posterior fossa	134 (16.0)	29 (15.0)	4 (8.9)	62 (23.0)	39 (12.0)
Other	138 (17.0)	25 (13.0)	4 (8.9)	60 (22.0)	49 (16.0)
Outcome of pregnancy, n (%)					
Death	53 (6.5)	0 (0.0)	0 (0.0)	53 (20.0)	0 (0.0)
Livebirth	642 (78.0)	194 (100.0)	45 (100.0)	87 (33.0)	316 (100.0)
Termination	124 (15.0)	0 (0.0)	0 (0.0)	124 (47.0)	0 (0.0)
The MERIDIAN study analysis population, n (%)					
ORD available	570 (69.0)	145 (75.0)	33 (73.0)	144 (53.0)	248 (79.0)
ORD unavailable	175 (21.0)	22 (11.0)	7 (16.0)	106 (39.0)	40 (13.0)
Excluded	84 (10.0)	27 (14.0)	5 (11.0)	24 (8.8)	28 (8.9)

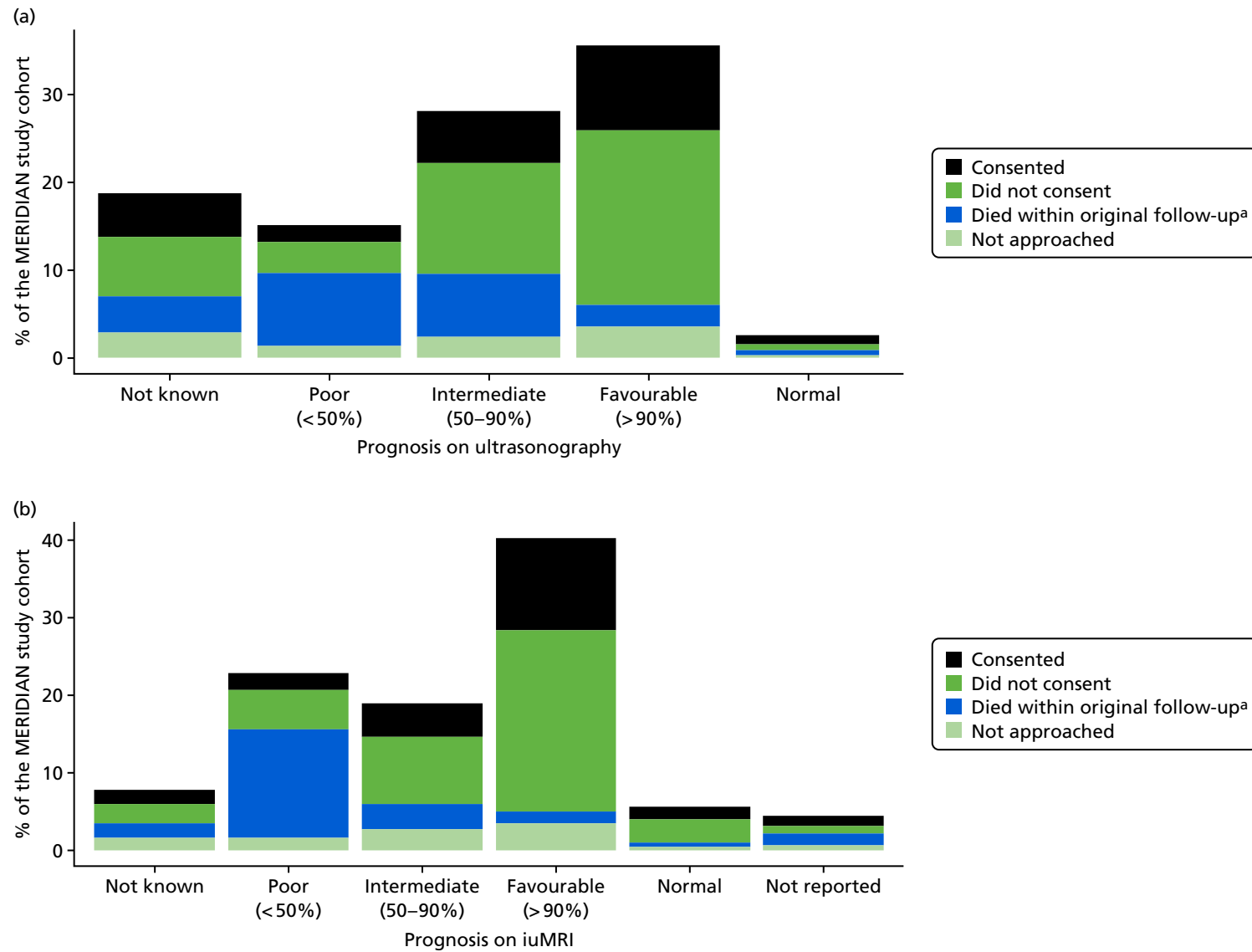


FIGURE 9 Bar chart showing the MERIDIAN study cohort ($n = 829$), by prognosis on (a) ultrasonography and (b) iuMRI for project 2. a, Death within original MERIDIAN include terminations, stillbirths, fetal deaths and deaths within 6 months of birth.

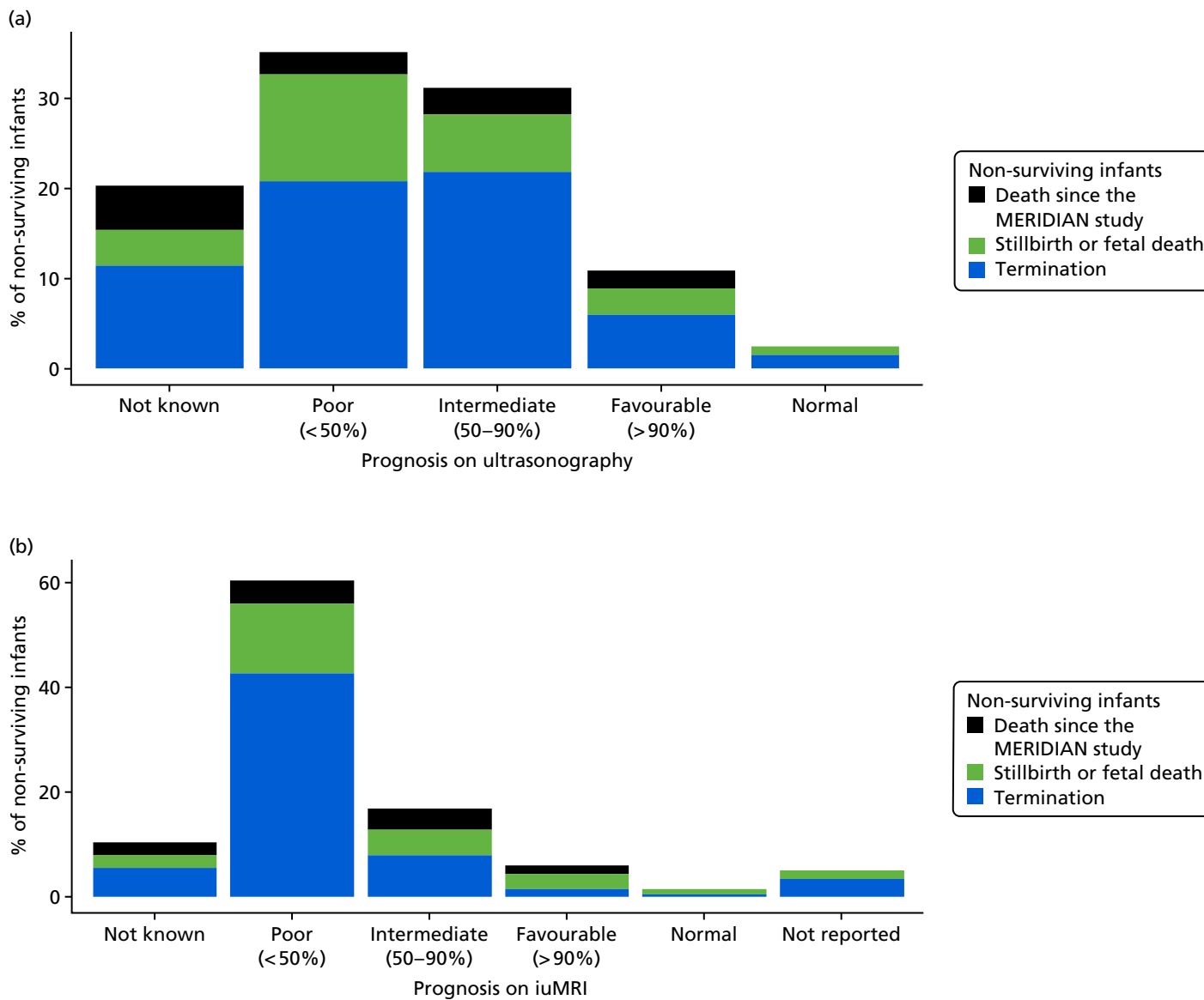


FIGURE 10 Bar chart showing the non-surviving infants in the MERIDIAN study cohort, by prognosis on (a) ultrasonography and (b) iuMRI ($n = 202$).

Project 1

Primary analysis project 1: diagnostic accuracy

Consent was obtained for 238 infants, of whom 81 had undergone further imaging since the original MERIDIAN study. As a result of this, 10 infants had an updated ORD. Five participants had updated imaging that did not show the same pathology as the previous ORD and the additional pathologies seen had not been acquired since birth. Of these five, the original ORD agreed with the ultrasonography in two cases and with the iuMRI in five cases; the revised outcome for all five participants was discrepant with both ultrasonography and iuMRI. Outcome data were obtained for a further five participants for whom there had been no previous ORD.

The primary analysis (comprising fetuses undergoing iuMRI within 14 days of ultrasonography) included four children with a revised ORD and four children who previously had no ORD. *Table 36* shows the updated diagnostic accuracy using updated imaging. The overall difference in diagnostic accuracy was 24.8% (95% CI 21.0% to 28.5%; $p < 0.0001$), which is very similar to the original difference of 24.9% and still statistically significant.

When including all participants with an ORD, regardless of the timing of the iuMRI (*Table 37*), there was still a significant difference of 23.3% (95% CI 19.8% to 26.9%; $p < 0.0001$).

Project 2

Developmental assessment scores and categories

Box plots showing the distribution of BSID3 scores appear to show very little difference between an ultrasonography prognosis and an iuMRI prognosis (*Figure 11*). Median BSID3 composite scores are fairly similar in those patients with a normal or favourable prognosis on ultrasonography and those with the same prognosis on iuMRI. There is a small difference between ultrasonography and iuMRI in median composite scores in those patients with a poor or an intermediate prognosis. iuMRI demonstrates slightly lower medians in each of the composite subscale scores. Lower scores indicate lower developmental ability. Box plots of ASQ-3 and SDQ can be found in *Appendix 12*.

TABLE 36 Diagnostic accuracy of ultrasonography and iuMRI by age of fetus in the 574 cases with ORD, using updated ORD ($n = 5$) and new ORD ($n = 4$) for participants who completed iuMRI within 2 weeks

Fetal age	Ultrasonography correct, n (%)	iuMRI correct, n (%)	Percentage difference (95% CI)	p -value ^a
18–23 weeks ($N = 372$)	257 (69.1)	339 (91.1)	22.0 (17.4 to 26.7)	< 0.0001
≥ 24 weeks ($N = 202$)	130 (64.4)	190 (94.1)	29.7 (22.9 to 36.5)	< 0.0001
Combined ($N = 574$)	387 (67.4)	529 (92.2)	24.8 (21.0 to 28.5)	< 0.0001

a McNemar's test between ultrasonography and iuMRI.

TABLE 37 Diagnostic accuracy of ultrasonography and iuMRI by age of fetus in 643 cases with ORDs, using updated and new ORDs. A total of 638 cases had previously had an ORD and a further five cases had a new ORD at 2–3 years

Fetal age	Ultrasonography correct, n (%)	iuMRI correct, n (%)	Percentage difference (95% CI)	p -value ^a
18–23 weeks ($N = 403$)	277 (68.7)	362 (89.8)	21.1 (16.7 to 25.5)	< 0.0001
≥ 24 weeks ($N = 240$)	160 (66.7)	225 (93.8)	27.1 (20.9 to 33.2)	< 0.0001
Combined ($N = 643$)	437 (68)	587 (91.3)	23.3 (19.8 to 26.9)	< 0.0001

a McNemar's test between ultrasonography and iuMRI.

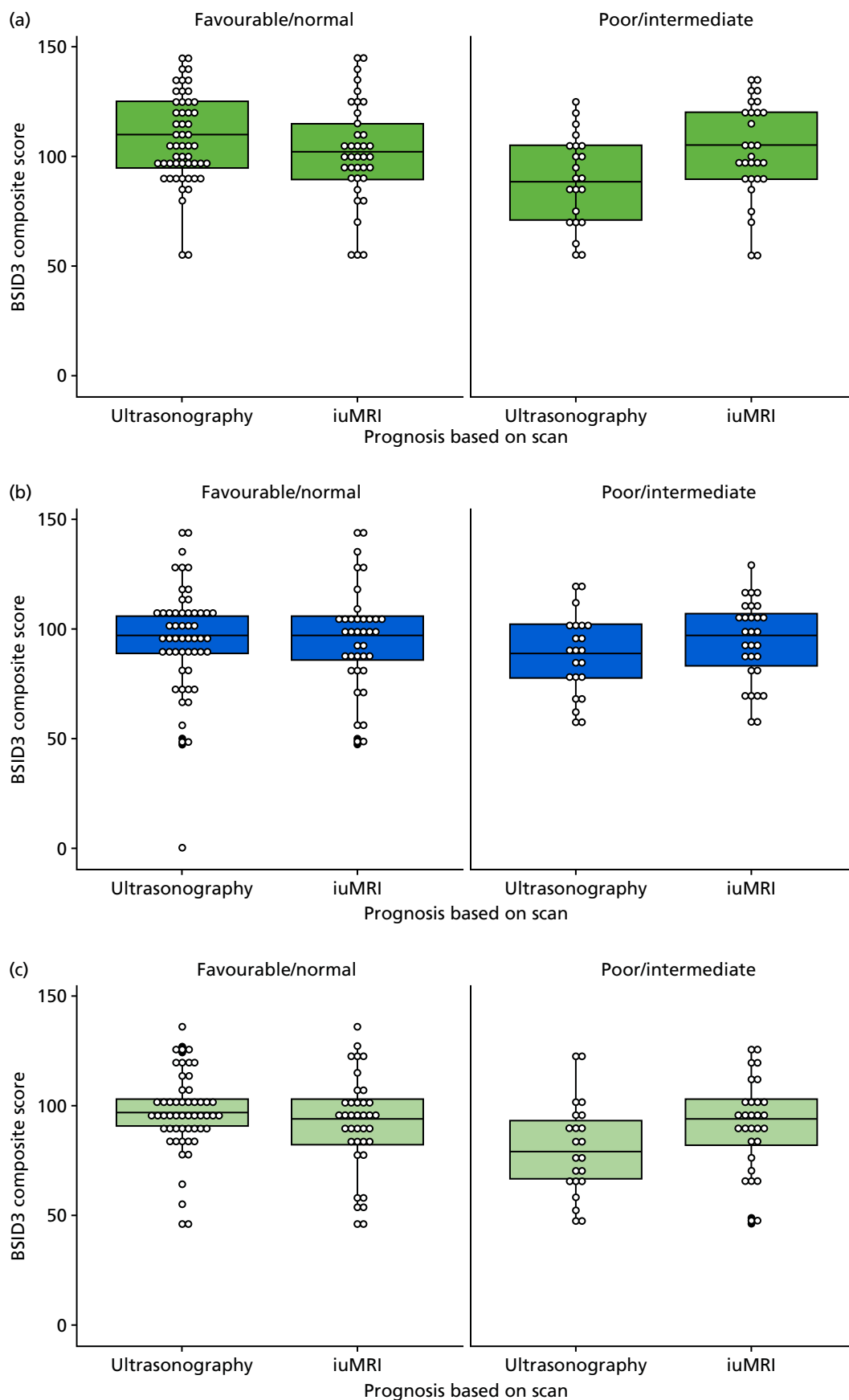


FIGURE 11 Box plots of BSID3 subscale composite scores, by ultrasound and iuMRI prognosis in the MERIDIAN study. (a) Cognitive index; (b) language index; and (c) motor index.

The majority of participants (106/190) who were followed up were assessed as normal (*Table 38*). A total of 58 out of 190 (31%) participants were assessed as abnormal. Descriptive statistics tables for the developmental scores are presented in *Appendix 13*.

Primary analysis: project 2

Our primary analysis of project 2 (*Table 39*) indicates that there is no difference between ultrasonography and iuMRI in prognosticating abnormal development in the surviving infants (difference of 5%, 95% CI -7% to 18%; $p = 0.505$). This observed result is likely to be influenced by the small number of infants from the original cohort with a poor prognosis being recruited into the study. A high proportion of these infants died.

In utero magnetic resonance imaging demonstrated a higher number of correct prognoses in surviving infants who were developmentally assessed as 'normal' or 'at risk' than did ultrasonography (ultrasonography, 51%; iuMRI, 70%). This difference was statistically significant (difference of 19%, 95% CI 10% to 30%; $p = 0.0004$).

TABLE 38 Number and percentage of participants in each development category overall and according to each individual assessment tool

Development tool ($N = 190$)	n (%)
Overall development category	
Abnormal	58 (31.0)
Borderline	26 (14.0)
Normal	106 (56.0)
BSID3 category	
Abnormal (< 70)	21 (24.0)
Borderline (70 to < 85)	12 (14.0)
Normal (≥ 85)	55 (63.0)
ASQ-3 category	
Abnormal	49 (29.0)
Borderline	24 (14.0)
Normal	98 (57.0)
SDQ category	
Abnormal (≥ 16)	17 (17.0)
Borderline (13–15)	5 (5.0)
Normal (0–12)	79 (78.0)
GMFCS level	
I	26 (84.0)
II	2 (6.5)
III	1 (3.2)
IV	2 (6.5)
V	0 (0.0)
Cerebral palsy	
Yes	4 (2.1)

TABLE 39 Primary analysis of project 2: prognosis on ultrasonography and iuMRI compared with the neurodevelopmental category at 2- to 3-year follow-up in surviving infants only

Prognosis	Abnormal development	Normal/at-risk development
Prognosis on ultrasonography, <i>n</i> (%)	<i>N</i> = 58	<i>N</i> = 132
Poor	11 (19.0)	5 (3.8)
Intermediate	14 (24.0)	35 (27)
Favourable	17 (29.0)	60 (46.0)
Normal	1 (1.7)	7 (5.3)
Not known	15 (26.0)	25 (19.0)
Ultrasonography correctly classified (%)	19	51
Prognosis on iuMRI, <i>n</i> (%)	<i>N</i> = 54	<i>N</i> = 125
Poor	13 (24.0)	5 (4.0)
Intermediate	15 (28.0)	21 (17.0)
Favourable	20 (37.0)	76 (61.0)
Normal	1 (1.9)	12 (9.6)
Not known	5 (9.3)	11 (8.8)
iuMRI correctly classified (%)	24	70
Both ultrasonography and iuMRI prognosis	<i>N</i> = 54	<i>N</i> = 125
Difference (95% CI)	5 (–7 to 18)	19 (10 to 30)
<i>p</i> -value ^a	0.505	0.0004

a McNemar's test between ultrasonography and iuMRI prognosis.

When including non-surviving infants (*Table 40*), iuMRI was slightly better at correctly classifying a prognosis (difference of 11%, 95% CI –3% to 21%). However, with the weighting applied to the non-surviving cases this difference decreased and was not statistically significant (difference of 8.37%, 95% CI –4% to 17%; $p = 0.2717$).

A further sensitivity analysis was conducted for which 'normal' and 'at-risk' cases were grouped together (*Table 41*). In this analysis, the difference in sensitivity was 6% (95% CI –1% to 21%; $p = 0.099$), which is greater than the difference in sensitivity shown in *Table 6*. The difference in specificity was 20% (95% CI 9% to 31%; $p = 0.0005$).

A breakdown of ultrasonography and iuMRI prognoses in abnormal, normal and borderline cases and a visual summary of *Table 39* can be seen in *Appendix 14*.

Ages and Stages Questionnaire, Third Edition, and Bayley Scales of Infant and Toddler Development, Third Edition

In total, ASQ-3 categorised 25 cases as abnormal, of which 17 cases were in agreement with BSID3, providing a PPV of 68% (*Table 42*). The ASQ-3 categorised 47 cases as normal and 12 cases as borderline. The BSID3 was in agreement with 57 out of 59 of these cases, providing a NPV of 97%.

The ASQ-3 correctly categorised 17 out of 19 abnormal cases (*Table 43*). This corresponds to a sensitivity of 90% (95% CI 67% to 99%). A total of 43 out of 52 normal cases were correctly categorised by the ASQ-3, providing a specificity of 83% (95% CI 70% to 92%).

Although the numbers are small, this provides some limited evidence that the ASQ-3 performs fairly well in categorising developmental impairment when compared with BSID3; however, the ASQ-3 had a relatively low PPV owing to eight false positives.

TABLE 40 Primary analysis of project 2: prognosis on ultrasonography and iuMRI compared with the neurodevelopmental category at 2- to 3-year follow-up in surviving and non-surviving infants

Prognosis	Fetal or infant death	Combined surviving and non-surviving infants (weighting of 0.306)
Prognosis on ultrasonography, <i>n</i> (%)	<i>N</i> = 78	<i>N</i> = 81.26
Poor	29 (37.0)	19.87 (25.0)
Intermediate	19 (24.0)	19.81 (24.0)
Favourable	10 (13.0)	20.06 (25.0)
Normal	2 (2.6)	1.61 (2.0)
Not known	18 (23.0)	20.51 (25.0)
Ultrasonography correctly classified (%)	38.00	30.63
Prognosis on iuMRI, <i>n</i> (%)	<i>N</i> = 73	<i>N</i> = 76.34
Poor	36 (48.0)	24 (31.0)
Intermediate	18 (24.0)	21 (28.0)
Favourable	9 (12.0)	23 (30.0)
Normal	2 (2.7)	2 (2.6)
Not known	10 (13.0)	8 (11.0)
iuMRI correctly classified (%)	49.00	38.99
Both ultrasonography and iuMRI, <i>n</i> (%)	<i>N</i> = 73	<i>N</i> = 76.34
Difference (95% CI)	11 (–3 to 21)	8.37 (–4 to 17)
<i>p</i> -value ^a	0.1456	0.2717

a McNemar's test between ultrasonography and iuMRI prognosis.

TABLE 41 Primary analysis of project 2 repeated combining abnormal and at-risk groups

Prognosis	Abnormal/at-risk development	Normal development
Prognosis on ultrasonography, <i>n</i> (%)	<i>N</i> = 84	<i>N</i> = 106
Poor	11 (13.0)	5 (4.7)
Intermediate	20 (24.0)	29 (27.0)
Favourable	28 (33.0)	49 (46.0)
Normal	2 (2.4)	6 (5.7)
Not known	23 (27.0)	17 (16.0)
Ultrasonography correctly classified (%)	13	52
Prognosis on iuMRI, <i>n</i> (%)	<i>N</i> = 80	<i>N</i> = 99
Poor	15 (19.0)	3 (3.0)
Intermediate	20 (25.0)	16 (16.0)
Favourable	34 (43.0)	62 (63.0)
Normal	4 (5.0)	9 (9.1)
Not known	7 (8.8)	9 (9.1)
iuMRI correctly classified (%)	19	72
Both ultrasonography and iuMRI prognosis	<i>N</i> = 80	<i>N</i> = 99
Difference (95% CI)	6 (–1 to 21)	20 (9 to 31)
<i>p</i> -value ^a	0.099	0.0005

a McNemar's test between ultrasonography and iuMRI prognosis.

TABLE 42 Comparison of development categories according to BSID3 and ASQ-3 in participants who completed both assessments ($n = 84$)

ASQ-3	BSID3			Total
	Abnormal	Borderline	Normal	
Abnormal	17	7	1	25
Borderline	2	2	8	12
Normal	0	3	44	47
Total	19	12	53	84

TABLE 43 The number of cases correctly diagnosed as abnormal and normal using ASQ-3 alone when compared with BSID3

Overall	Abnormal development cases correctly diagnosed		Normal development cases correctly diagnosed	
	Number of cases	Sensitivity (95% CI)	Number of cases	Specificity (95% CI)
	17/19	90 (67 to 99)	44/53	83 (70 to 92)

Project 3

Prevalence of impairment in participants with mild isolated ventriculomegaly

Of the 190 children included in project 2, 86 children had received a diagnosis of mild isolated VM on ultrasonography and 71 children received this diagnosis on iuMRI; 60 children had the same diagnosis on both modalities. In total, 19 (22%) participants diagnosed with isolated mild VM on ultrasonography had an abnormal neurodevelopment, with a further 12 (14%) participants considered 'at risk'. Among those patients with a diagnosis of isolated mild VM on iuMRI, the incidences were 18% on ultrasonography and 13% on iuMRI. None of the children had received a poor prognosis on either modality (*Table 44*).

TABLE 44 Developmental assessment in relation to the original prognosis

Prognosis	Included in follow-up, n (%)			Not included, n (%)	
	Abnormal ($N = 19$)	At risk ($N = 12$)	Normal ($N = 55$)	Did not consent to project 2 ($N = 173$)	Non-livebirth ($N = 38$)
Based on ultrasonography					
Poor (< 50%)	0 (0)	0 (0)	0 (0)	1 (1)	2 (5)
Intermediate (50–90%)	2 (11)	2 (17)	8 (15)	32 (18)	5 (13)
Favourable (> 90%)	12 (63)	8 (66)	40 (73)	130 (75)	28 (74)
Normal	1 (5)	0	2 (4)	2 (1)	1 (3)
Not known	4 (21)	2 (16)	5 (9)	8 (5)	2 (5)
Based on iuMRI					
Poor (< 50%)	0 (0)	0 (0)	0 (0)	0 (0)	3 (8)
Intermediate (50–90%)	5 (26)	0 (0)	7 (13)	21 (12)	5 (13)
Favourable (> 90%)	11 (58)	10 (83)	42 (76)	132 (76)	23 (61)
Normal	0 (0)	1 (8)	0 (0)	9 (5)	2 (5)
Not known	2 (11)	1 (8)	2 (4)	4 (2)	4 (11)
Not recorded	1 (5)	0 (0)	4 (7)	7 (4)	1 (3)

Isolated mild VM cases showed a higher incidence of abnormal neurodevelopment than the commonly quoted 5–10% incidence attributed to VM, irrespective of whether the diagnoses were based on ultrasonography or iuMRI. These differences may arise from features other than VM as described below.

Severity

The severity of VM was similar across the normal, at-risk and abnormal groups, with the largest trigone sizes being, on average, 11.0 mm, 11.0 mm and 11.1 mm, respectively, among the 86 cases imaged by ultrasonography. The largest trigone, as measured by iuMRI, showed more separation, with means of 11.3 mm, 11.5 mm and 12.1 mm, respectively, among the 71 cases, but with substantial overlap in size between the groups. A total of 62 out of the original 86 cases had ventricles within normal limits on ORD, with fewer instances of a resolved VM in the abnormal group (9/19 in the abnormal group vs. 2/11 in the at-risk group and 11/54 in the normal group).

Additional brain diagnoses

Ten (12%) children had additional brain diagnoses noted on ORD, although not all of these additional findings were considered to be of significance when assessing the overall diagnostic agreement. Eight of these diagnoses were also found on iuMRI. Proportionately more additional diagnoses were noted in the abnormal neurodevelopmental group, in which 4 out of 19 children had an additional diagnosis, compared with 2 out of 11 children in the at-risk group and 4 out of 54 children in the normal group. All 10 of the additional diagnoses were made postnatally and no new diagnoses were reported at the 2- to 3-year case review.

Other diagnoses and life events

A total of 15 children had either (non-brain) conditions or significant life events (e.g. cardiac complications, chromosomal abnormalities or prolonged hospitalisation) that are likely to have led to a delayed development. These were proportionately more common among children assessed as having abnormal neurodevelopment (7/19) than among children assessed as having at-risk (4/12) or normal development (4/54). Four of these children also had additional brain abnormalities (two in the abnormal group and two in the at-risk group).

Table 45 shows the prevalence of abnormal and at-risk neurodevelopment for children diagnosed with VM for the subset with no additional diagnoses and for the subset with no diagnoses or other complications.

TABLE 45 Prevalence of abnormal and at-risk neurodevelopment at 2- to 3-year follow-up among cases of isolated mild VM

VM diagnosis details	Abnormal		At risk	
	<i>n</i>	Prevalence, % (95% CI)	<i>n</i>	Prevalence, % (95% CI)
VM on ultrasonography				
All cases (<i>N</i> = 86)	19	22 (14 to 32)	12	14 (7 to 23)
Excluding additional brain diagnoses (<i>N</i> = 76)	15	20 (11 to 30)	10	13 (6 to 23)
Excluding all additional diagnoses or complications (<i>N</i> = 65)	10	15 (8 to 26)	8	12 (17 to 40)
VM on iuMRI				
All cases (<i>N</i> = 71)	13	18 (10 to 29)	9	13 (6 to 23)
Excluding additional brain diagnoses (<i>N</i> = 67)	12	18 (10 to 29)	9	13 (6 to 24)
Excluding all additional diagnoses or complications (<i>N</i> = 57)	7	12 (5 to 24)	8	14 (6 to 26)

Chapter 6 Add-on study

Introduction

Background

Recruitment into the MERIDIAN study was restricted to women with fetuses for which a brain abnormality was seen on ultrasonography. This allowed an assessment of the diagnostic accuracy of ultrasonography and iuMRI in terms of the false-positive rate of detecting a fetal brain abnormality; however, the false-negative rate cannot be estimated by the main MERIDIAN study. The add-on study addressed this by recruiting pregnant women carrying a fetus without any form of abnormality detected on detailed ultrasonography. This allowed us to describe the false-negative rate of ultrasonography.

Study aims and objectives

The aim of the add-on study was to answer the question 'What is the false-negative rate of antenatal ultrasound performed by fetal medicine specialists when trying to detect fetal brain abnormalities?'

Methods

Participants and eligibility criteria

Inclusion criteria

Participants were recruited to the add-on study based on the following criteria:

- Had an ongoing singleton or multifetal pregnancy of at least 18⁺⁰ weeks' gestation by ultrasonography dating.
- Was thought to be carrying a fetus with no abnormality whatsoever (i.e. no brain or somatic abnormality). This criterion was based on a normal mid-pregnancy abnormality scan in a fetal medicine centre or was confirmed as normal by a senior sonographer or specialist midwife.

Exclusion criteria

A participant was excluded from the study if any of the following criteria were met:

- had a history of a fetal brain abnormality in a previous pregnancy
- was unable to give informed consent
- had a cardiac pacemaker, an intraorbital metallic foreign body or a recent surgery with metallic sutures or implant
- had previously experienced or is likely to suffer severe anxiety or claustrophobia in relation to the iuMRI examination
- was unable or unwilling to travel to Sheffield for specialist iuMRI
- was unable to understand English (except where satisfactory translation services were available)
- was < 16 years old
- was unwilling for the GP to be informed about the study and given copies of scan reports.

Participating MERIDIAN study sites displayed posters and leaflets in public areas that included details of the add-on study. Potential participants could then ask staff about the study and receive the full participant information sheet (PIS). Fetal medicine specialists and research midwives also informed women about the study during their clinics and provided them with the PIS and leaflet. Clinic lists were also screened and women who had attended the clinic for a detailed ultrasonography abnormality scan within the last 7 days were sent a copy of the PIS by post. If the woman was interested in participating, the research midwife or

fetal medicine specialist gained permission to pass the woman's contact details on to the research team at the Academic Unit of Radiology, University of Sheffield; alternatively, the woman could make direct contact herself as details were provided on the study leaflet and PIS.

Local media, such as radio advertisements and newspapers, were also used to promote the study and included details of how to express an interest to the research team.

The research team at the Academic Unit of Radiology contacted the potential participants by telephone to complete a screening assessment and, if deemed eligible, arranged for them to attend the unit for the iuMRI study.

When the participant attended for the scan, further background information was collected and written informed consent was obtained.

Sample size

Starting from the assumption that no ultrasonography false negatives will be found, the study aimed to recruit 200 fetuses on the basis of the $3/n$ rule,⁸⁹ a large sample approximation of the upper 95% CI for very rare events. This allowed the NPV of ultrasonography to be estimated to an upper confidence limit of 1.5% in the absence of any abnormal scans and to within a standard error of $\leq 2\%$ for an incidence of $< 10\%$.

Statistical methods

In this add-on study, the diagnostic accuracy of ultrasonography was defined as the probability that a negative scan was correct; this corresponds to the NPV. For iuMRI, diagnostic accuracy was the percentage of cases correctly diagnosed, and the proportion of normal diagnoses corresponds to the NPV of iuMRI following ultrasonography. No attempt was made to estimate the sensitivity or specificity as the study was not a prospective sample of all pregnancies.

The fetal brain was assumed to be normal if the iuMRI study was normal, but in cases when an abnormality was suspected by iuMRI the pregnancy was followed up to ascertain whether ultrasonography, iuMRI or neither were correct.

Results

Recruitment and scanning took place between November 2013 and May 2017. A total of 225 pregnant women enquired about the study, of whom three did not meet the inclusion criteria because of pregnancy complications, 21 failed to attend or cancelled their appointment, two gave birth before their appointment and one iuMRI was abandoned owing to the participant feeling unwell. A total of 198 participants with 205 fetuses (14 twin pregnancies) were scanned successfully, of whom 26% were between 18 and 23 gestational weeks at the time of iuMRI.

The iuMRI studies were reported as normal for 203 cases but brain abnormalities were reported in two fetuses from separate pregnancies. In the first, a pathology consistent with a focal cortical dysplasia or cortical tuber was suggested on iuMRI and confirmed on postnatal MRI, although its exact nature remained uncertain. The second difference was a mild VM with the trigones of both lateral ventricles measuring between 10 mm and 11 mm; the fetus also had macrocephaly (biparietal diameter > 97 th centile and occipitofrontal on the 97th centile). The latter has been interpreted as a borderline and evolving diagnosis for which both ultrasonography and iuMRI reached appropriate diagnoses; however, the first is a missed negative for ultrasonography, resulting in a NPV of 99.5% (exact 95% CI 97.3% to 100.0%).

Chapter 7 Discussion

Main findings

We conducted a large, multicentre, observational, cohort study designed to assess the diagnostic accuracy, clinical impact and acceptability of iuMRI as an adjunct to detailed antenatal ultrasonography when diagnosing fetal brain abnormalities. The primary findings are summarised below, along with some unresolved questions.

Accuracy of diagnosis

Our main finding was that the overall diagnostic accuracies of ultrasonography and iuMRI were 67.9% and 92.8%, respectively, meaning that iuMRI showed an improvement in diagnostic accuracy of 24.9%. The diagnostic accuracy of ultrasonography was lower in fetuses with a gestational age of > 24 weeks (64.2%) than in those with a gestational age of 18–23 weeks (69.9%), whereas iuMRI remained similar at 93.5% for those > 24 weeks' gestational age and 92.4% for those with a gestational age of 18–23 weeks. This indicates an improvement in diagnostic accuracy of 22.5% in the 18–23 weeks group and of 29.3% in the > 24 weeks group. The reduced accuracy of ultrasonography observed in fetuses at > 24 weeks' gestation may be a result of several factors, including difficulties caused by the ongoing ossification of the fetal skull, the increased physical size of the woman and the descent of the fetal head into the maternal pelvis.

The accuracy of positive ultrasonography in diagnosing fetal brain abnormalities has been the subject of several previous studies.^{15–18} Our finding that ultrasonography was accurate in 69.9% of fetuses at < 24 weeks' gestation and in 64% of fetuses in later stages of gestation is consistent with those reports. This leads us to suggest that iuMRI significantly increases the accuracy of fetal brain diagnoses compared with ultrasonography alone in all fetuses at \geq 18 weeks' gestation.

The three commonest ultrasonography diagnoses were VM as the only intracranial abnormality (54%), an abnormality restricted to the contents of the posterior fossa (15%) and failed commissuration (14%). There were statistically significant improvements in diagnostic accuracy for iuMRI compared with ultrasonography in all three subgroups (VM, 8.5%; posterior fossa, 22%; failed commissuration, 60.7%), indicating that iuMRI has advantages in diagnosing a broad range of developmental abnormalities of the brain and should be included in the diagnostic pathway regardless of the suspected pathology type. Notably, additional brain abnormalities were present in 10.1% (31/306) of cases in which ultrasonography diagnosed isolated VM. iuMRI accurately recognised 27 of these additional brain abnormalities, resulting in a diagnostic accuracy of 98.4%, compared with ultrasonography alone at 89.9%.

Further to this, we were able to recruit women with low-risk pregnancies and normal fetuses on ultrasonography to undergo iuMRI of the fetal brain. To our knowledge, this is the first study of this population and our results confirm the ability of both ultrasonography and iuMRI to correctly confirm when the brain development of the fetus is normal. This highlights the validity of ultrasonography as the primary screening imaging method for pregnancy and further supports the need for additional iuMRI when abnormalities are detected.

Diagnostic confidence

The importance of diagnostic confidence as well as diagnostic accuracy when assessing an imaging technology is often overlooked and less well studied, even though Ng and Palmer⁵⁸ have explained its relevance. To the authors' knowledge, there are no published studies that report on the diagnostic confidence of iuMRI. Our study was designed to address this issue and our data show that iuMRI increased the proportion of high-confidence diagnoses by 13%, from 82% on ultrasonography alone to 95% after iuMRI. As highlighted by Ng and Palmer,⁵⁸ an incorrect diagnosis made with high confidence can result in

an inappropriate change in management (in this case an inappropriate TOP). We found that iuMRI gave fewer high-confidence but incorrect diagnoses than ultrasonography (6% vs. 22% of the total number of cases). Alternatively, TOP may not be offered if an imaging diagnosis is made with low confidence. This again can bring about medical errors if the diagnosis is found to be correct, as withholding an intervention may have detrimental effects. The MERIDIAN study data demonstrate that iuMRI resulted in fewer low-confidence diagnoses (5% iuMRI vs. 18% ultrasonography alone) and fewer incorrect low-confidence diagnoses (3% iuMRI vs. 8% ultrasonography alone).⁴⁴

Counselling and management

As far as we were aware, this is the first prospective analysis of the clinical impact of iuMRI in fetuses with brain abnormalities; in this paper we have shown that iuMRI brings about changes in counselling and management in a high percentage of cases.⁴⁴ Specifically, iuMRI provided additional diagnostic information in 49% of cases, caused a documented change in prognosis in 44% of cases and had major effects on counselling in 15% of cases. The contribution of iuMRI to overall clinical management was judged to be 'significant' in 26% of cases and had either a 'decisive' or 'major' influence in a further 9% of cases. This is a considerably larger effect than the 5% change in management anticipated at the design stage, an increase that can be attributed to the greater than predicted improvement in diagnostic accuracy.⁴⁴

Prognostic importance

The importance of the long-term functional significance of the suspected brain abnormality must not be overlooked and is of great importance in preparing parents for the potential outcomes. There are limited data available for many brain abnormalities (or combinations) of fetal abnormalities from which fetal medicine experts can draw to predict outcome⁶⁶ and to inform their counselling, although data on outcomes for certain abnormalities, such as isolated VM and ACC, are more plentiful. The 2- to 3-year follow-up study aimed to address this gap. Although recruitment figures were lower than anticipated, our findings indicate that, despite the increase in diagnostic accuracy and confidence provided by iuMRI, fetal medicine experts still find it difficult to prognosticate accurately, especially in terms of predicting abnormal development. There are a number of possible explanations for this uncertainty when providing a prognosis, including the experience of the fetal medicine expert in reading and interpreting the iuMRI report. It is possible that a neonatologist or paediatric neurologist, who are more familiar with these types of reports or conditions and may see affected children for long-term follow-up, may be more able to predict a prognosis with greater accuracy. The growing culture of litigation in health care may also be leading to concerns among fetal medicine experts regarding the counselling and management of a PND.⁹⁰ The impact of a brain abnormality, even if the abnormality is confirmed, can vary markedly between individual infants. The prognosis of an abnormality can include a great deal of uncertainty and the information and steps required to reach a decision can be complex.⁶⁴

Our key finding in relation to long-term prognostic importance is that iuMRI appears to be better at ruling out an abnormal developmental pattern than ultrasonography, most likely by excluding the diagnoses that were suspected on ultrasonography. iuMRI was associated with a statistically significant improvement of 19% in predicting a 'normal' and 'at-risk' developmental outcome than ultrasonography alone. By contrast, there was a smaller, non-statistically significant difference observed between iuMRI and ultrasonography in predicting an abnormal development. This finding may be a result of the small number of participants from the original cohort with a poor prognosis being included in the follow-up study.

Among the 54 children identified as having abnormal neurodevelopment at 2–3 years and for whom a prognosis was available for both ultrasonography and iuMRI, 11 (19%) and 13 (24%) had received a poor prognosis from ultrasonography and iuMRI, respectively (difference of 5%, 95% CI –7% to 19%; $p = 0.505$). A total of 78 children died, either in utero or prior to the 2- to 3-year follow-up, of whom 73 had received a prognosis on both modalities. Ultrasonography identified 29 (37%) children as having a poor prognosis compared with 36 (48%) children on iuMRI (difference of 11%, 95% CI –3% to 22%; $p = 0.146$).

Isolated mild VM is often associated with a favourable outcome;⁹¹ however, within our cohort 22% of cases of isolated mild VM diagnosed by ultrasonography had abnormal development at 2–3 years of age compared with 18% of cases on iuMRI. These findings are higher than the 11% identified in a review of 90 articles investigating the counselling and outcome when isolated mild VM is diagnosed on prenatal screening,⁹² yet it remains unclear whether or not this value is higher than in the general population. Our finding could be as a result of its association with other abnormalities.⁹³ The majority of children with isolated mild VM on either screening modality were recorded as having a resolved VM and no further diagnoses on ORD, but proportionately more unresolved VMs were found among those children with abnormal development at 2–3 years. Fifteen children had additional (non-brain-related) conditions or significant life events (e.g. cardiac complications, chromosomal abnormalities or prolonged hospitalisation) that are likely to have been associated with delayed development. These were proportionately more common among children assessed as having abnormal neurodevelopment (7/19) than among those assessed as having at-risk (4/12) or normal development (4/54).

Selection bias may play a role in this finding. Parents who have a particular concern about their child may be more likely to enrol in a study such as this, especially if they have already been discharged from local services or feel that their concerns are not listened to by health professionals.

Acceptability to patients

Considerable work has already been undertaken and published confirming the safety of iuMRI. The general risks of MRI apply and include scanning people with contraindications (e.g. pacemakers) and the risk of taking ferromagnetic objects into the scan room, which become a projectile threat. Rigorous screening procedures are conducted prior to the procedure to mitigate these risks. Screening was performed successfully in our study and led to a high completion rate of iuMRI (> 99%). Acceptability of the procedure was also high, with at least 95% of women reporting that they would have an iuMRI study again if a future pregnancy was complicated by a fetal brain abnormality.

Radiologists' experience

The radiologists had a wide range of iuMRI-reporting experience before the start of the study. The most experienced radiologist was from the host site (Sheffield), where approximately two-thirds of the cases in the study were performed. The overall iuMRI error rate was 7.0% and our analysis has shown that those with more experience were less likely to make errors. It must be noted that even the less experienced radiologists consistently had an error rate of < 10%. The analysis of errors has allowed us to investigate the factors that may influence the diagnostic accuracy of iuMRI, and this work has helped to identify practices that could be implemented into a national screening programme to reduce the risk of diagnostic errors.

Health economics

The health economic analysis estimated the incremental cost and benefits from iuMRI compared with ultrasonography alone. As the MERIDIAN study was not a controlled trial, the outcomes following ultrasonography alone had to be estimated. Three scenarios were considered to estimate the proportion of people who had a TOP after iuMRI or ultrasonography. The base case used the proportion of people who had a TOP after iuMRI for iuMRI, and the proportion of people who were offered TOP after ultrasonography (and before iuMRI) for ultrasonography. Scenario 1 used the proportion of people offered a TOP after iuMRI for iuMRI, and the proportion of people who were offered TOP after ultrasonography (and before iuMRI) for ultrasonography. Scenario 2 used the proportion of people who had a TOP after iuMRI for iuMRI, and adjusted the proportion of people offered a TOP after ultrasonography (and before iuMRI) by the proportion of those who were offered TOP following iuMRI and who had a TOP for ultrasonography. Therefore, the addition of iuMRI to the pathway changed the proportion of people who had a TOP instead of a delivery. In the base case, more people had a TOP after iuMRI than after ultrasonography, but in scenarios 1 and 2 more people had a TOP after ultrasonography than after iuMRI.

The addition of iuMRI to one ultrasonography scan increases costs per person by £264. The resource use for investigations, further imaging and further consultations differed in the MERIDIAN study for participants who had a TOP and for those who did not, with higher costs for those who had a TOP. However, a delivery is much more expensive than a TOP, meaning that the total cost for a participant who had a TOP is less than that for a participant who had a delivery. This means that if iuMRI led to sufficiently more participants having a TOP than a delivery, it would be cost-saving. In the base case, the proportion of participants having TOPs after iuMRI was much smaller than after ultrasonography, and so there was an incremental cost of £480 and a cost per management decision correctly revised of £2203. In scenarios 1 and 2, the proportions of participants having a TOP were much closer after iuMRI and ultrasonography, so the incremental costs were much lower.

In addition, the difference in costs and in the cost per management decision appropriately changed are sensitive to the assumptions around which costs should be included. Inclusion of participant-borne costs increases the incremental cost as participants have to spend additional time travelling to and attending iuMRI scans, but participant-borne costs are not routinely included in the evaluation of health technologies conducted by the National Institute for Health and Care Excellence in England. The exclusion of TOP and delivery costs reduces the uncertainty associated with the approach taken to estimate the number of TOPs with iuMRI and ultrasonography, as the major cost driver is the cost of the iuMRI itself.

In all scenarios, the incremental cost is consistently < £600 and the cost per management decision appropriately changed is always < £3000. In a scenario comparing iuMRI with repeat ultrasonography, iuMRI is cost-saving.

Whether or not iuMRI represents value for money depends on the willingness to pay per management decision appropriately changed. The probabilistic sensitivity analysis demonstrated that there is 100% probability that iuMRI is cost-effective if decision-makers are willing to pay \geq £4000 per management decision appropriately changed.

Although the parameters for costs and effectiveness are uncertain, the maximum EVPIs for the base case and scenarios 1 and 2 are very low, indicating that there is little value in conducting additional research to determine more precise estimates for the parameters.

Sociological substudy

Prenatal diagnosis technologies attract significant social science attention because they constitute a case study of intrinsic interest for those exploring ethical, legal and social dilemmas, an opportunity to interrogate how social contexts shape understandings of health and medical knowledge. Our findings provide insights from both the service users' and professionals' perspectives on how these issues affect a PND of brain abnormality in the context of iuMRI. Service users provide insights into lay perspectives on acceptability of and barriers to care, and professionals provide insights into the possible implementation of NHS health-care policies. We consider both perspectives when exploring the social issues that, as Williams³¹ notes, are often neglected when rolling out new technologies in fetal medicine. In both service users' and health professionals' accounts, there was much similarity in how the understandings of acceptability of care were conceptualised in complex ways, acknowledging trade-offs and weighing up options.

Our findings suggest that iuMRI in a PND for brain abnormality is broadly acceptable to both service users and health professionals, but it is not considered a self-evident good. The acceptability of iuMRI was conditional, with service users' and professionals' accounts demonstrating careful consideration of the costs and benefits of iuMRI, and providing suggestions for how service user and professional acceptability could be assured and enhanced in future implementation. Nevertheless, professionals found it easy to recruit women to the trial, and the quantitative data from the patient survey demonstrate the high level of women's baseline acceptance of iuMRI as part of their care. Interviews with women showed that they were usually 'information hungry' and prepared to tolerate significant discomfort to maximise information prior to decision-making about their pregnancy. Apart from specific issues, such as claustrophobia,

the main barriers to patient acceptability centred on having to travel (linked to, but not limited to, socioeconomic costs). Professionals also conceptualised iuMRI as useful in the PND of brain abnormality, but with diverse views on the types of brain abnormality for which iuMRI was considered to 'add value' to the PND process.

The balancing of the advantages and disadvantages of iuMRI in the PND for brain abnormality reveals three main issues: accessibility, triangulation and quality assurance.

Accessibility

Accessibility was conceptualised in both geographic and temporal terms. Geographically, iuMRI provision has emerged in areas with facilitating factors (e.g. interested clinicians and sufficient resources). Overall case numbers do not warrant specialist centres in every region, so it is inevitable that although some patients will live near a centre with expert iuMRI provision, many will have to travel much further to access it. This is not unique to PND in the context of the need for cost-efficient use of NHS resources. However, PND has unique characteristics that make geographical barriers more problematic. PND is highly sensitive to temporal delay in accessing services because decisions about TOPs for fetal abnormality are time sensitive. Both clinicians and service users rely on diagnostic and prognostic information to inform decisions about whether or not ending a pregnancy is, first, legal to offer, and, second, the option that the parents select. Delays in accessing information from iuMRI have serious consequences for those making these sensitive, high-stakes decisions. The timely access to and reporting of iuMRI were seen as critical to service user and professional experiences. Adequate resourcing was needed to ensure timely access to iuMRI.

Triangulation

Triangulation encompasses many of the perceived advantages of iuMRI in the PND for brain abnormality. First, iuMRI was seen as complementary to ultrasonography, not as a replacement. iuMRI represented a comparative view of the fetal brain, offering reassurance in the face of uncertainty after ultrasonography. This triangulation effect had particular power in managing the inherent uncertainties associated with PND, although definitive conclusions often remained elusive.⁷³ Participants found iuMRI especially helpful in cases that were closer to the borderline of normality (e.g. VM). The triangulation effect was evident in clinicians' views that MDT discussion was valuable. However, although visual representations were often seen as a useful way of generating shared understandings, the magnetic resonance images were by no means the only or the dominant form of communication supporting shared meaning-making. Many fetal medicine specialists appreciated iuMRI as it provided access to a second professional opinion or an opinion from a different imaging modality. Those in larger units (perhaps with more opportunities for informal second opinions) tended to see the value as limited to specific brain anomalies.

Quality assurance

Quality assurance of iuMRI image interpretation was seen as crucial to its successful implementation, with discussion focusing on transparency and reputation. For service users, transparency in the PND process could be enhanced by using iuMRI images; those with the opportunity to view iuMRI images were more likely to value iuMRI information. The interview findings suggested that it was the in-depth explanations that were used as narration for viewing the iuMRI images that were of particular value to women and their partners, which may mean that the valued transparency is an effect of having iuMRI images to discuss, rather than inherent in the images themselves. Some health professionals also identified the visual element of iuMRI as useful (e.g. seeing the brain as symmetrical). More commonly, clinicians identified the transparency in reporting images (of both normality and abnormality) as their priority. In fetal medicine, reporting of ultrasonography images is usually assured by using a standardised list and repeat scanning; whereas in radiology, iuMRI is usually assured by the use of prescribed imaging protocols and the double reporting of images. Clinicians were more confident of the value of the quality assurance mechanisms associated with their home specialty, but all saw that transparently assuring standards was critical, given the importance of the information in terms of decision-making. In the current system, when working across specialty boundaries, much depends on respecting the reputation of the person interpreting the image, in line with Reed *et al.*'s⁷⁶ concept of the 'bridging' effect between professional groups.

Concepts of acceptability and satisfaction with care are context specific. Parents and health professionals are working in challenging circumstances to produce the 'least worst' outcome rather than a truly positive experience. The patient survey data demonstrate broad acceptability of iuMRI in the PND of brain anomalies, with at least 95% of respondents indicating that they would choose iuMRI again if in similar circumstances. But acceptability is strongly influenced by perceived utility in a difficult situation, with at least 80% of respondents feeling that the information from iuMRI was useful to them. Interviews suggested that parents appreciated that iuMRI offered them details that could help inform judgements about the severity of the fetal brain abnormality. Similarly, $\approx 90\%$ of patients felt that their care was either excellent or very good. However, the broad acceptability of iuMRI needs to be balanced with the inconveniences that patients and professionals will tolerate in the PND pathway. Such inconveniences are significant but tolerated because of the lack of more palatable options.

Strengths and weaknesses

Strengths and weaknesses of the study

The MERIDIAN study was designed to address the weaknesses in the current literature by conducting an appropriately powered, pragmatic, prospective, multicentre study. This allowed us to complete a robust evaluation of the extent to which iuMRI improves the diagnostic accuracy of ultrasonography alone and to assess the clinical impact on the counselling and management of the pregnancy. The pragmatic nature of the study allowed us to assess and demonstrate how iuMRI could be incorporated in the diagnostic pathway in the UK for diagnosing fetal brain abnormalities.

Our primary analysis included only cases for which the iuMRI was performed within 2 weeks of the ultrasonography and an ORD was available, which addressed a weakness of many prior studies that often did not rely on an ORD. To our knowledge, this is the first study to present an analysis of diagnostic confidence and assess its impact on appropriate management decisions.

Within this study, we have not only been able to confirm the improved diagnostic performance of iuMRI when diagnosing fetal brain abnormalities but we have also shown that ultrasonography and iuMRI are both capable of confirming when the development of a fetal brain is normal. We believe that this is the first study within this population to confirm this and will have implications for designing a national antenatal screening programme.

The limitations of the study include the potential for reporting bias from the referring clinicians on the diagnostic and prognostic outcomes, as well as the information about their confidence in the findings. However, changes in the prognosis were under-reported by clinicians, suggesting that they were not systematically going back to review their initial decisions at the time of reviewing the iuMRI results.

Two further limitations were the high proportion of participants without an ORD (24%) and the high proportion of iuMRI scans undertaken at the host institution (67%). The first limitation arose because we were unable to undertake diagnostic investigations solely for research purposes, meaning that our ORD relied on collecting data from investigations completed for clinical purposes, which was not always done. This was particularly an issue following TOP or fetal demise, which accounted for over half of the cases of missing ORDs. The second issue is an important limitation because two-thirds of the scans were undertaken by two radiologists who were experienced in undertaking MRI specifically in fetal abnormalities; other centres will, at present, not have this level of expertise. To address this, we undertook two sensitivity analyses. The first found that the improvement in diagnostic accuracy following iuMRI could not be attributed to missing data, with even extreme imputation of missing data showing iuMRI following ultrasonography to be superior to ultrasonography alone. The second showed only a marginally higher accuracy at the host institution compared with other centres. Both of these analyses appear to confirm the robustness of the difference in favour of iuMRI.

Strengths and weaknesses of the 2- to 3-year follow-up study

The main limitation of the follow-up study was the small number of children included and for whom developmental data could be collected. Owing to delays in study initiation, many of the children were out of the age range to be offered the BSID3 assessment, owing to the ceiling effect. Therefore, we relied on ASQ-3 data for many children. The ASQ-3 is a screening tool and its sensitivity, specificity, PPV and NPV are not perfect. It is designed not to miss true positives at the risk of erroneously diagnosing some children as abnormal who would then receive a correct diagnosis on further assessment by a clinician. Our analysis of ASQ-3 classification compared with the BSID3 assessment confirmed that the PPV was relatively low. Although this affects the generalisability of the data and limits the conclusions that can be drawn, it does highlight the importance of further research into the uncertainty of the prognostic importance of brain abnormalities diagnosed antenatally. There are a number of factors that influenced our recruitment rate; these included a delayed start to the study (including delay with funding approval and issues with site set-up). These delays meant that a number of the children were out of the age range for the BSID3 assessment and the ASQ-3 and, therefore, only a medical case note review could be completed. There was also a high proportion of families who could not be contacted. The approach protocol was updated during the study to increase the number and type of contact options available to the research team (as detailed in *Chapter 5, Protocol amendments after study initiation*); however, many families had moved and there was no follow-up address available or contact telephone numbers were out of use.

The BSID3 is a well-validated tool for assessing development in early infancy; it is widely used and generates standardised scores that allow corrections for differences in age at measurement. We wish to highlight that despite the widespread use of BSID3 in follow-up studies, and the lack of better alternatives, it remains a tool to assess developmental stage and not a tool to accurately predict the final developmental or cognitive outcome later in life; however, it is currently widely considered to be the reference standard. It is entirely possible for a child who performs well on the BSID3 between 2 and 3 years to manifest developmental or learning difficulties later in life and vice versa. Owing to the complexity of early childhood development, the BSID3 may also fail to detect or predict the presence of important behavioural or psychological problems in later life. It is also important to note that the cognitive measure is affected by the child's motor ability and, therefore, a lower score may be obtained that may not accurately reflect the child's ultimate level of cognitive function. The BSID3 is also time-consuming and more than half of the children did not undergo the full BSID3 assessment, with the development instead ascertained by a brief questionnaire-based assessment. Although unavoidable, this is likely to have contributed a degree of measurement error to the final classifications.

Strengths and weaknesses of the economic analysis

The key limitations of this analysis lie in the form that the analysis takes and the outcomes that are used. The outcome of a management decision appropriately revised combines both cases in which a decision to continue the pregnancy was changed to TOP and cases in which a decision to terminate was changed to a continuation of pregnancy. In reality, the willingness to pay per appropriately revised decision may differ between those situations in which the intention to terminate a pregnancy was revised to a continuation of the pregnancy and those situations in which the continuation of the pregnancy was revised to a TOP. The analysis does not consider the implications of inappropriately changing a management decision (e.g. when iuMRI was incorrect and changed the prognosis in 1.81% of cases). The impacts on decision-makers, parents and children are clearly profound, both when a decision is taken to terminate a healthy fetus or when the decision is to continue the pregnancy when a fetal abnormality exists. These limitations should not be overlooked when considering the cost implications. Findings relating to these aspects of the study are reported in other sections of this report and should be considered in conjunction with evidence of cost-effectiveness.

Strengths and weaknesses of the sociological substudy

There are limitations to this study. The use of purposive sampling for maximum variation in the substudy samples allowed the small samples to encompass broadly representative ranges of views. The findings from the analysis of the data are not intended to be generalised in a simplistic way to the relevant population, but they do provide a robust account of the likely range of views within that population.⁷⁸ Although all eligible participants were included in the survey, not all participants responded to both surveys. Therefore, the results of the surveys do not include information on non-respondents and we do not know whether or not that subgroup had different views from those expressed by participants. In addition, all of our participants were recruited from a subgroup that actively chose to get involved in the MERIDIAN study (whether as a health professional or as a patient), and so the findings do not encompass the views of those who did not take part in the larger clinical study.

Strengths and weaknesses of the add-on study

The add-on study was a prospective study of fetuses with no brain abnormalities on ultrasonography. Although the sample size of 205 participants is notable, it was small in terms of its ability to estimate the NPV of ultrasonography. The fact that one abnormality was found (a NPV of 99.5%, 95% CI 97.3% to 100.0%) is reassuring in terms of the low error rate but may raise concern in mothers who may not grasp the possibility that ultrasonography is potentially an imperfect rule-out in 2% of pregnancies. Although the false-negative rate is likely to be lower than this, our study was not of a sufficient size to estimate the error rate more precisely.

Although the main focus of this study was on the use of iuMRI in fetuses with a suspected brain abnormality, it is likely that iuMRI will be considered for fetuses that are considered 'at risk' but whose brain appears normal on ultrasonography. The incident of a fetus with an abnormality not being detected by ultrasonography (and indeed the second case with mild VM) is a timely reminder that ultrasonography is imperfect. For the foreseeable future, it is not viable for iuMRI to be used routinely in the manner in which ultrasonography currently is, but future research may be useful in assessing in which circumstances the addition of iuMRI to the diagnostic pathway would be of benefit.

Chapter 8 Conclusions

Our results indicate an overall absolute improvement in diagnostic accuracy of 24.9% when iuMRI is used as an adjunct to ultrasonography to diagnose fetal brain abnormalities. Statistically significant differences were found in diagnostic accuracy in both gestational age groups studied, specifically with an increase of 22.5% in the 18–23 weeks group and an increase of 29.4% in the ≥ 24 weeks group.

The use of iuMRI in the diagnostic pathway also improved diagnostic confidence when compared with ultrasonography alone. The combination of improved diagnostic accuracy and confidence leads to changes in counselling and management in a high proportion of cases. Our subgroup analysis demonstrates the value of iuMRI across the three anatomical subgroups studied. iuMRI for the diagnosis of fetal brain abnormalities has also been shown to be acceptable for both service users and health professionals. However, our results suggest that despite the increase in diagnostic accuracy and confidence provided by iuMRI, fetal medicine experts still find it difficult to prognosticate accurately, especially in terms of predicting abnormal development.

The inclusion of iuMRI in the diagnostic pathway resulted in additional costs compared with ultrasonography alone, owing to the cost of iuMRI itself and the costs from resulting management decisions. Across a range of scenarios, the incremental cost was consistently $< \pounds 600$ and the cost per management decision appropriately changed was always $< \pounds 3000$. Whether or not iuMRI represents value for money depends on willingness to pay per management decision appropriately changed; if decision-makers are willing to pay $\geq \pounds 4000$ per management decision correctly revised, then there is a 100% probability that iuMRI is a cost-effective addition to the diagnostic pathway.

These results, along with the high completion rates for the procedure in the MERIDIAN study and the favourable safety profile of the technique, indicate that iuMRI has a major diagnostic and clinical impact on the PND of fetal brain abnormalities.

Implications for practice and policy

This study provides robust data indicating that there would be significant benefits from the routine use of iuMRI as an adjunct to ultrasonography in the diagnostic pathway for identifying fetal brain abnormalities.

The analysis of the types of diagnostic errors made by iuMRI has allowed us to identify and address several clinical considerations of how a national programme with the best diagnostic accuracy achievable can be provided. These include providing an iuMRI service at a small number of supraregional centres and/or promoting collaboration between radiologists to form expert panels for cases, formal training to improve radiologists' knowledge base, supervision of the iuMRI study by the radiologist and the double reporting of images. A number of these proposals have been supported by the views of patients and health professionals, as detailed below.

The sociological substudy included not only the patients' views of the iuMRI but also those of the health professionals. Insights from both perspectives are important in informing the future development of service provision. Although decisions on funding and policy may be made for a particular service, for them to be effectively implemented they must be acceptable to patients, and also have the formal and informal support of clinicians who deliver care, because professionals retain a relatively high level of professional

discretion in how they carry out their work.^{94,95} The following discussion points represent the most important implications for policy development in this area:

- Small changes to the organisation of iuMRI facilities have the potential to have a positive impact on patient experience (e.g. allowing a supportive person to sit with the woman in the iuMRI room during the scan and providing good information for parents). Most patients involved in the PND pathway are anxious but information hungry. Providing more information for women, including reassurances about the impact of iuMRI on themselves and on the fetus, and providing support while the women await results, as well as a written summary of the iuMRI report, may help to reduce negative impacts on women during an already anxiety-provoking process.
- Clear and transparent reporting of iuMRI images would have a significant benefit for both patients and professionals. Women value clear explanations using a wide range of resources, including visual representations (not necessarily iuMRI images). Professionals value robust quality assurance mechanisms in image reporting, facilitating working across the geographical distances that cost-efficient use of resources for specialist services requires. Consideration could be given to incorporating into the quality assurance process the elements that are familiar to the professional groups involved [e.g. standardised reporting of features (from fetal medicine) and double reporting of images (from radiology)]. A modest investment in visibly robust quality assurance mechanisms could enhance professional engagement with whatever cost-effective national provision framework is implemented.
- The timing of care provision is compressed in PND and, therefore, providing appropriately timed access to iuMRI could help encourage professional and lay engagement with iuMRI, whereas lengthy delays in accessing iuMRI may result in the proliferation of local arrangements, reducing opportunities for cost-effectiveness and national oversight of care.
- For clinicians working across specialty boundaries, collaboration is considered to be crucial. Many participants saw the value of more collaborative working across fetal medicine and radiology on the issue of fetal imaging and in relation to specialty training, but also with regard to collaborative working in MDTs.
- All of the issues listed above link back, in some way, to the issue of adequate resourcing. Demand for MRI provision appears to be a limiting factor on both user experience and service development, but increasing capacity in trained personnel and in terms of the number of MRI scanners is a lengthy and expensive investment. It may be helpful to consider some shorter-term and longer-term plans for dealing with the resource issues associated with implementing policies that increase the demand for MRI provision.

Recommendations for future research

The MERIDIAN study recruited women carrying fetuses with a suspected brain abnormality shown on ultrasonography, and the add-on study (see *Chapter 6*) included fetuses thought to have no form of abnormality whatsoever. Therefore, cases of isolated microcephaly were not studied in the MERIDIAN study. Microcephaly is often associated with other brain abnormalities and the severity correlates with the risk of a poor outcome,^{96,97} yet there is a paucity of literature with regard to the use of iuMRI in fetuses with microcephaly.

The use of iuMRI in the diagnosis of fetal spine abnormalities has not been well studied; however, data from a small study of 33 participants⁹⁸ suggest that iuMRI may have a role in the diagnosis of these abnormalities. Further research is required to evaluate the use of iuMRI in the diagnosis of spine abnormalities.

As identified in the 2- to 3-year follow-up study (see *Chapter 5*), the functional significance of the brain abnormality is more important to parents than the specific structural or anatomical diagnosis. We have presented data from children aged 2–3 years and also discussed the limitations of the tools available to assess development at this age. Therefore, we propose that longer-term follow-up studies of children diagnosed with fetal brain abnormalities are required.

Acknowledgements

We gratefully acknowledge the hard work, support and advice from all of the MERIDIAN research midwives and research nurses who contributed to the screening, recruitment and outcome data collection. We thank Daniel Beever (University of Sheffield CTRU) for study monitoring and site close-down activity. We thank Amanda Loban, Kirsty Pemberton and Emily Turton (University of Sheffield CTRU) for their support on data-management concerns. We thank Leanne Armstrong (University of Sheffield, Academic Unit of Radiology) and Rita Lynch (University of Sheffield CTRU) for their administrative and clerical support. We thank Aimee Card (Research Co-ordinator, Research and Development Department, Sheffield Teaching Hospitals NHS Foundation Trust) for her support on ethics and governance issues.

We offer special thanks to the members of our two oversight committees, the TSC and the DMEC.

Trial Steering Committee

Zarko Alfirovic (Liverpool Women's Hospital), Harriet Joy (University Hospital Southampton NHS Foundation Trust), Jane Fisher (ARC), Gill Yaz (SHINE), Samantha Johnson (University of Leicester) and Julius Sim (Keele University).

Data Monitoring and Ethics Committee

Julius Sim (Keele University), David Howe (University Hospital Southampton NHS Foundation Trust) and Alan Montgomery (University of Nottingham).

The MERIDIAN study research team

Writing group

Paul D Griffiths (Academic Unit of Radiology, University of Sheffield), Michael Bradburn (CTRU, School of Health and Related Research, University of Sheffield), Michael J Campbell (School of Health and Related Research, University of Sheffield), Cindy L Cooper (CTRU, School of Health and Related Research, University of Sheffield), Nicholas Embleton (Newcastle Neonatal Service, Royal Victoria Infirmary), Ruth Graham (School of Geography, Politics and Sociology, Newcastle University), Anthony R Hart (Department of Perinatal and Paediatric Neurology, Sheffield Children's Hospital NHS Foundation Trust), Deborah Jarvis (Academic Unit of Radiology, University of Sheffield), Mark D Kilby [(1) Centre for Women's & Newborn Health, Institute of Metabolism & Systems Research, University of Birmingham; (2) Fetal Medicine Centre, Birmingham Women's Foundation Trust (Birmingham Health Partners)], Mabel Lie (Institute of Cellular Medicine, Newcastle University), Gerald Mason (Leeds Teaching Hospitals NHS Trust), Laura Mandefield (CTRU, School of Health and Related Research, University of Sheffield), Cara Mooney (CTRU, School of Health and Related Research, University of Sheffield), Rebekah Pennington (Health Economics and Decision Science, School of Health and Related Research, University of Sheffield), Stephen C Robson (Institute of Cellular Medicine, Newcastle University) and Allan Wailoo (Health Economics and Decision Science, School of Health and Related Research, University of Sheffield).

Trial Management Group

Paul D Griffiths (Academic Unit of Radiology, University of Sheffield), Michael Bradburn (CTRU, School of Health and Related Research, University of Sheffield), Michael J Campbell (School of Health and Related Research, University of Sheffield), Cindy L Cooper (CTRU, School of Health and Related Research, University of Sheffield), Nicholas Embleton (Newcastle Neonatal Service, Royal Victoria Infirmary), Ruth Graham (Newcastle University), Anthony R Hart (Department of Perinatal and Paediatric Neurology, Sheffield Children's Hospital

NHS Foundation Trust), Gerald Mason (Leeds Teaching Hospitals NHS Trust), Stephen C Robson (Newcastle University) and Allan Wailoo (Health Economics and Decision Science, School of Health and Related Research, University of Sheffield).

Co-applicants

Paul D Griffiths (Academic Unit of Radiology, University of Sheffield), Michael Bradburn (CTRU, School of Health and Related Research, University of Sheffield), Michael J Campbell (School of Health and Related Research, University of Sheffield), Cindy L Cooper (CTRU, School of Health and Related Research, University of Sheffield), Nicholas Embleton (Newcastle Neonatal Service, Royal Victoria Infirmary), Ruth Graham (Newcastle University), Gerald Mason (Leeds Teaching Hospitals NHS Trust), Stephen C Robson (Newcastle University) and Allan Wailoo (Health Economics and Decision Science, School of Health and Related Research, University of Sheffield).

Local investigators

Mark D Kilby (Birmingham Women's Hospital NHS Foundation Trust), Janet Wright (Bradford Teaching Hospitals NHS Foundation Trust), Christoph Lees and Jeremy Brockelsby (Cambridge University Hospitals NHS Foundation Trust), Anne-Marie Coady (Hull and East Yorkshire Hospitals NHS Trust), Christoph Lees (Imperial College Healthcare NHS Trust), Graham Tydeman (NHS Fife), Richard Smith (Norfolk & Norwich University NHS Foundation Trust), Pam Loughna and George Bugg (Nottingham University Hospitals NHS Trust), Gerald Mason and Kelly Cohen (Leeds Teaching Hospitals NHS Trust), Samina Dornan (Royal Maternity Hospital Belfast), Stephen C Robson (Royal Victoria Infirmary), Dilly Anumba (Sheffield Teaching Hospitals NHS Foundation Trust), Edward Johnstone (St Mary's Hospital, Central Manchester University Hospitals NHS Foundation Trust), Ian Scudamore and Hatem Mousa (University Hospitals of Leicester NHS Trust), Audrey Long and Geraldine Masson (University of North Midlands NHS Trust).

Local investigators for the 2- to 3-year follow-up study

Manobi Borooah (Birmingham Women's Hospital NHS Foundation Trust), Roopali Soni and Sunita Seal (Bradford Teaching Hospitals NHS Foundation Trust), Kathryn Beardsall (Cambridge University Hospitals NHS Foundation Trust), Hassan Gaili (Hull and East Yorkshire Hospitals NHS Trust), Nigel Basheer (Imperial College Healthcare NHS Trust), Sean Ainsworth (NHS Fife), Priya Muthukumar (Norfolk & Norwich University NHS Foundation Trust), Jon Dorling (Nottingham University Hospitals NHS Trust), Catherine Harrison (Leeds Teaching Hospitals NHS Trust), David Sweet (Royal Maternity Hospital Belfast), Richard Hearn (Royal Victoria Infirmary), Kirsteen Mackay (Sheffield Teaching Hospitals NHS Foundation Trust), Edward Johnstone (St. Mary's Hospital, Central Manchester University Hospitals NHS Foundation Trust), Elaine Boyle (University Hospitals of Leicester NHS Trust) and Lee Abbott (University of North Midlands NHS Trust).

Contributions of authors

The following authors produced the first draft of the report: Paul D Griffiths (Professor of Radiology), Michael Bradburn (Statistician), Cindy L Cooper (Professor of Health Services Research and Clinical Trials), Ruth Graham (Senior Lecturer in Sociology), Laura Mandefield (Statistician) and Cara Mooney (Study Manager).

The following authors conceived or designed the study: Paul D Griffiths (Professor of Radiology), Michael Bradburn (Statistician), Michael J Campbell (Professor of Medical Statistics), Cindy L Cooper (Professor of Health Services Research and Clinical Trials), Nicholas Embleton (Consultant Neonatal Paediatrician),

Ruth Graham (Senior Lecturer in Sociology), Gerald Mason (Consultant in Feto Maternal Medicine), Stephen C Robson (Consultant in Fetal Medicine) and Allan Wailoo (Professor of Health Economics).

The following authors were involved in the acquisition of data for the study: Paul D Griffiths (Professor of Radiology), Nicholas Embleton (Consultant Neonatal Paediatrician), Ruth Graham (Senior Lecturer in Sociology), Anthony R Hart (Consultant Paediatric Neurologist), Deborah Jarvis (Senior Radiographer), Mark D Kilby (Consultant in Maternal and Fetal Medicine), Mabel Lie (Research Associate in Sociology and Social Policy), Gerald Mason (Consultant in Feto Maternal Medicine) and Stephen C Robson (Consultant in Fetal Medicine).

The following authors were involved in the analysis of the data: Paul D Griffiths (Professor of Radiology), Michael Bradburn (Statistician), Ruth Graham (Senior Lecturer in Sociology), Deborah Jarvis (Senior Radiographer), Mabel Lie (Research Associate in Sociology and Social Policy), Laura Mandefield (Statistician), Cara Mooney (Study Manager) and Rebekah Pennington (Research Fellow, Health Economist).

The following authors were involved in the interpretation of data for the study: Paul D Griffiths (Professor of Radiology), Michael Bradburn (Statistician), Nicholas Embleton (Consultant Neonatal Paediatrician), Ruth Graham (Senior Lecturer in Sociology), Anthony R Hart (Consultant Paediatric Neurologist), Mark D Kilby (Consultant in Maternal and Fetal Medicine), Mabel Lie (Research Associate in Sociology and Social Policy), Gerald Mason (Consultant in Feto Maternal Medicine) and Stephen C Robson (Consultant in Fetal Medicine).

The following authors drafted the report: Paul D Griffiths (Professor of Neuroradiology), Michael Bradburn (Statistician), Ruth Graham (Senior Lecturer in Sociology), Laura Mandefield (Statistician), Cara Mooney (Trial Manager) and Rebekah Pennington (Research Fellow, Health Economist).

The following authors revised the report critically for important intellectual content: Paul D Griffiths (Professor of Radiology), Michael Bradburn (Statistician), Michael J Campbell (Professor of Medical Statistics), Cindy L Cooper (Professor of Health Services Research and Clinical Trials), Nicholas Embleton (Consultant Neonatal Paediatrician), Ruth Graham (Senior Lecturer in Sociology), Anthony R Hart (Consultant Paediatric Neurologist), Deborah Jarvis (Senior Radiographer), Mark D Kilby (Consultant in Maternal and Fetal Medicine), Mabel Lie (Research Associate in Sociology and Social Policy), Gerald Mason (Consultant in Feto Maternal Medicine), Laura Mandefield (Statistician), Cara Mooney (Trial Manager), Stephen C Robson (Consultant in Fetal Medicine) and Allan Wailoo (Professor of Health Economics).

The following authors were involved in the final approval of the version of the report to be published: Paul D Griffiths (Professor of Radiology), Michael Bradburn (Statistician), Cindy L Cooper (Professor of Health Services Research and Clinical Trials), Ruth Graham (Senior Lecturer in Sociology), Laura Mandefield (Statistician) and Cara Mooney (Trial Manager).

All authors agree to be accountable for all aspects of the report in ensuring that questions related to the accuracy or integrity of any part of the report are appropriately investigated and resolved.

Publications

Griffiths PD, Brackley K, Bradburn M, Connolly DJA, Gawne-Cain ML, Griffiths DI, *et al.* Anatomical subgroup analysis of the MERIDIAN cohort: failed commissuration. *Ultrasound Obstet Gynecol* 2017;**50**:753–60.

Griffiths PD, Brackley K, Bradburn M, Connolly DJA, Gawne-Cain ML, Griffiths DI, *et al.* Anatomical subgroup analysis of the MERIDIAN cohort: ventriculomegaly. *Ultrasound Obstet Gynecol* 2017;**50**:736–44.

Griffiths PD, Brackley K, Bradburn M, Connolly DJA, Gawne-Cain ML, Kilby M, *et al.* Anatomical subgroup analysis of the MERIDIAN cohort: posterior fossa abnormalities. *Ultrasound Obstet Gynecol* 2017;**50**:745–52.

Griffiths PD, Bradburn M, Campbell MJ, Connolly DJA, Cooper CL, Jarvis D, *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.

Griffiths PD, Bradburn M, Campbell MJ, Cooper CL, Graham R, Jarvis D, *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.

Lie MLS, Graham RH, Robson SC, Griffiths PD, MERIDIAN Collaborative Group. MRI for fetal developmental brain abnormalities: perspectives from the pregnant patient. *Qual Health Res* 2018;**28**:1295–307.

Batty R, Gawne-Cain ML, Mooney C, Mandefield L, Bradburn M, Mason G, Griffiths PD. Analysis of errors made on in utero MR studies of the foetal brain in the MERIDIAN. *Eur Radiol* 2019;**29**:195–201.

Griffiths PD, Bradburn M, Mandefield L, Mooney C, Jarvis D. The rate of brain abnormalities on in utero MR studies in fetuses with normal ultrasound examinations of the brain and calculation of indicators of diagnostic performance. *Clin Radiol* 2019;**74**:527–33.

Lie MLS, Graham RH, Robson S, Griffiths P. ‘He looks gorgeous’: iuMR images and the transforming of foetal and parental identities. *Soc Health* 2019;**41**:360–77.

Griffiths PD on behalf of the MERIDIAN study group. Protocol 11PRT/2491: Magnetic Resonance Imaging to Enhance the Diagnosis of Fetal Developmental Brain Abnormalities in Utero (MERIDIAN) (ISRCTN27626961). *The Lancet*. URL: www.thelancet.com/protocol-reviews/11PRT-2491 (accessed 9 August 2019).

Data-sharing statement

All data requests should be submitted to the corresponding author for consideration. Access to available anonymised data may be granted following review.

Patient data

This work uses data provided by patients and collected by the NHS as part of their care and support. Using patient data is vital to improve health and care for everyone. There is huge potential to make better use of information from people’s patient records, to understand more about disease, develop new treatments, monitor safety, and plan NHS services. Patient data should be kept safe and secure, to protect everyone’s privacy, and it’s important that there are safeguards to make sure that it is stored and used responsibly. Everyone should be able to find out about how patient data are used. #datasaveslives You can find out more about the background to this citation here: <https://understandingpatientdata.org.uk/data-citation>.

References

1. EUROCAT. *Prevalence Tables for Nervous System Anomalies (Per 10,000 Births)*. 2012. URL: www.eurocat-network.eu/prentalascreeingand%20diagnosis (accessed 12 March 2018).
2. Whitby E, Paley MN, Davies N, Sprigg A, Griffiths PD. Ultrafast magnetic resonance imaging of central nervous system abnormalities in utero in the second and third trimester of pregnancy: comparison with ultrasound. *BJOG* 2001;**108**:519–26. <https://doi.org/10.1111/j.1471-0528.2001.00115.x>
3. Hubbard AM, Adzick NS, Crombleholme TM, Coleman BG, Howell LJ, Haselgrove JC, Mahboubi S. Congenital chest lesions: diagnosis and characterization with prenatal MR imaging. *Radiology* 1999;**212**:43–8. <https://doi.org/10.1148/radiology.212.1.r99j3143>
4. Levine D, Barnes P, Korf B, Edelman R. Tuberous sclerosis in the fetus: second-trimester diagnosis of subependymal tubers with ultrafast MR imaging. *AJR Am J Roentgenol* 2000;**175**:1067–9. <https://doi.org/10.2214/ajr.175.4.1751067>
5. Levine D, Barnes PD. Cortical maturation in normal and abnormal fetuses as assessed with prenatal MR imaging. *Radiology* 1999;**210**:751–8. <https://doi.org/10.1148/radiology.210.3.r99mr47751>
6. Levine D, Barnes PD, Madsen JR, Li W, Edelman RR. Fetal central nervous system anomalies: MR imaging augments sonographic diagnosis. *Radiology* 1997;**204**:635–42. <https://doi.org/10.1148/radiology.204.3.9280237>
7. Girard N, Gire C, Sigaudy S, Porcu G, d'Ercole C, Figarella-Branger D, et al. MR imaging of acquired fetal brain disorders. *Childs Nerv Syst* 2003;**19**:490–500. <https://doi.org/10.1007/s00381-003-0761-x>
8. Heilmann G. *MRI of the Fetal Brain*. Berlin: Springer; 2004.
9. Malinger G, Ben-Sira L, Lev D, Ben-Aroya Z, Kidron D, Lerman-Sagie T. Fetal brain imaging: a comparison between magnetic resonance imaging and dedicated neurosonography. *Ultrasound Obstet Gynecol* 2004;**23**:333–40. <https://doi.org/10.1002/uog.1016>
10. Malinger G, Lev D, Lerman-Sagie T. Fetal central nervous system: MR imaging versus dedicated US – need for prospective, blind, comparative studies. *Radiology* 2004;**232**:306. <https://doi.org/10.1148/radiol.2321032051>
11. Whitby EH, Paley MN, Sprigg A, Rutter S, Davies NP, Wilkinson ID, Griffiths PD. Comparison of ultrasound and magnetic resonance imaging in 100 singleton pregnancies with suspected brain abnormalities. *BJOG* 2004;**111**:784–92. <https://doi.org/10.1111/j.1471-0528.2004.00149.x>
12. Griffiths PD, Reeves MJ, Morris JE, Mason G, Russell SA, Paley MN, Whitby EH. A prospective study of fetuses with isolated ventriculomegaly investigated by antenatal sonography and in utero MR imaging. *AJNR Am J Neuroradiol* 2010;**31**:106–11. <https://doi.org/10.3174/ajnr.A1767>
13. Launay S, Robert Y, Valat AS, Thomas D, Devisme L, Rocourt N, Vaast P. Cerebral fetal MRI and ventriculomegaly. *J Radiol* 2002;**83**(6):723–30.
14. Salomon LJ, Ouahba J, Delezoide AL, Vuillard E, Oury JF, Sebag G, Garel C. Third-trimester fetal MRI in isolated 10- to 12-mm ventriculomegaly: is it worth it? *BJOG* 2006;**113**:942–7. <https://doi.org/10.1111/j.1471-0528.2006.01003.x>
15. Mundy L, Hiller J, Braunack-Mayer A, Merlin T. *MRI for the Detection of Foetal Abnormalities*. Adelaide, SA Adelaide Health Technology Assessment; 2007.

16. 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. <https://doi.org/10.1002/uog.13429>
17. van Doorn M, Oude Rengerink K, Newsum EA, Reneman L, Majoie CB, Pajkrt E. 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. <https://doi.org/10.3109/14767058.2015.1109621>
18. Jarvis D, Mooney C, Cohen J, Papaioannou D, Bradburn M, Sutton A, Griffiths PD. 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. <https://doi.org/10.1007/s00330-016-4563-4>
19. Simon EM, Goldstein RB, Coakley FV, Filly RA, Broderick KC, Musci TJ, Barkovich AJ. Fast MR imaging of fetal CNS anomalies in utero. *AJNR Am J Neuroradiol* 2000;**21**:1688–98.
20. Levine D, Barnes PD, Robertson RR, Wong G, Mehta TS. Fast MR imaging of fetal central nervous system abnormalities. *Radiology* 2003;**229**:51–61. <https://doi.org/10.1148/radiol.2291020770>
21. Amini H, Axelsson O, Raiend M, Wikstrom J. The clinical impact of fetal magnetic resonance imaging on management of CNS anomalies in the second trimester of pregnancy. *Acta Obstet Gynecol Scand* 2010;**89**:1571–81. <https://doi.org/10.3109/00016349.2010.526184>
22. Saleem SN, Said AH, Abdel-Raouf M, El-Kattan EA, Zaki MS, Madkour N, et al. Fetal MRI in the evaluation of fetuses referred for sonographically suspected neural tube defects (NTDs): impact on diagnosis and management decision. *Neuroradiology* 2009;**51**:761–72. <https://doi.org/10.1007/s00234-009-0549-0>
23. Twickler DM, Magee KP, Caire J, Zaretsky M, Fleckenstein JL, Ramus RM. Second-opinion magnetic resonance imaging for suspected fetal central nervous system abnormalities. *Am J Obstet Gynecol* 2003;**188**:492–6. <https://doi.org/10.1067/mob.2003.100>
24. Whitby EH, Variend S, Rutter S, Paley MNJ, Wilkinson ID, Davies NP, et al. Corroboration of in utero MRI using post-mortem MRI and autopsy in foetuses with CNS abnormalities. *Clin Radiol* 2004;**59**:1114–20. <https://doi.org/10.1016/j.crad.2004.04.018>
25. Yuh WT, Nguyen HD, Fisher DJ, Turgut Tali E, Gao F, Simonson TM, et al. MR of fetal central nervous system abnormalities. *AJNR Am J Neuroradiol* 1994;**15**:459–64.
26. Hagmann CF, Robertson NJ, Leung WC, Chong KW, Chitty LS. Foetal brain imaging: ultrasound or MRI. A comparison between magnetic resonance imaging and a dedicated multidisciplinary neurosonographic opinion. *Acta Paediatr* 2008;**97**:414–9. <https://doi.org/10.1111/j.1651-2227.2008.00689.x>
27. Ismail KM, Ashworth JR, Martin WL, Chapman S, McHugo J, Whittle MJ, et al. Fetal magnetic resonance imaging in prenatal diagnosis of central nervous system abnormalities: 3-year experience. *J Matern Fetal Neonatal Med* 2002;**12**:185–90. <https://doi.org/10.1080/jmf.12.3.185.190>
28. Peruzzi P, Corbitt RJ, Raffel C. Magnetic resonance imaging versus ultrasonography for the in utero evaluation of central nervous system anomalies. *J Neurosurg Pediatr* 2010;**6**:340–5. <https://doi.org/10.3171/2010.7.PEDS09511>
29. We JS, Young L, Park IY, Shin JC, Im SA. Usefulness of additional fetal magnetic resonance imaging in the prenatal diagnosis of congenital abnormalities. *Arch Gynecol Obstet* 2012;**286**:1443–52. <https://doi.org/10.1007/s00404-012-2474-4>
30. Hosny IA, Elghawabi HS. Ultrafast MRI of the fetus: an increasingly important tool in prenatal diagnosis of congenital anomalies. *Magn Reson Imaging* 2010;**28**:1431–9. <https://doi.org/10.1016/j.mri.2010.06.024>

31. Williams C. Dilemmas in fetal medicine: premature application of technology or responding to women's choice? *Sociol Health Illn* 2006;**28**:1–20.
32. Zakaria S, Brandt KR, Degnim AC, Thomsen KM. Patients' perceptions of breast MRI: a single-center study. *AJR Am J Roentgenol* 2009;**192**:1149–54. <https://doi.org/10.2214/AJR.08.1243>
33. MacKenzie R, Sims C, Owens RG, Dixon AK. Patients' perceptions of magnetic resonance imaging. *Clin Radiol* 1995;**50**:137–43. [https://doi.org/10.1016/S0009-9260\(05\)83042-9](https://doi.org/10.1016/S0009-9260(05)83042-9)
34. Tabak N, Yahel A. Adolescents' prior knowledge about medical imaging and their degree of cooperation with imaging procedures. *Med Law* 2006;**25**:627–45.
35. Middleton WD, Payne WT, Teefey SA, Hildebolt CF, Rubin DA, Yamaguchi K. Sonography and MRI of the shoulder: comparison of patient satisfaction. *AJR Am J Roentgenol* 2004;**183**:1449–52. <https://doi.org/10.2214/ajr.183.5.1831449>
36. Yan T, Wen SW, Walker MC, Beduz MA, Kim PC. Women's satisfaction with the current state of prenatal care for pregnancies complicated by fetal anomalies: a survey of five academic perinatal units in Ontario. *J Obstet Gynaecol Can* 2007;**29**:308–14. [https://doi.org/10.1016/S1701-2163\(16\)32436-7](https://doi.org/10.1016/S1701-2163(16)32436-7)
37. Walker LV, Miller VJ, Dalton VK. The health-care experiences of families given the prenatal diagnosis of trisomy 18. *J Perinatol* 2008;**28**:12–19. <https://doi.org/10.1038/sj.jp.7211860>
38. Leithner K, Assem-Hilger E, Fischer-Kern M, Löffler-Stastka H, Thien R, Ponocny-Seliger E. Prenatal care: the patient's perspective. A qualitative study. *Prenat Diagn* 2006;**26**:931–7. <https://doi.org/10.1002/pd.1529>
39. Leithner K, Pornbacher S, Assem-Hilger E, Krampfl E, Ponocny-Seliger E, Prayer D. Psychological reactions in women undergoing fetal magnetic resonance imaging. *Obstet Gynecol* 2008;**111**:396–402. <https://doi.org/10.1097/AOG.0b013e3181610281>
40. Leithner K, Pörnbacher S, Assem-Hilger E, Krampfl-Bettelheim E, Prayer D. Prenatal magnetic resonance imaging: towards optimized patient information. *Ultrasound Obstet Gynecol* 2009;**34**:182–7. <https://doi.org/10.1002/uog.6391>
41. Adamsbaum C, Garel C, Legros J. Fetal MR: evaluation of patient's experiences and practical implications. *J Radiol* 2008;**89**:791–6. [https://doi.org/10.1016/S0221-0363\(08\)73785-5](https://doi.org/10.1016/S0221-0363(08)73785-5)
42. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ* 2015;**351**:h5527. <https://doi.org/10.1136/bmj.h5527>
43. Likert R. A technique for the measurement of attitudes. *Arch Psychol* 1932;**22**:5–55.
44. Griffiths PD, Bradburn M, Campbell MJ, Cooper CL, Graham R, Jarvis D, 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. [https://doi.org/10.1016/S0140-6736\(16\)31723-8](https://doi.org/10.1016/S0140-6736(16)31723-8)
45. Great Britain. *Abortion Act 1967. Section 1*. London: The Stationery Office; 1967. URL: www.legislation.gov.uk/ukpga/1967/87/section/1 (accessed 25 May 2019).
46. Carroll SG, Porter H, Abdel-Fattah S, Kyle PM, Soothill PW. Correlation of prenatal ultrasound diagnosis and pathologic findings in fetal brain abnormalities. *Ultrasound Obstet Gynecol* 2000;**16**:149–53. <https://doi.org/10.1046/j.1469-0705.2000.00199.x>
47. Sun CC, Grumbach K, DeCosta DT, Meyers CM, Dungan JS. Correlation of prenatal ultrasound diagnosis and pathologic findings in fetal anomalies. *Pediatr Dev Pathol* 1999;**2**:131–42. <https://doi.org/10.1007/s100249900101>

48. Akgun H, Basbug M, Ozgun MT, Canoz O, Tokat F, Murat N, Ozturk F. Correlation between prenatal ultrasound and fetal autopsy findings in fetal anomalies terminated in the second trimester. *Prenat Diagn* 2007;**27**:457–62. <https://doi.org/10.1002/pd.1710>
49. Amini H, Antonsson P, Papadogiannakis N, Ericson K, Pilo C, Eriksson L, et al. Comparison of ultrasound and autopsy findings in pregnancies terminated due to fetal anomalies. *Acta Obstet Gynecol Scand* 2006;**85**:1208–16. <https://doi.org/10.1080/00016340600880886>
50. Boyd PA, Tondi F, Hicks NR, Chamberlain PF. Autopsy after termination of pregnancy for fetal anomaly: retrospective cohort study. *BMJ* 2004;**328**:137. <https://doi.org/10.1136/bmj.37939.570104.EE>
51. Fadda GM, Capobianco G, Balata A, Litta P, Ambrosini G, D'Antona D, et al. Routine second trimester ultrasound screening for prenatal detection of fetal malformations in Sassari University Hospital, Italy: 23 years of experience in 42,256 pregnancies. *Eur J Obstet Gynecol Reprod Biol* 2009;**144**:110–14. <https://doi.org/10.1016/j.ejogrb.2009.02.045>
52. Grandjean H, Larroque D, Levi S. The performance of routine ultrasonographic screening of pregnancies in the Eurofetus Study. *Am J Obstet Gynecol* 1999;**181**:446–54. [https://doi.org/10.1016/S0002-9378\(99\)70577-6](https://doi.org/10.1016/S0002-9378(99)70577-6)
53. Richmond S, Atkins J. A population-based study of the prenatal diagnosis of congenital malformation over 16 years. *BJOG* 2005;**112**:1349–57. <https://doi.org/10.1111/j.1471-0528.2005.00660.x>
54. Northern Regional Survey Steering Group. Fetal abnormality: an audit of its recognition and management. *Arch Dis Child* 1992;**67**:770–4. <https://doi.org/10.1136/adc.67.7.Spec.No.770>
55. Goldstein I, Copel JA, Makhoul IR. Mild cerebral ventriculomegaly in fetuses: characteristics and outcome. *Fetal Diagn Ther* 2005;**20**:281–4. <https://doi.org/10.1159/000085086>
56. Little RJA, Rubin DB. *Statistical Analysis with Missing Data*. 2nd edn. Hoboken, NJ: John Wiley & Sons, Inc.; 2014.
57. Von Hippel PT. Regression with missing ys: an improved strategy for analyzing multiply imputed data. *Sociological Methodology* 2007;**37**:83–117. <https://doi.org/10.1111/j.1467-9531.2007.00180.x>
58. Ng CS, Palmer CR. Analysis of diagnostic confidence and diagnostic accuracy: a unified framework. *Br J Radiol* 2007;**80**:152–60. <https://doi.org/10.1259/bjr/64096611>
59. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand* 1983;**67**:361–70. <https://doi.org/10.1111/j.1600-0447.1983.tb09716.x>
60. Curtis L, Burns A. *Unit Costs of Health and Social Care 2016*. Canterbury: PPSRU, University of Kent; 2016. URL: www.pssru.ac.uk/pub/uc/uc2016/full.pdf?uc=2016-full (accessed 1 October 2017).
61. Department of Health and Social Care. *National Schedule of Reference Costs: The Main Schedule*. 2016. URL: www.gov.uk/government/publications/nhs-reference-costs-2015-to-2016 (accessed 30 October 2017).
62. UK Government. *Expenses and Benefits: Business Travel Mileage for Employees' Own Vehicles*. 2019. URL: www.gov.uk/expenses-and-benefits-business-travel-mileage/rules-for-tax (accessed 17 July 2019).
63. Smith R, Office for National Statistics. *Earnings and Hours Worked, Age Group: ASHE Table 6*. 2017. URL: www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/datasets/agegroupashetable6 (accessed 31 October 2017).

64. Royal College of Obstetricians and Gynaecologists. *Termination of Pregnancy for Fetal Abnormality in England, Scotland and Wales*. 2010. URL: www.rcog.org.uk/en/guidelines-research-services/guidelines/termination-of-pregnancy-for-fetal-abnormality-in-england-scotland-and-wales/ (accessed 12 March 2018).
65. Petrou S, Johnson S, Wolke D, Hollis C, Kochhar P, Marlow N. Economic costs and preference-based health-related quality of life outcomes associated with childhood psychiatric disorders. *Br J Psychiatry* 2010;**197**:395–404. <https://doi.org/10.1192/bjp.bp.110.081307>
66. National Institute for Health and Care Excellence (NICE). *Guide to the Methods of Technology Appraisal 2013*. London: NICE; 2013. URL: www.nice.org.uk/process/pmg9/chapter/the-reference-case (accessed 20 December 2017).
67. Griffiths PD, Brackley K, Bradburn M, Connolly DJA, Gawne-Cain ML, Griffiths DI, et al. Anatomical subgroup analysis of the MERIDIAN cohort: ventriculomegaly. *Ultrasound Obstet Gynecol* 2017;**50**:736–44. <https://doi.org/10.1002/uog.17475>
68. Griffiths PD, Brackley K, Bradburn M, Connolly DJA, Gawne-Cain ML, Griffiths DI, et al. Anatomical subgroup analysis of the MERIDIAN cohort: failed commissuration. *Ultrasound Obstet Gynecol* 2017;**50**:753–60. <https://doi.org/10.1002/uog.17502>
69. Griffiths PD, Brackley K, Bradburn M, Connolly DJA, Gawne-Cain ML, Kilby MD, et al. Anatomical subgroup analysis of the MERIDIAN cohort: posterior fossa abnormalities. *Ultrasound Obstet Gynecol* 2017;**50**:745–52. <https://doi.org/10.1002/uog.17485>
70. Fanou EM, Reeves MJ, Howe DT, Joy H, Morris S, Russell S, Griffiths PD. In utero magnetic resonance imaging for diagnosis of dural venous sinus ectasia with thrombosis in the fetus. *Pediatr Radiol* 2013;**43**:1591–8. <https://doi.org/10.1007/s00247-013-2745-7>
71. Batty R, Gawne-Cain ML, Mooney C, Mandfield L, Bradburn M, Mason G, Griffiths PD. Analysis of errors made on in utero MR studies of the foetal brain in the MERIDIAN study. *Eur Radiol* 2018;**29**:195–201. <https://doi.org/10.1007/s00330-018-5508-x>
72. Leithner K, Prayer D, Porstner E, Kapusta ND, Stammer-Safar M, Krampfl-Bettelheim E, Hilger E. Psychological reactions related to fetal magnetic resonance imaging: a follow-up study. *J Perinat Med* 2013;**41**:273–6. <https://doi.org/10.1515/jpm-2012-0218>
73. Reed K, Kochetkova I, Whitby E. Visualising uncertainty: examining women’s views on the role of magnetic resonance imaging (MRI) in late pregnancy. *Soc Sci Med* 2016;**164**:19–26. <https://doi.org/10.1016/j.socscimed.2016.07.012>
74. Joyce K. Appealing images: magnetic resonance imaging and the production of authoritative knowledge. *Soc Stud Sci* 2005;**35**:437–62. <https://doi.org/10.1177/0306312705050180>
75. Joyce KA. From numbers to pictures: the development of magnetic resonance imaging and the visual turn in medicine. *Science as Culture* 2006;**15**:1–22. <https://doi.org/10.1080/09505430600639322>
76. Reed K, Kochetkova I, Molyneux-Hodgson S. ‘You’re looking for different parts in a jigsaw’: foetal MRI (magnetic resonance imaging) as an emerging technology in professional practice. *Sociol Health Illn* 2016;**38**:736–52. <https://doi.org/10.1111/1467-9566.12398>
77. Jutel A. Sociology of diagnosis: a preliminary review. *Sociol Health Illn* 2009;**31**:278–99. <https://doi.org/10.1111/j.1467-9566.2008.01152.x>
78. Patton MQ. *Qualitative Evaluation and Research Methods*. 2nd edn. Newbury Park, CA: Sage; 1990.
79. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;**3**:77–101. <https://doi.org/10.1191/1478088706qp0630a>

80. Braun V, Clarke V. What can 'thematic analysis' offer health and wellbeing researchers? *Int J Qual Stud Health Well-being* 2014;**9**:26152. <https://doi.org/10.3402/qhw.v9.26152>
81. Silverman D. *Doing Qualitative Research: A Practical Handbook*. London: Sage; 2000.
82. Silverman D. *Interpreting Qualitative Data: Methods for Analysing Talk, Text and Interaction*. 2nd edn. London: Sage; 2001. <https://doi.org/10.1177/146879410200200111>
83. Glaser BG, Strauss AL. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago, IL: Aldine; 1967. <https://doi.org/10.1097/00006199-196807000-00014>
84. Pope C, Ziebland S, Mays N. Analysing qualitative data. *Qualitative research in health care. BMJ* 2000;**320**:114–6. <https://doi.org/10.1136/bmj.320.7227.114>
85. Miquel-Verges F, Woods SL, Aucott SW, Boss RD, Sulpar LJ, Donohue PK. Prenatal consultation with a neonatologist for congenital anomalies: parental perceptions. *Pediatrics* 2009;**124**:e573–9. <https://doi.org/10.1542/peds.2008-2865>
86. McCoyd JL. "I'm not a saint": burden assessment as an unrecognized factor in prenatal decision making. *Qual Health Res* 2008;**18**:1489–500. <https://doi.org/10.1177/1049732308325642>
87. Palisano R, Rosenbaum P, Bartlett D, Livingston M. *Gross Motor Function Classification System – Expanded and Revised* Hamilton, ON: CanChild Centre for Childhood Disability Research; 2007.
88. Newcombe RG. Improved confidence intervals for the difference between binomial proportions based on paired data. *Stat Med* 1998;**17**:2635–50. [https://doi.org/10.1002/\(SICI\)1097-0258\(19981130\)17:22<2635::AID-SIM954>3.0.CO;2-C](https://doi.org/10.1002/(SICI)1097-0258(19981130)17:22<2635::AID-SIM954>3.0.CO;2-C)
89. Eypasch E, Lefering R, Kum CK, Troidl H. Probability of adverse events that have not yet occurred: a statistical reminder. *BMJ* 1995;**311**:619–20. <https://doi.org/10.1136/bmj.311.7005.619>
90. Loughna P. Congenital abnormalities: failure to detect and treat. *The Obstetrician & Gynaecologist* 2008;**10**:33–7. <https://doi.org/10.1576/toag.10.1.033.27375>
91. Kelly EN, Allen VM, Seaward G, Windrim R, Ryan G. Mild ventriculomegaly in the fetus, natural history, associated findings and outcome of isolated mild ventriculomegaly: a literature review. *Prenat Diagn* 2001;**21**:697–700. <https://doi.org/10.1002/pd.138>
92. Melchiorre K, Bhide A, Gika AD, Pilu G, Papageorgiou AT. Counselling in isolated mild fetal ventriculomegaly. *Ultrasound Obstet Gynecol* 2009;**34**:212–24. <https://doi.org/10.1002/uog.7307>
93. Gaglioti P, Danelon D, Bontempo S, Mombro M, Cardaropoli S, Todros T. Fetal cerebral ventriculomegaly: outcome in 176 cases. *Ultrasound Obstet Gynecol* 2005;**25**:372–7. <https://doi.org/10.1002/uog.1857>
94. Baggot R. *Health and Health Care in Britain*. 3rd edn. Basingstoke: Palgrave Macmillan; 2004.
95. Ham C. *Health Policy in Britain*. 6th edn. Basingstoke: Palgrave Macmillan; 2009. <https://doi.org/10.1007/978-1-137-01397-2>
96. Gilmore EC, Walsh CA. Genetic causes of microcephaly and lessons for neuronal development. *Wiley Interdiscip Rev Dev Biol* 2013;**2**:461–78. <https://doi.org/10.1002/wdev.89>
97. Deloison B, Chalouhi GE, Bernard JP, Ville Y, Salomon LJ. Outcomes of fetuses with small head circumference on second-trimester ultrasonography. *Prenat Diagn* 2012;**32**:869–74. <https://doi.org/10.1002/pd.3923>
98. von Koch CS, Glenn OA, Goldstein RB, Barkovich AJ. Fetal magnetic resonance imaging enhances detection of spinal cord anomalies in patients with sonographically detected bony anomalies of the spine. *J Ultrasound Med* 2005;**24**:781–9. <https://doi.org/10.7863/jum.2005.24.6.781>
99. Department of Health and Social Care. *NHS Reference Costs 2015 to 2016*. 2016. URL: www.gov.uk/government/publications/nhs-reference-costs-2015-to-2016 (accessed 30 October 2017).

100. Great Ormond Street Hospital for Children. *Pricing 2014/15*. URL: www.labs.gosh.nhs.uk/media/534948/nhs_joint_price_list_201415_dec14.pdf (accessed 5 June 2019).
101. Birmingham Women's and Children's NHS Foundation Trust. *Private Tests Guide Price List: West Midlands Regional Genetics Laboratory*. Birmingham: Birmingham Women's and Children's NHS Foundation Trust; 2014. URL: www.bwnft.nhs.uk/wp-content/uploads/2014/10/PP_Prices_2014_v2.pdf (accessed 1 October 2017).
102. NHS UK Genetic Texting Network. *Skeletal Dysplasias, FGFR3-Related (NIPD – Diagnostic Testing): Prenatal Diagnosis by NIPD at London North East RGC GOSH in 5 Days*. 2017. URL: <https://ukgtn.nhs.uk/find-a-test/search-by-disorder-gene/details/4741/> (accessed 1 October 2017).
103. Rutherford B. *Which Screening Tests are Available Privately?* 2016. URL: www.babycentre.co.uk/x557433/which-screening-tests-are-available-privately (accessed 1 October 2017).
104. Sagoo GS, Mohammed S. *Array CGH Testing for Learning Disability: When is it Worth it?* Cambridge: PHG Foundation; 2014. URL: https://ukgtn.nhs.uk/fileadmin/uploads/ukgtn/Documents/Resources/Library/Reports_Guidelines/Array%20CGH%20testing%20for%20learning%20disability-when%20is%20it%20worth%20it%20Nov%202014.pdf (accessed 1 October 2017).

Appendix 1 Unit costs (base case)

Cost component	Details	Cost (£)	Source
Consultation	First consultation with consultant	161	<i>NHS Reference Costs 2015 to 2016</i> : ⁹⁹ WF01B (Obstetrics)
	Subsequent consultation with consultant	121	<i>NHS Reference Costs 2015 to 2016</i> : ⁹⁹ WF01A (Obstetrics)
Ultrasonography scan	Two-dimensional ultrasonography	106	<i>NHS Reference Costs 2015 to 2016</i> : ⁹⁹ NZ21Z Outpatient procedure (Obstetrics)
	Three-dimensional ultrasonography	146	<i>NHS Reference Costs 2015 to 2016</i> : ⁹⁹ NZ22Z Outpatient procedure (Obstetrics)
MRI	iuMRI (≥ 19 years)	146	<i>NHS Reference Costs 2015 to 2016</i> : ⁹⁹ RA01A – single area, no contrast
	iuMRI (< 19 years)	143	<i>NHS Reference Costs 2015 to 2016</i> : ⁹⁹ RA01B – single area, no contrast
Termination	TOP with feticide	1581	<i>NHS Reference Costs 2015 to 2016</i> : ⁹⁹ MA20Z – ≥ 20 -week gestation. Elective inpatient plus 0.84 excess bed-days
	TOP without feticide	1136	<i>NHS Reference Costs 2015 to 2016</i> : ⁹⁹ MA18D – 14- to 20-week gestation. Elective inpatient plus 0.82 excess bed-days
Delivery	Delivery	2647	<i>NHS Reference Costs 2015 to 2016</i> : ⁹⁹ NZ25Z-NZ15C weighted average
Investigations	Karyotype	150	Great Ormond Street Hospital ¹⁰⁰
	Fetal echocardiography	106	Assumed the same as a two-dimensional ultrasonography
	TORCH screen	8	Sheffield Teaching Hospitals NHS Foundation Trust ^a
	Antibody screen	9	Sheffield Teaching Hospitals NHS Foundation Trust ^a
	Cystic fibrosis screen	257	West Midlands Regional Genetics Laboratory ¹⁰¹
	<i>FGFR3</i> – related skeletal dysplasias panel test	600	NHS UK Genetic Testing Network ¹⁰²
	Non-invasive prenatal testing	400	BabyCentre ¹⁰³
	Array	291	Sagoo and Mohammed ¹⁰⁴
	CMV	17	Sheffield Teaching Hospitals NHS Foundation Trust ^a
Toxoplasmosis	9	Sheffield Teaching Hospitals NHS Foundation Trust ^a	

CMV, cytomegalovirus; *FGFR3*, fibroblast growth factor receptor 3; TORCH, toxoplasmosis, rubella cytomegalovirus, herpes simplex and human immunodeficiency virus (HIV).

a Sheffield Teaching Hospitals, personal communication, 2016.

Appendix 2 Probability distributions for input parameters

Parameter category	Parameter	Distribution	Mean	Standard error	Alpha	Beta
Cost (£)	Consultant consultation: first	Normal	161	0		
	Consultant consultation: subsequent	Normal	121	0		
	Two-dimensional ultrasonography	Normal	106	0		
	Three-dimensional ultrasonography	Normal	146	0		
	MRI (≥ 19 years)	Normal	146	0		
	MRI (6–18 years)	Normal	143	0		
	TOP with feticide	Normal				
	TOP average stay (≥ 20 weeks' gestation)	Normal	1232	28		
	TOP excess bed-day (≥ 20 weeks' gestation)	Normal	416	10		
	TOP without feticide	Normal				
	TOP average stay (14–20 weeks' gestation)	Normal	910	42		
	TOP excess bed-day (14–20 weeks' gestation)	Normal	277	36		
	Delivery cost (average)	Normal	2647	0		
	Karyotype	Normal	152	30		
	Fetal echocardiography	Normal	106	0		
	TORCH screen	Normal	8	2		
	Antibody screen	Normal	9	2		
	Cystic fibrosis screen	Normal	263	53		
	<i>FGFR3</i> – related skeletal dysplasias panel test	Normal	600	120		
	Non-invasive prenatal testing	Normal	400	80		
	Array	Normal	298	60		
	Cytomegalovirus	Normal	17	3		
	Toxoplasmosis	Normal	9	2		
Resource use	Postcode miles: no TOP (miles)	Log-normal	58.39	3.75		
	Postcode miles: TOP (miles)	Log-normal	82.57	13.97		
	Hourly wage: no TOP (£)	Log-normal	13.32	0.20		
	Hourly wage: TOP (£)	Log-normal	14.68	0.49		
	Participant time: no TOP (minutes)	Log-normal	229.67	12.19		
	Participant time: TOP (minutes)	Log-normal	283.25	46.82		
	Proportion of patients, two-dimensional ultrasonography: no TOP	Log-normal	0.97	0.01		
	Proportion of patients, two-dimensional ultrasonography: TOP	Log-normal	1.00	0.00		
	Proportion of patients, three-dimensional ultrasonography: no TOP	Log-normal	0.11	0.02		

Parameter category	Parameter	Distribution	Mean	Standard error	Alpha	Beta
	Proportion of patients, three-dimensional ultrasonography: TOP	Log-normal	0.05	0.05		
	Proportion of patients, < 19 years: no TOP	Log-normal	1.00	0.00		
	Proportion of patients, < 19 years: TOP	Log-normal	1.00	0.00		
	Proportion of patients, follow-up two-dimensional ultrasonography: no TOP	Log-normal	0.72	0.03		
	Proportion of patients, follow-up two-dimensional ultrasonography: TOP	Log-normal	0.20	0.09		
	Proportion of patients, follow-up iuMRI: no TOP	Log-normal	0.12	0.02		
	Proportion of patients, follow-up iuMRI: TOP	Log-normal	0.15	0.08		
	Additional visits: no TOP (n)	Log-normal	3.16	0.11		
	Additional visits: TOP (n)	Log-normal	2.10	0.07		
	Proportion of patients, karyotype: no TOP	Log-normal	0.38	0.04		
	Proportion of patients, karyotype: TOP	Log-normal	0.80	0.12		
	Proportion of patients, fetal echocardiography: no TOP	Log-normal	0.26	0.03		
	Proportion of patients, fetal echocardiography: TOP	Log-normal	0.20	0.09		
	Proportion of patients, infection screen: no TOP	Log-normal	0.85	0.03		
	Proportion of patients, infection screen: TOP	Log-normal	0.75	0.12		
	Proportion of patients, NIPT: no TOP	Log-normal	0.26	0.03		
	Proportion of patients, NIPT: TOP	Log-normal	0.20	0.09		
	Proportion of patients, antibody testing: no TOP	Log-normal	0.02	0.01		
	Proportion of patients, antibody testing: TOP	Log-normal	0.10	0.07		
	Proportion of patients, cystic fibrosis screen: no TOP	Log-normal	0.02	0.01		
	Proportion of patients, cystic fibrosis screen: TOP	Log-normal	0.05	0.05		
	Proportion of patients, platelets: no TOP	Log-normal	0.00	0.00		
	Proportion of patients, platelets: TOP	Log-normal	0.00	0.00		
	Proportion of patients, FGFR3: no TOP	Log-normal	0.00	0.00		
	Proportion of patients, FGFR3: TOP	Log-normal	0.00	0.00		
	Proportion of patients, NIFTY: no TOP	Log-normal	0.00	0.00		
	Proportion of patients, NIFTY: TOP	Log-normal	0.00	0.00		
	Proportion of patients, array: no TOP	Log-normal	0.00	0.00		
	Proportion of patients, array: TOP	Log-normal	0.00	0.00		
	Proportion of patients, cytomegalovirus: no TOP	Log-normal	0.00	0.00		
	Proportion of patients, cytomegalovirus: TOP	Log-normal	0.00	0.00		
	Proportion of patients, toxoplasmosis: no TOP	Log-normal	0.00	0.00		
	Proportion of patients, toxoplasmosis: TOP	Log-normal	0.00	0.00		
	Proportion of patients, TOP with feticide: no TOP	Log-normal	0.00	0.00		
	Proportion of patients, TOP with feticide: TOP	Log-normal	0.55	0.11		

Parameter category	Parameter	Distribution	Mean	Standard error	Alpha	Beta
	Proportion of patients, TOP without feticide: no TOP	Log-normal	0.00	0.00		
	Proportion of patients, TOP without feticide: TOP	Log-normal	0.20	0.09		
	Proportion of patients, delivery: no TOP	Log-normal	1.00	0.00		
	Proportion of patients, delivery: TOP	Log-normal	0.00	0.00		
Proportions	iuMRI: proportion having TOP	Beta	0.119		68	502
	Ultrasonography: proportion of patients offered TOP	Beta	0.247		135	412
	Proportion of management decisions appropriately changed	Beta	0.218		120	431
	iuMRI: proportion of patients offered TOP	Beta	0.352		189	348
	iuMRI: proportion of patients offered TOP who then had TOP	Beta	0.307		58	131

FGFR3, fibroblast growth factor receptor 3; NIFTY, non-invasive fetal trisomy test (a type of NIPT); NIPT, non-invasive prenatal testing; TORCH, toxoplasmosis, rubella, cytomegalovirus, herpes simplex, and human immunodeficiency virus (HIV).

Appendix 3 Recruitment and participant flow

The reasons given on the approach form for not asking about entry to study and reasons given by mothers who declined to be in the study.

Reason	n (%)
<i>N</i>	198
Not asked to be in the study (N = 86)	
Considered clinically inappropriate/unhelpful	50 (58.0)
Likely to cause undue delay	6 (7.0)
Other	30 (35.0)
Asked to be in study but refused (N = 112)	
Decision already made	18 (16.0)
Unwilling to travel	13 (12.0)
Does not want iuMRI	28 (25.0)
Too upset or distressed	11 (9.8)
Other	35 (31.0)
No reason given	7 (6.2)

Appendix 4 Ultrasonography technical factors

The technical features reported during ultrasonography.

Technical feature	<i>n</i> (%)
Number completing ultrasonography and iuMRI (<i>N</i>)	821
Two-dimensional ultrasonography only	681 (83.0)
Both two- and three-dimensional ultrasonography	140 (17.0)
Technical factors contributing to low confidence	237 (29.0)
Fetal position	115 (14.0)
High body mass index	107 (13.0)
Oligohydramnios	7 (0.9)
Gestational age	21 (2.6)
Other	50 (6.1)

Appendix 5 Diagnostic accuracy

The diagnostic accuracy of ultrasonography and iuMRI in 570 cases and three subgroups within the primary analysis population (i.e. first, single and repeat pregnancies within the MERIDIAN study population).

Pregnancy characteristic	Ultrasonography correct, <i>n</i> (%)	iuMRI correct, <i>n</i> (%)	Percentage difference (95% CI)	<i>p</i> -value ^a
Overall (<i>N</i> = 570)	388 (68.1)	530 (93.0)	24.9 (21.1 to 28.7)	< 0.0001
First pregnancies (<i>N</i> = 563)	383 (68.0)	525 (93.3)	25.3 (21.4 to 29.0)	< 0.0001
Singleton pregnancies (<i>N</i> = 539)	363 (67.3)	499 (92.6)	25.3 (21.3 to 29.2)	< 0.0001
Repeat pregnancies (<i>N</i> = 7)	5 (71.4)	5 (71.4)	0	

^a McNemar's test between ultrasonography and iuMRI.

Appendix 6 Prognosis in relation to inclusion/exclusion in primary analysis

Data are presented for all mothers who had complete data for clinical impact ($n = 783$). For multiple pregnancies, prognosis is given in relation to the pregnancy as a whole.

Prognosis	ORD available, n (%) ($N = 544$)	ORD unavailable, n (%) ($N = 162$)	Excluded, n (%) ($N = 77$)	Overall, n (%) ($N = 784$) ^a
Prognosis on ultrasonography				
Not known	90 (17.0)	38 (24.0)	18 (23.0)	146 (19.0)
Poor (< 50%)	72 (13.0)	42 (26.0)	6 (7.8)	120 (15.0)
Intermediate (50–90%)	155 (29.0)	47 (29.0)	18 (23.0)	220 (28.0)
Favourable (> 90%)	214 (39.0)	29 (18.0)	34 (44.0)	278 (36.0)
Normal	13 (2.4)	6 (3.7)	1 (1.3)	20 (2.6)
Prognosis on iuMRI				
Not known	41 (7.5)	15 (9.3)	9 (12.0)	65 (8.3)
Poor (< 50%)	112 (20.0)	64 (40.0)	12 (16.0)	188 (24.0)
Intermediate (50–90%)	116 (21.0)	27 (17.0)	12 (16.0)	155 (20.0)
Favourable (> 90%)	251 (46.0)	39 (24.0)	38 (49.0)	329 (42.0)
Normal	24 (4.4)	17 (11.0)	6 (7.8)	47 (6.0)

^a One participant had one fetus in the 'ORD available' category and one fetus in the 'ORD unavailable' category.

Appendix 7 Prognosis on ultrasonography alone versus in utero magnetic resonance imaging following ultrasonography for all mothers who completed in utero magnetic resonance imaging

Ultrasonography alone	iuMRI prognosis (confidence), <i>n</i> (%)					Total, <i>n</i> (%)
	Not known	Poor (< 50%)	Intermediate (50–90%)	Favourable (> 90%)	Normal	
Not known	32 (4.1)	28 (3.6)	32 (4.1)	40 (5.1)	13 (1.7)	145 (18.5)
Poor (< 50%)	4 (0.5)	94 (12)	15 (1.9)	5 (0.6)	2 (0.3)	120 (15.3)
Intermediate (50–90%)	16 (2.0)	57 (7.3)	88 (11.2)	53 (6.8)	6 (0.8)	220 (28.1)
Favourable (> 90%)	12 (1.5)	6 (0.8)	17 (2.2)	222 (28.4)	21 (2.7)	278 (35.5)
Normal	1 (0.1)	2 (0.3)	3 (0.4)	9 (1.1)	5 (0.6)	20 (2.6)
Total	65 (8.3)	187 (23.9)	155 (19.8)	329 (42.0)	47 (6.0)	783 (100.0)

Shading denotes where prognosis on ultrasonography and iuMR match.

Appendix 8 Change in prognosis from ultrasonography to in utero magnetic resonance imaging following ultrasonography by category when prognosis is unknown on one modality

Prognosis	<i>n</i> (%)
Number of fetuses	178
Not known on both ultrasonography and iuMRI	32 (18.0)
Not known on ultrasonography, known on iuMRI	113 (63.0)
Prognosis on iuMRI	
Poor (< 50%)	28 (16.0)
Intermediate (50–90%)	32 (18.0)
Favourable (> 90%)	40 (23.0)
Normal	13 (7.3)
Known on ultrasonography, not known on iuMRI	33 (18.5)
Prognosis on ultrasonography	
Poor (< 50%)	4 (2.2)
Intermediate (50–90%)	16 (9.0)
Favourable (> 90%)	12 (6.7)
Normal	1 (0.6)

Appendix 9 Change in counselling based on ultrasonography alone and in utero magnetic resonance imaging following ultrasonography in 774 cases with complete data

Counselling	TOP not discussed, <i>n</i> (%)	TOP discussed, <i>n</i> (%)			Total, <i>n</i> (%)
		Not offered	Offered based on substantial disability risk	Offered, other reason	
Ultrasonography					
TOP not discussed	285 (73.0)	41 (42.0)	52 (19.0)	3 (21.0)	381 (49.0)
TOP discussed					
Not offered	87 (22.0)	39 (40.0)	64 (24.0)	4 (29.0)	194 (25.0)
Offered based on substantial disability risk	18 (4.6)	10 (10.2)	143 (53.0)	5 (36.0)	176 (22.0)
Offered, other reason	3 (0.8)	8 (8.2)	10 (3.7)	2 (14.0)	23 (2.9)
Total	393 (51.0)	98 (13.0)	269 (35.0)	14 (1.8)	774

Appendix 10 Composite case study

Composite case study

Anne and Jo were looking forward to their 20-week scan, as they were keen to find out if their baby was going to be a girl. They had already had two sons, and they had tried for some time for a third child without success. At the abnormality scan, the sonographer was more quiet than usual and this troubled them. She then left suddenly and came back with a midwife who also looked concerned. Not knowing what to think, Anne started to become tearful and the midwife tried to explain by saying that there was something not quite right with the baby's brain, and that it could mean that the baby had Down syndrome. She could not say much more when asked, and told them that the consultant would be able to give them a fuller explanation. Unfortunately, as it was a Friday, the next available appointment with the consultant was on Monday. They endured a terrible weekend worrying about the condition and all that it could mean for their baby.

On the Monday, the consultant did the scan herself and explained the features in the brain that were of concern. They were told that there was no other test apart from an amniocentesis that could determine if the baby had Down syndrome, so they went ahead with the test. The results came back clear, so the consultant explained the option of having a MRI scan through the MERIDIAN study, a diagnostic accuracy trial. The scan would be able to determine if the results of the ultrasound scan were accurate, and would be able to look at other parts of the brain to see if it was developing normally. Anne was somewhat claustrophobic so she was quite unwilling at first, but then, because they were desperate to know the condition of their baby's brain, they agreed to participate in the study. The scan was arranged for the following Monday.

Before the scan, Anne read the generic MRI information sheet that indicated that pregnant women should take extra care, especially in the first trimester. When she brought it up, she was told that no studies have been conducted on those in the second trimester. This still raised alarm bells for her and caused her to have doubts. She arrived at the room with her husband, expecting her husband to be allowed in with her, but as they ushered him into the waiting room, she felt she could not object. Going into the machine, she started to have a panic attack and asked to be taken out. She was given a pillow and asked to look out of the tunnel or to close her eyes altogether, which helped her. But she still missed her husband's presence. The noise from the machine was so loud that she felt the baby reacting to it, which caused her some concern. But she tried to keep as still as she could to minimise any problems they were having in taking images of the baby's brain. This was difficult for her because she had problems with her pelvis, which was giving her pain, and the heat in the machine did not help. She was in the machine for over 1 hour.

It took over 1 week before the results of the scan were discussed with her. At that appointment, they were told that the 'ventriculomegaly' was borderline and that she would have to be monitored over the course of the pregnancy, as the condition could stay the same, worsen or improve. The idea that it could worsen and could mean ending the pregnancy hung over them, and because of the uncertainty of everything, they preferred to keep it to themselves. As they went in for scans every fortnight, the size of the ventricles varied throughout the pregnancy. Eventually, their baby girl was born with no observable problems. But because of their whole experience, they found that the bond with their baby had been affected and that the whole pregnancy was ruined.

Following the birth, there was very little follow-up on their baby's case. Anne had to push for an ultrasound scan to be done and to explain why it needed to be done. The baby was scanned at 15 days old, but Anne and Jo had to wait until she was 12 weeks old before they were able to find out at a fetal medicine consultation that the ventricles were within the normal range. Now, 6 months after the MRI, there are signs that the baby has a mild hydrocephalus and her development needs to be monitored. They have been looking for a written report and images of the MRI scan but were disappointed that these were not provided. They are also no clearer as to how the condition arose.

Appendix 11 Participant characteristic tables

TABLE 46 Inclusion in the follow-on study project 1 in relation to original prognosis

Prognosis	Overall, n (%)	Consented, n (%)	Did not consent, n (%)	Died within original follow-up, ^a n (%)	Not approached, n (%)
<i>N</i>	829	239	316	188	86
Prognosis on ultrasonography					
Not known	155 (19.0)	45 (19.0)	52 (17.0)	34 (18.0)	24 (28.0)
Poor (< 50%)	125 (15.0)	19 (7.9)	26 (8.2)	69 (37.0)	11 (13.0)
Intermediate (50–90%)	233 (28.0)	59 (25.0)	95 (30.0)	59 (31.0)	20 (23.0)
Favourable (> 90%)	295 (36.0)	107 (45.0)	138 (44.0)	21 (11.0)	29 (34.0)
Normal	21 (2.5)	9 (3.8)	5 (1.6)	5 (2.7)	2 (2.3)
Prognosis on ultrasonography and iuMRI					
Not known	65 (8.2)	16 (7.1)	20 (6.5)	16 (9.1)	13 (16.0)
Poor (< 50%)	189 (24.0)	22 (9.7)	38 (12.0)	116 (66.0)	13 (16.0)
Intermediate (50–90%)	157 (20.0)	43 (19.0)	65 (21.0)	27 (15.0)	22 (27.0)
Favourable (> 90%)	334 (42.0)	129 (57.0)	164 (53.0)	12 (6.9)	29 (36.0)
Normal	47 (5.9)	16 (7.1)	23 (7.4)	4 (2.3)	4 (4.9)

^a Deaths within original MERIDIAN include terminations, stillbirths, fetal deaths and deaths within 6 months of birth.

Note

Values do not total 100% owing to rounding.

TABLE 47 Inclusion in the follow-on study project 2 in relation to the original prognosis

Prognosis	Overall, n (%)	Consented, n (%)	Did not consent, n (%)	Died within original follow-up, ^a n (%)	Not approached, n (%)
<i>N</i>	829	194	361	188	86
Prognosis on ultrasonography					
Not known	155 (19.0)	41 (21.0)	56 (16.0)	34 (18.0)	24 (28.0)
Poor (< 50%)	125 (15.0)	16 (8.2)	29 (8.0)	69 (37.0)	11 (13.0)
Intermediate (50–90%)	233 (28.0)	49 (25.0)	105 (29.0)	59 (31.0)	20 (23.0)
Favourable (> 90%)	295 (36.0)	80 (41.0)	165 (46.0)	21 (11.0)	29 (34.0)
Normal	21 (2.5)	8 (4.1)	6 (1.7)	5 (2.7)	2 (2.3)
Prognosis on iuMRI					
Not known	65 (8.2)	16 (8.7)	20 (5.7)	16 (9.1)	13 (16.0)
Poor (< 50%)	189 (24.0)	18 (9.8)	42 (12.0)	116 (66.0)	13 (16.0)
Intermediate (50–90%)	157 (20.0)	36 (20.0)	72 (20.0)	27 (15.0)	22 (27.0)
Favourable (> 90%)	334 (42.0)	99 (54.0)	194 (55.0)	12 (6.9)	29 (36.0)
Normal	47 (5.9)	14 (7.7)	25 (7.1)	4 (2.3)	4 (4.9)

^a Deaths within original MERIDIAN include terminations, stillbirths, fetal deaths and deaths within 6 months of birth.

Note

Values do not total 100% owing to rounding.

TABLE 48 Inclusion in the follow-on study project 2 in relation to the original prognosis

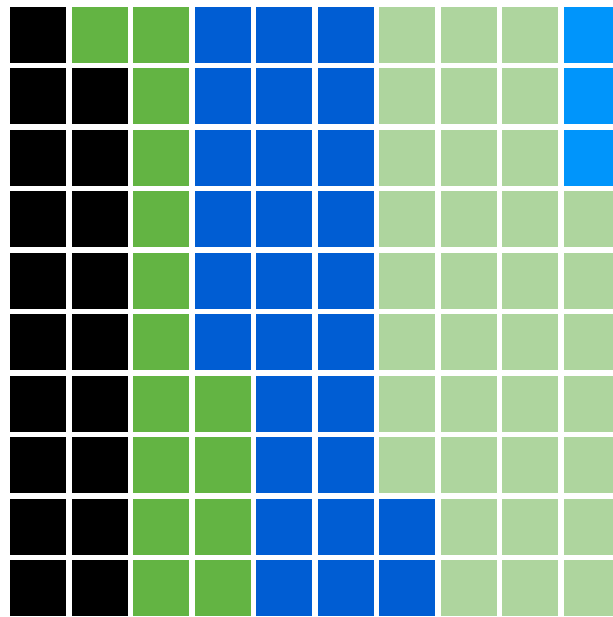
Prognosis	<i>N</i>	Consented, <i>n</i> (%)	Did not consent, <i>n</i> (%)	Died within original follow-up, ^a <i>n</i> (%)	Not approached, <i>n</i> (%)
Ultrasonography prognosis					
Not known	155	41 (27.0)	56 (36.0)	34 (22.0)	24 (16.0)
Poor (< 50%)	125	16 (13.0)	29 (23.0)	69 (55.0)	11 (8.8)
Intermediate (50–90%)	233	49 (21.0)	105 (45.0)	59 (25.0)	20 (8.6)
Favourable (> 90%)	295	80 (27.0)	165 (56.0)	21 (7.1)	29 (9.8)
Normal	21	8 (38.0)	6 (29.0)	5 (24.0)	2 (9.5)
iuMRI prognosis (%)					
Not known	65	16 (25.0)	20 (31.0)	16 (25.0)	13 (20.0)
Poor (< 50%)	189	18 (9.5)	42 (22.0)	116 (61.0)	13 (6.9)
Intermediate (50–90%)	157	36 (23.0)	72 (46.0)	27 (17.0)	22 (14.0)
Favourable (> 90%)	334	99 (30.0)	194 (58.0)	12 (3.6)	29 (8.7)
Normal	47	14 (30.0)	25 (53.0)	4 (8.5)	4 (8.5)

^a Deaths within original MERIDIAN include terminations, stillbirths, fetal deaths and deaths within 6 months of birth.

Note

Values do not total 100% owing to rounding.

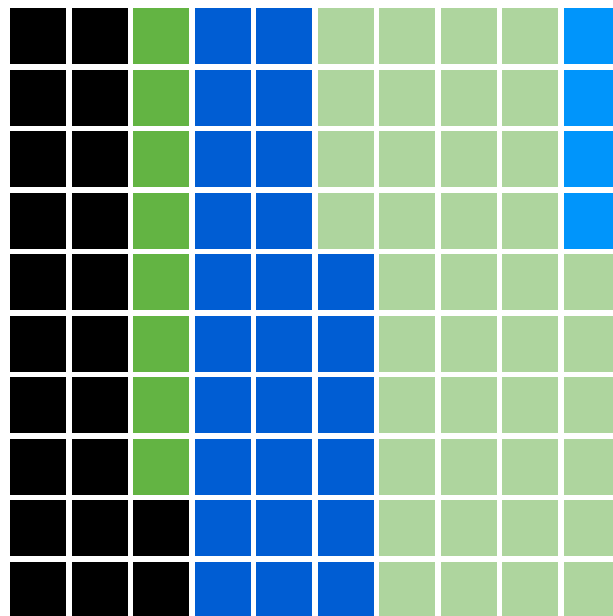
(a)



Prognosis

- Not known
- Poor (<50%)
- Intermediate (50-90%)
- Favourable (>90%)
- Normal

(b)



Prognosis

- Not known
- Poor (<50%)
- Intermediate (50-90%)
- Favourable (>90%)
- Normal

FIGURE 12 A waffle chart showing the MERIDIAN study cohort ($n = 829$). (a) By prognosis on ultrasonography when compared with (b) those participants recruited to project 2 ($n = 190$).

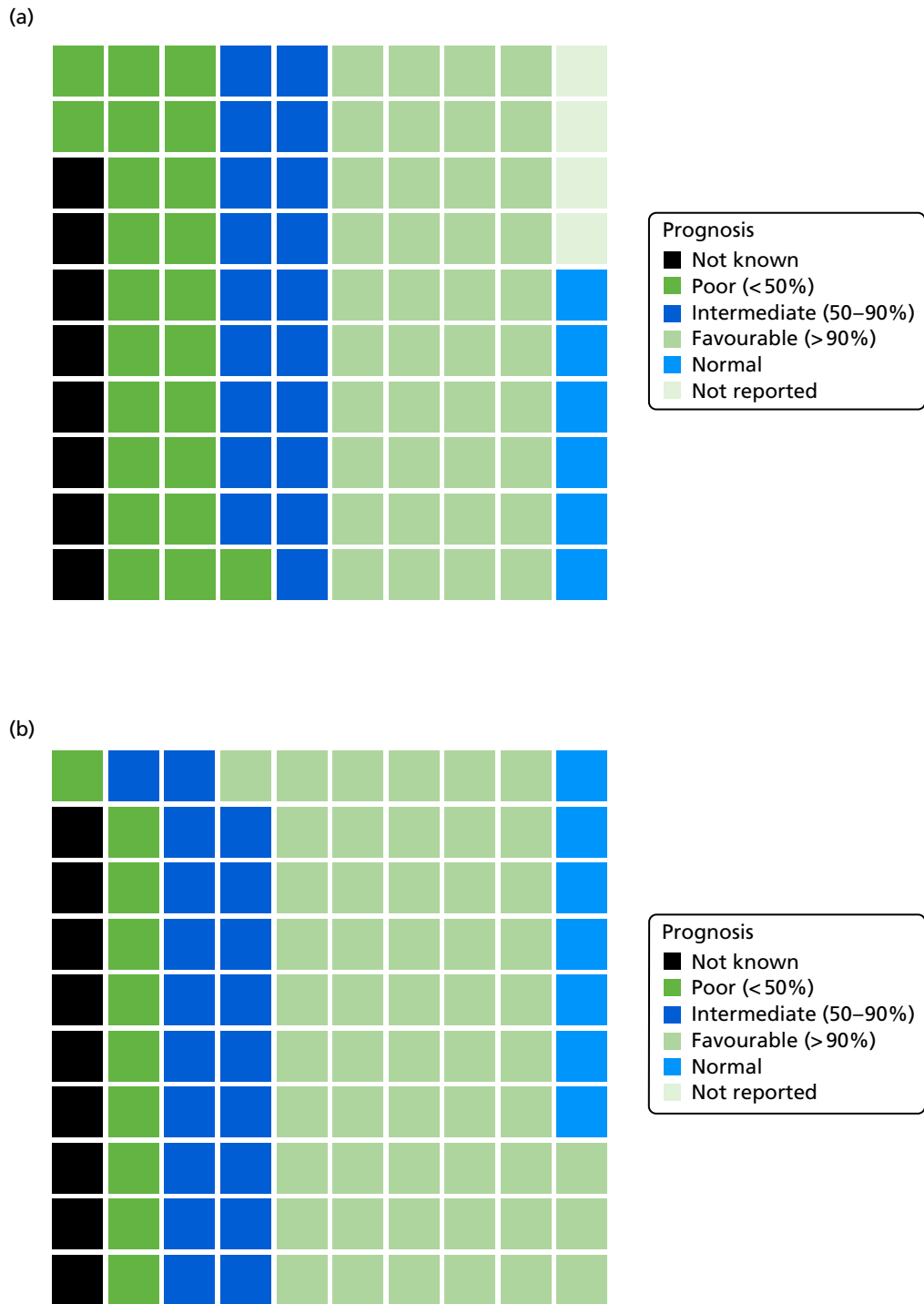


FIGURE 13 A waffle chart showing the MERIDIAN study cohort ($n = 829$). (a) By prognosis on iuMRI when compared with (b) those participants recruited to project 2 ($n = 190$).

Appendix 12 Box plots of Ages and Stages Questionnaire, Third Edition, and Strengths and Difficulties Questionnaire scores

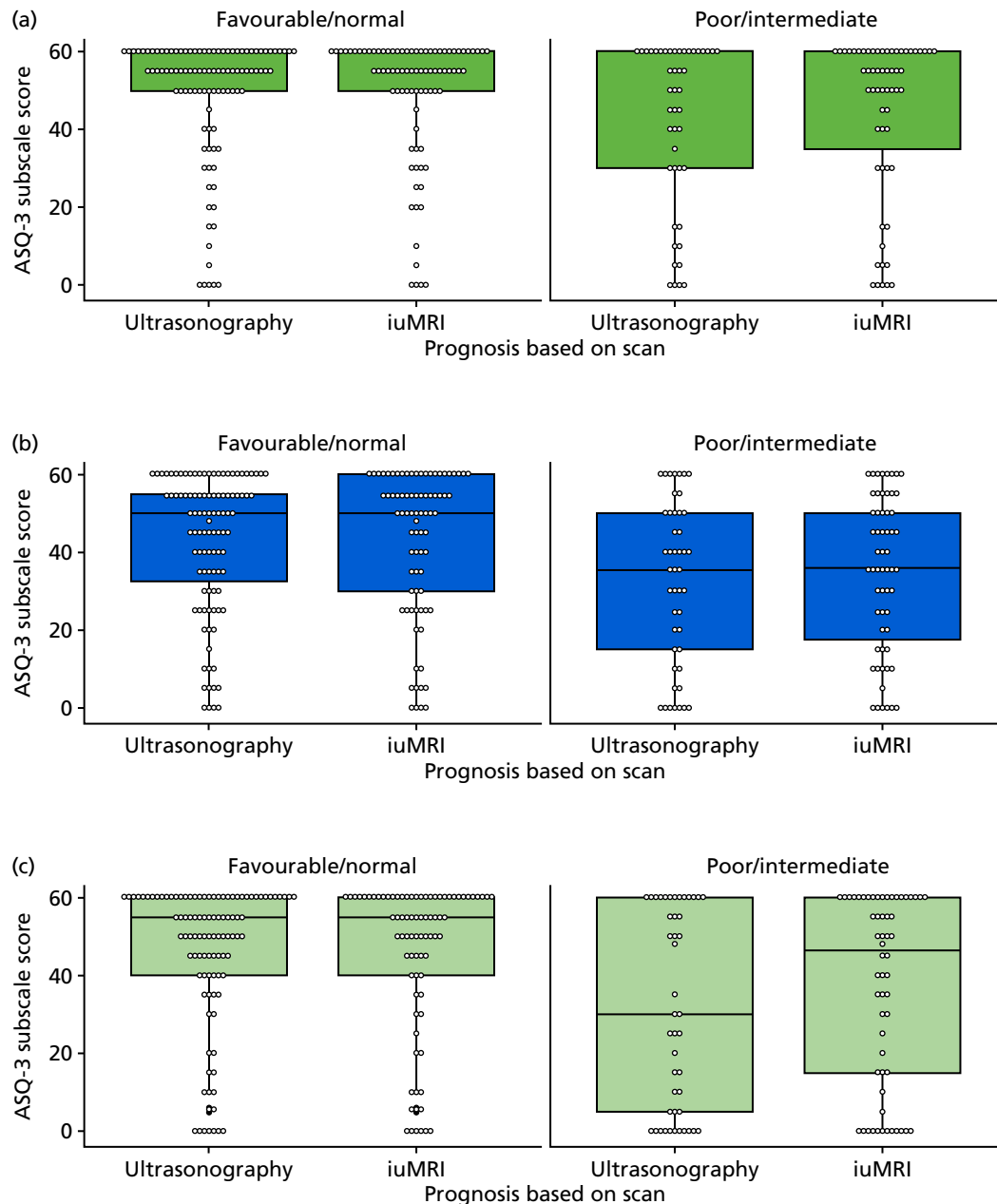


FIGURE 14 Box plots of ASQ-3 subscale scores by ultrasonography and iuMRI prognosis in the MERIDIAN study. (a) Communication, (b) fine motor, (c) gross motor, (d) problem-solving and (e) prosocial. (*continued*)

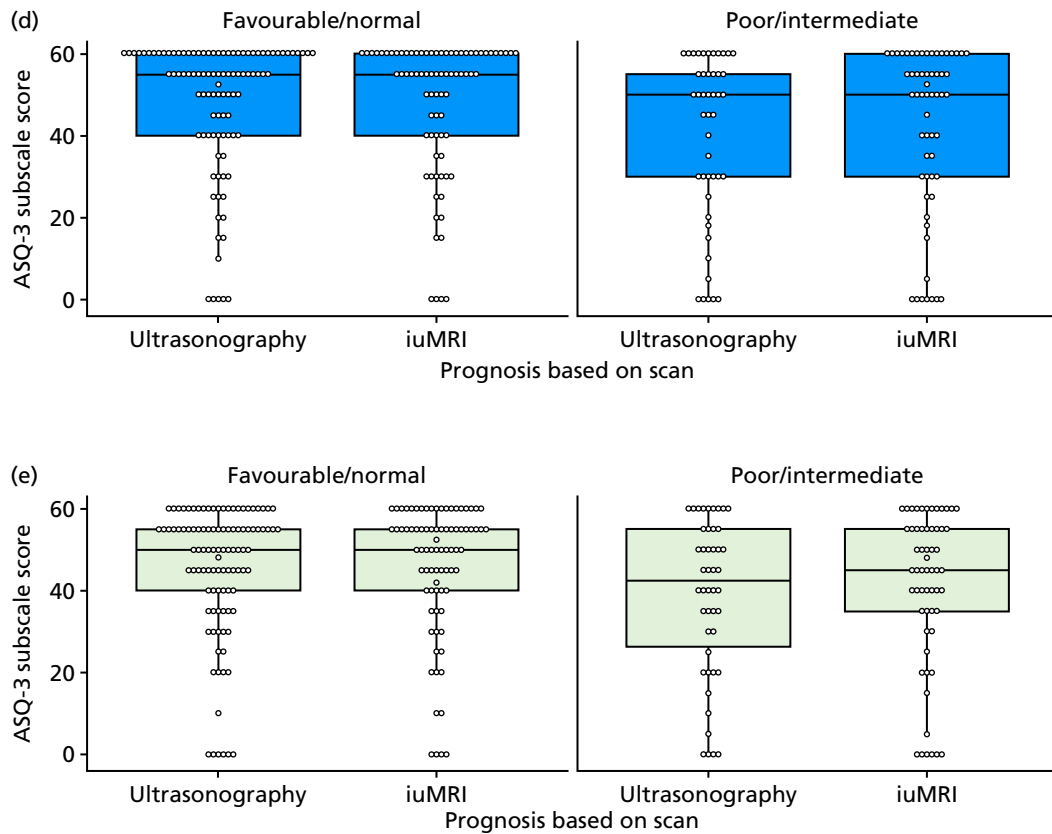


FIGURE 14 Box plots of ASQ-3 subscale scores by ultrasonography and iuMRI prognosis in the MERIDIAN study. (a) Communication, (b) fine motor, (c) gross motor, (d) problem-solving and (e) prosocial.

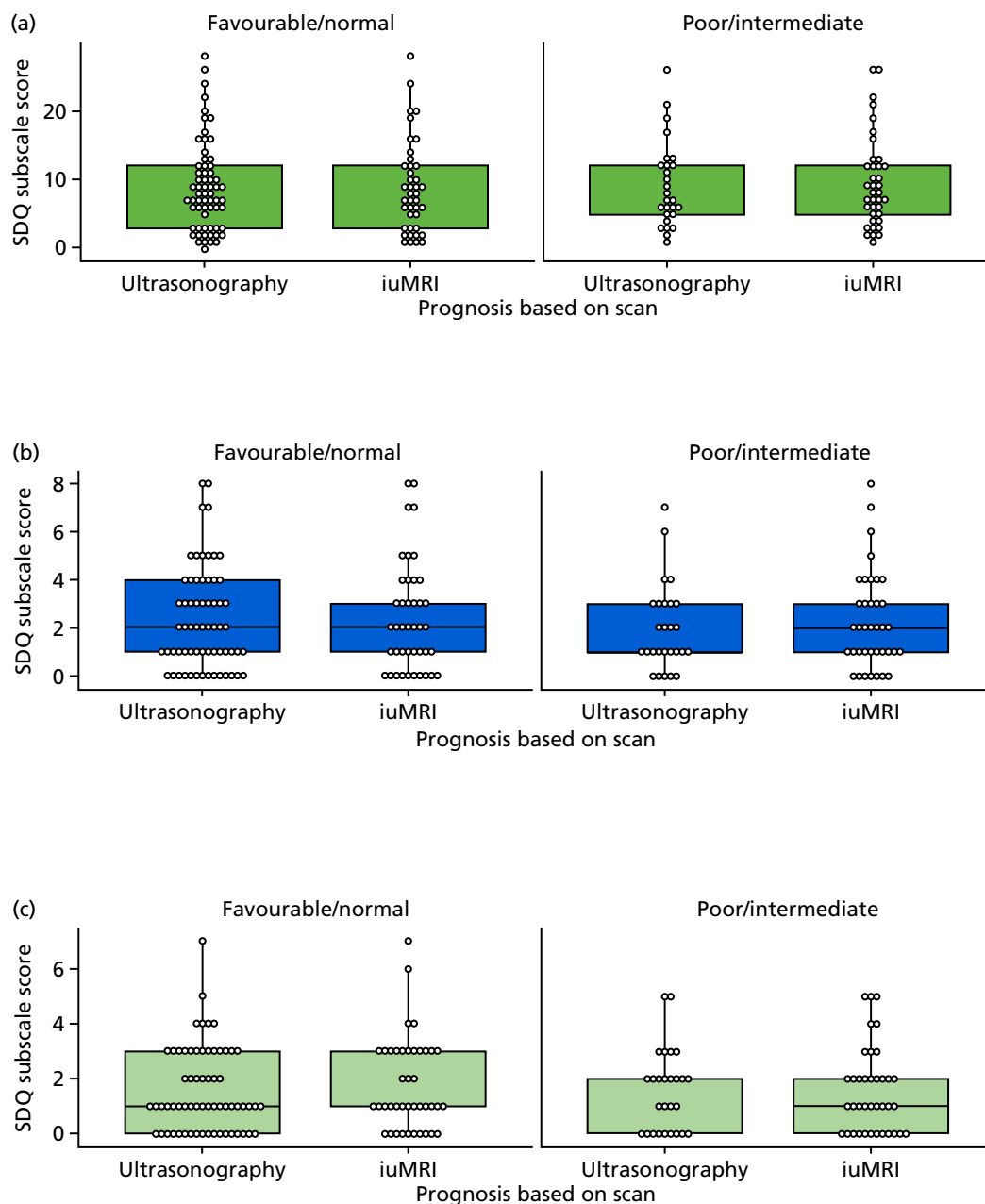


FIGURE 15 Box plots of SDQ subscale scores by ultrasonography and iuMRI prognosis in the MERIDIAN study. (a) Overall, (b) conduct, (c) emotional, (d) hyperactivity, (e) peer problems and (f) prosocial. (*continued*)

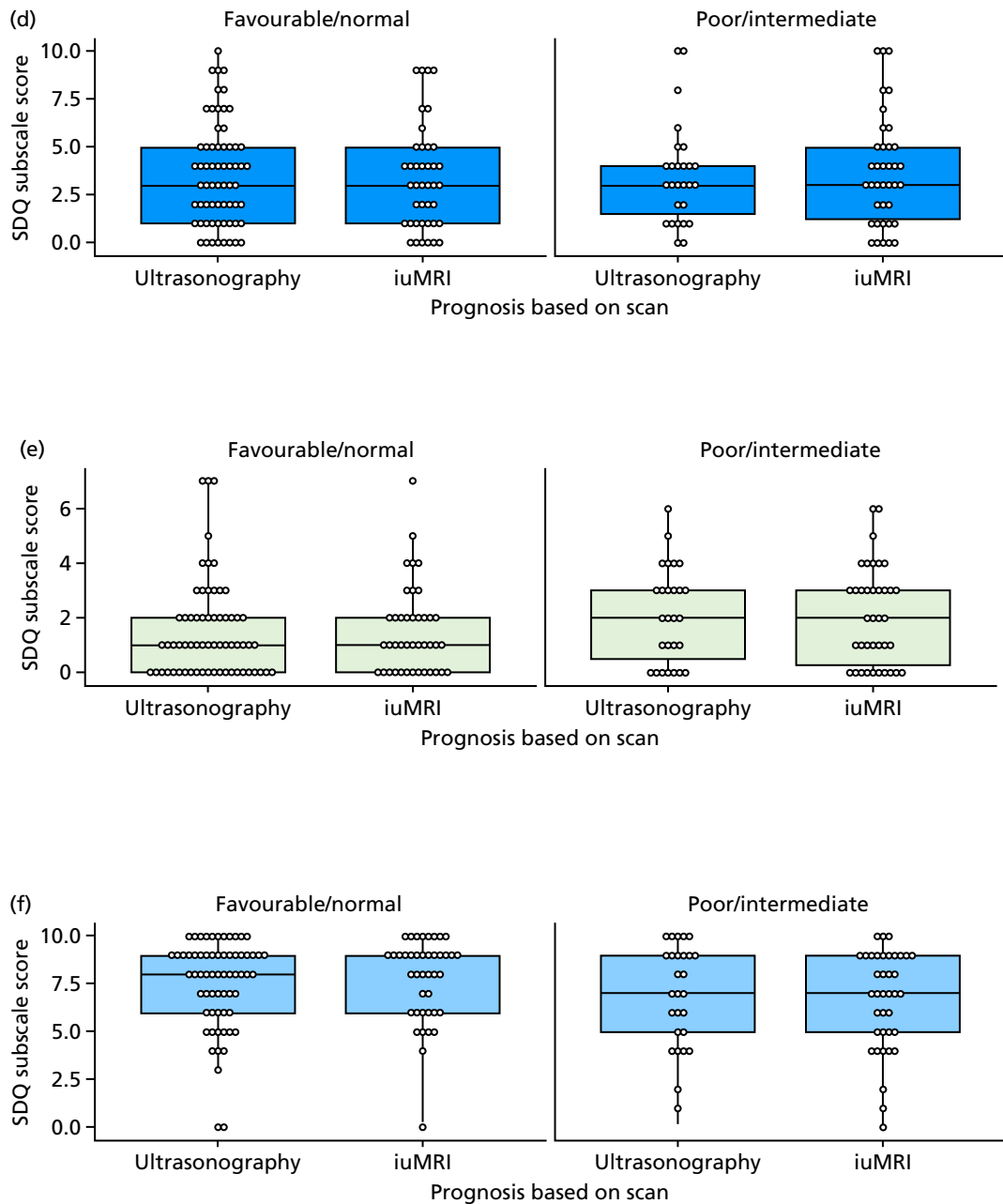


FIGURE 15 Box plots of SDQ subscale scores by ultrasonography and iuMRI prognosis in the MERIDIAN study. (a) Overall, (b) conduct, (c) emotional, (d) hyperactivity, (e) peer problems and (f) prosocial.

Appendix 13 Descriptive statistics tables for the developmental scores

TABLE 49 Summary of developmental outcome scores from the BSID assessment

BSID domain	BSID score
<i>n</i>	86
Cognitive raw score	
Mean (SD)	71.0 (12.6)
Median (IQR)	74 (67–80)
Min.–max.	23–87
Cognitive composite score	
Mean (SD)	10.6 (4.5)
Median (IQR)	10 (8–14)
Min.–max.	1–19
Language raw score	
Mean (SD)	102.3 (22.0)
Median (IQR)	100 (90–120)
Min.–max.	55–145
Language composite score	
Mean (SD)	30.8 (10.3)
Median (IQR)	32.5 (24–38)
Min.–max.	0–49
Motor raw score	
Mean (SD)	9.5 (4.0)
Median (IQR)	10 (7–12)
Min.–max.	0–19
Motor composite score	
Mean (SD)	32.1 (9.8)
Median (IQR)	34 (27–39)
Min.–max.	0–47
Max., maximum; min., minimum.	

TABLE 50 Summary of developmental outcome scores from the ASQ-3

ASQ-3 domain	ASQ-3 score
Communication	
<i>n</i>	176
Mean (SD)	45.3 (19.0)
Median (IQR)	55 (35–60)
Min.–max.	0–60
Gross motor	
<i>n</i>	175
Mean (SD)	40.8 (22.3)
Median (IQR)	50 (25–60)
Min.–max.	0–60
Fine motor	
<i>n</i>	175
Mean (SD)	38.3 (19.8)
Median (IQR)	45 (25–55)
Min.–max.	0–60
Problem-solving	
<i>n</i>	174
Mean (SD)	44.7 (18.9)
Median (IQR)	55 (35–60)
Min.–max.	0–60
Personal social	
<i>n</i>	175
Mean (SD)	42.6 (17.9)
Median (IQR)	50 (35–55)
Min.–max.	0–60
Max., maximum; min., minimum.	

TABLE 51 Summary of developmental outcome scores from the SDQ

SDQ domain	SDQ score
<i>n</i>	102
Overall	
Mean (SD)	9.1 (6.4)
Median (IQR)	7.5 (4–12)
Min.–max.	0–28
Emotional problems	
Mean (SD)	1.6 (1.5)
Median (IQR)	1 (0–3)
Min.–max.	0–7
Conduct problems	
Mean (SD)	2.2 (1.9)
Median (IQR)	2 (1–3)
Min.–max.	0–8
Hyperactivity	
Mean (SD)	3.6 (2.7)
Median (IQR)	3 (1–5)
Min.–max.	0–10
Peer problems	
Mean (SD)	1.8 (1.8)
Median (IQR)	1 (0–3)
Min.–max.	0–7
Prosocial	
Mean (SD)	7.3 (2.4)
Median (IQR)	8 (6–9)
Min.–max.	0–10
Max., maximum; min., minimum.	

Appendix 14 A breakdown of ultrasonography and in utero magnetic resonance imaging prognoses in 'abnormal', 'normal' and 'borderline' cases

Clinicians' developmental assessment and ultrasonography/iuMRI prognoses

TABLE 52 Ultrasonography prognosis category compared with iuMRI prognosis category in surviving infants who completed the developmental assessment and were assessed as 'abnormal' ($n = 53$)

Ultrasonography prognosis	iuMRI prognosis				Total
	Poor	Intermediate	Favourable/normal	Not known	
Poor	7	2	1	0	10
Intermediate	5	5	1	1	12
Favourable/normal	0	3	13	0	16
Not known	1	4	6	4	15
Total	13	14	21	5	53

TABLE 53 Ultrasonography prognosis category compared with iuMRI prognosis category in surviving infants who completed the developmental assessment and were assessed as 'normal' or 'borderline' ($n = 125$)

Ultrasonography prognosis	iuMRI prognosis				Total
	Poor	Intermediate	Favourable/normal	Not known	
Poor	2	1	2	0	5
Intermediate	1	11	17	4	33
Favourable/normal	1	3	56	3	63
Not known	1	6	13	4	24
Total	5	21	88	11	125

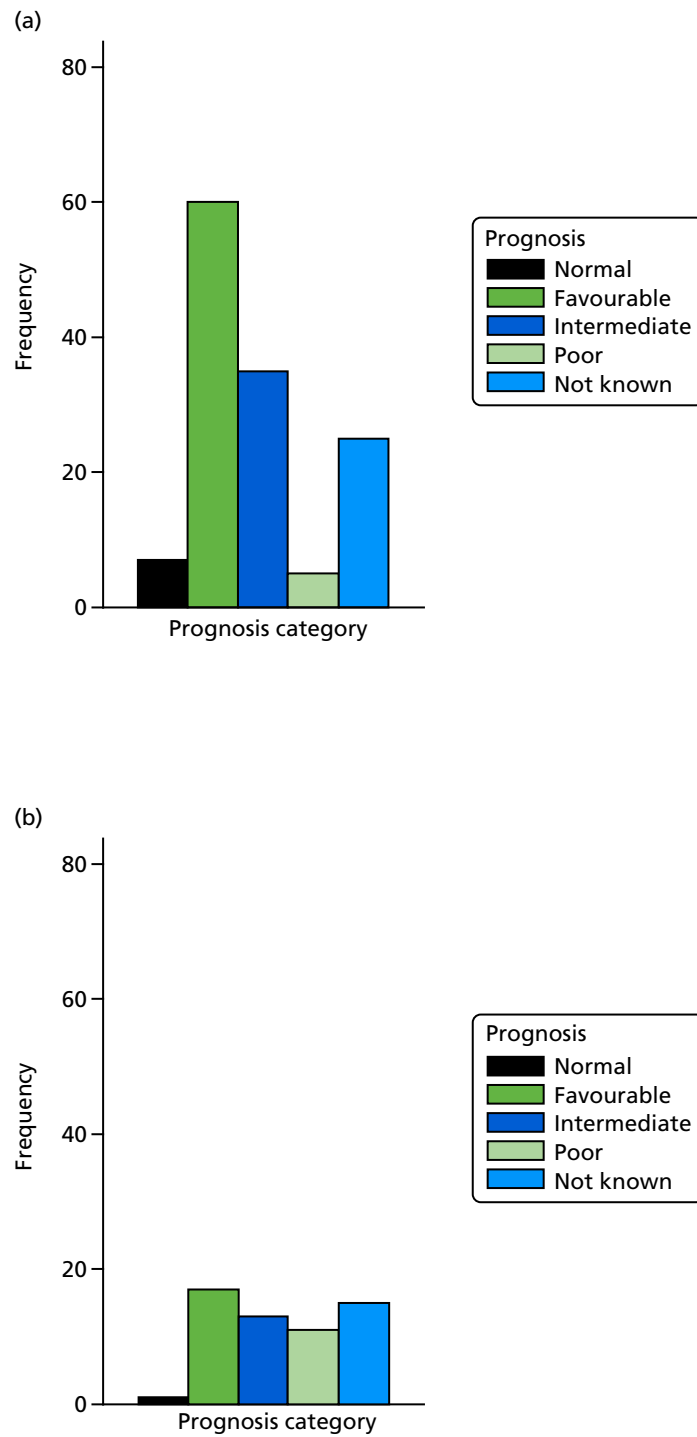


FIGURE 16 Frequency of each ultrasonography prognosis category by developmental outcome. (a) Normal/at risk, (b) abnormal and (c) death. (*continued*)

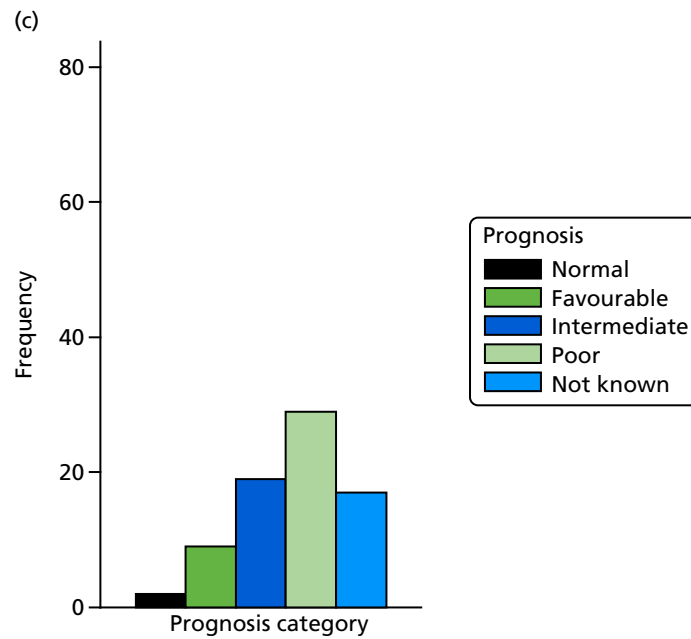


FIGURE 16 Frequency of each ultrasonography prognosis category by developmental outcome. (a) Normal/at risk, (b) abnormal and (c) death.

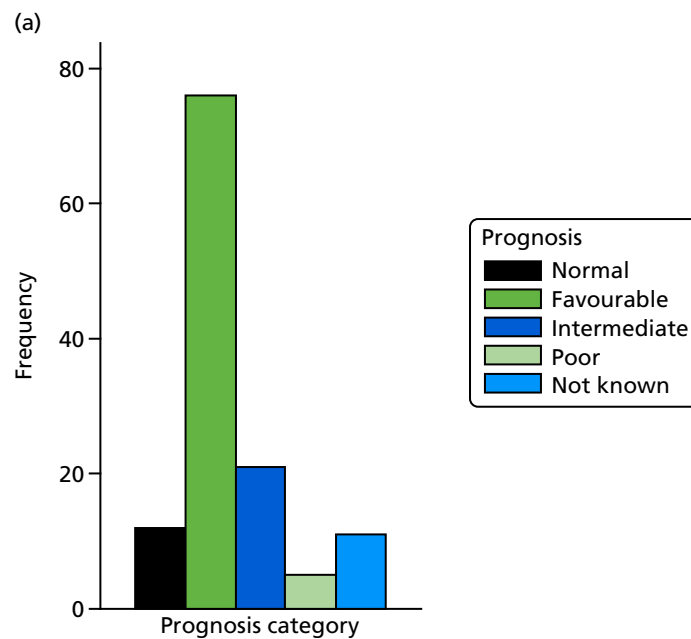


FIGURE 17 Frequency of each iuMRI prognosis category by developmental outcome. (a) Normal/at risk, (b) abnormal and (c) death. (*continued*)

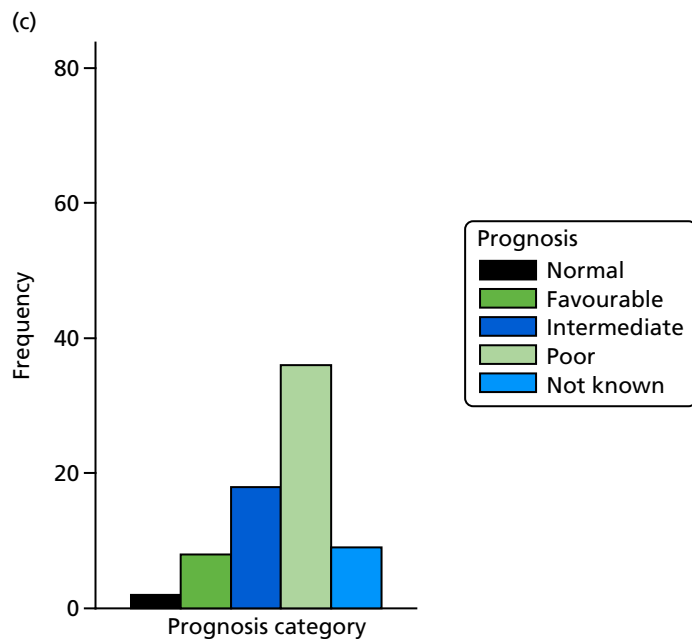
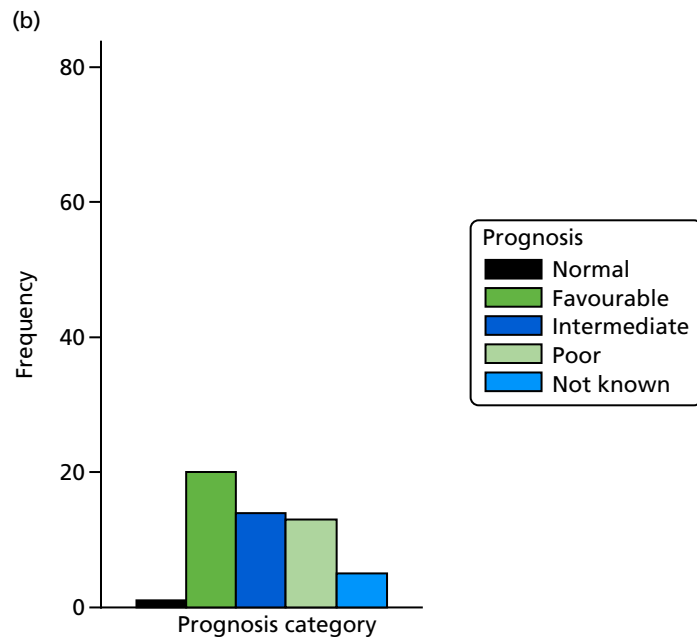


FIGURE 17 Frequency of each iuMRI prognosis category by developmental outcome. (a) Normal/at risk, (b) abnormal and (c) death.

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

*This report presents independent research funded by the National Institute for Health Research (NIHR).
The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the
Department of Health and Social Care*

Published by the NIHR Journals Library