NIHR | National Institute for Health Research

Journals Library

Check for updates

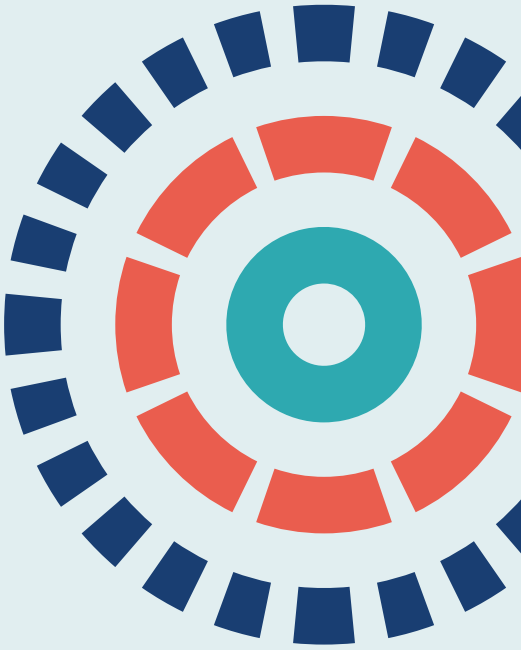
Health Technology Assessment

Volume 23 • Issue 66 • December 2019

ISSN 1366-5278

Whole-body MRI compared with standard pathways for staging metastatic disease in lung and colorectal cancer: the Streamline diagnostic accuracy studies

Stuart A Taylor, Susan Mallett, Anne Miles, Stephen Morris, Laura Quinn, Caroline S Clarke, Sandy Beare, John Bridgewater, Vicky Goh, Sam Janes, Dow-Mu Koh, Alison Morton, Neal Navani, Alfred Oliver, Anwar Padhani, Shonit Punwani, Andrea Rockall and Steve Halligan on behalf of the Streamline investigators



DOI 10.3310/hta23660

Whole-body MRI compared with standard pathways for staging metastatic disease in lung and colorectal cancer: the Streamline diagnostic accuracy studies

Stuart A Taylor[1]* Susan Mallett[2] Anne Miles[3] Stephen Morris[4] Laura Quinn[2] Caroline S Clarke[5] Sandy Beare[6] John Bridgewater[7] Vicky Goh[8] Sam Janes[9] Dow-Mu Koh[10] Alison Morton[11] Neal Navani[9] Alfred Oliver[11] Anwar Padhani[12] Shonit Punwani[1] Andrea Rockall[13] and Steve Halligan[1] on behalf of the Streamline investigators

[1]Centre for Medical Imaging, University College London, London, UK
[2]Institute of Applied Health Research, NIHR Birmingham Biomedical Research Centre, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK
[3]Department of Psychological Sciences, Birkbeck, University of London, London, UK
[4]Applied Health Research, University College London, London, UK
[5]Research Department of Primary Care and Population Health, and Priment Clinical Trials Unit, University College London, London, UK
[6]Cancer Research UK & UCL Cancer Trials Centre, University College London, London, UK
[7]UCL Cancer Institute, University College London, London, UK
[8]Department of Cancer Imaging, School of Biomedical Engineering and Imaging Sciences, King's College London, London, UK
[9]Lungs for Living Research Centre, UCL Respiratory, University College London, London, UK
[10]Department of Radiology, The Royal Marsden Hospital, Sutton, UK
[11]c/o Centre for Medical Imaging, University College London, London, UK
[12]Mount Vernon Centre for Cancer Treatment, Mount Vernon Hospital, Northwood, UK
[13]Imaging Department, Imperial College Healthcare NHS Trust, London, UK

*Corresponding author

Declared competing interests of authors: Stuart A Taylor reports personal fees from Robarts Clinical Trials Inc. (London, ON, Canada) outside the submitted work. Andrea Rockall reports personal fees from Guerbet (Paris, France) outside the submitted work. Vicky Goh and Anwar Padhani report grants from Siemens AG (Erlangen, Germany) outside the submitted work. Steve Halligan reports non-financial support from iCad Inc. (Nashua, NH, USA) outside the submitted work and was a member of the Health Technology Assessment (HTA) commissioning board (2008–14). Stephen Morris declares past membership of the National Institute for Health Research (NIHR) HTA Clinical Evaluation and Trials Board (2007–9), HTA Commissioning Board (2009–13), Public Health Research Board (2011–17), Health Services and Delivery Research (HSDR) Commissioning Board (2014–16) and HSDR Board (2014–18), and current membership of the NIHR Programme Grants for Applied Research expert subpanel (2015–present) and HSDR Evidence Synthesis Sub Board (2016–present).

Disclaimer: This report contains transcripts of interviews conducted in the course of the research and contains language that may offend some readers.

Published December 2019
DOI: 10.3310/hta23660

This report should be referenced as follows:

Taylor SA, Mallett S, Miles A, Morris S, Quinn L, Clarke CS, *et al*. Whole-body MRI compared with standard pathways for staging metastatic disease in lung and colorectal cancer: the Streamline diagnostic accuracy studies. *Health Technol Assess* 2019;**23**(66).

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

Health Technology Assessment

HTA/HTA TAR

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

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

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.

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 10/68/01. The contractual start date was in October 2012. The draft report began editorial review in December 2018 and was accepted for publication in May 2019. 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 Taylor *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 Senior Clinical Researcher, 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

Whole-body MRI compared with standard pathways for staging metastatic disease in lung and colorectal cancer: the Streamline diagnostic accuracy studies

Stuart A Taylor,[1*] Susan Mallett,[2] Anne Miles,[3] Stephen Morris,[4] Laura Quinn,[2] Caroline S Clarke,[5] Sandy Beare,[6] John Bridgewater,[7] Vicky Goh,[8] Sam Janes,[9] Dow-Mu Koh,[10] Alison Morton,[11] Neal Navani,[9] Alfred Oliver,[11] Anwar Padhani,[12] Shonit Punwani,[1] Andrea Rockall[13] and Steve Halligan[1] on behalf of the Streamline investigators

[1]Centre for Medical Imaging, University College London, London, UK
[2]Institute of Applied Health Research, NIHR Birmingham Biomedical Research Centre, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK
[3]Department of Psychological Sciences, Birkbeck, University of London, London, UK
[4]Applied Health Research, University College London, London, UK
[5]Research Department of Primary Care and Population Health, and Priment Clinical Trials Unit, University College London, London, UK
[6]Cancer Research UK & UCL Cancer Trials Centre, University College London, London, UK
[7]UCL Cancer Institute, University College London, London, UK
[8]Department of Cancer Imaging, School of Biomedical Engineering and Imaging Sciences, King's College London, London, UK
[9]Lungs for Living Research Centre, UCL Respiratory, University College London, London, UK
[10]Department of Radiology, The Royal Marsden Hospital, Sutton, UK
[11]c/o Centre for Medical Imaging, University College London, London, UK
[12]Mount Vernon Centre for Cancer Treatment, Mount Vernon Hospital, Northwood, UK
[13]Imaging Department, Imperial College Healthcare NHS Trust, London, UK

*Corresponding author stuart.taylor@ucl.ac.uk

Background: Whole-body magnetic resonance imaging is advocated as an alternative to standard pathways for staging cancer.

Objectives: The objectives were to compare diagnostic accuracy, efficiency, patient acceptability, observer variability and cost-effectiveness of whole-body magnetic resonance imaging and standard pathways in staging newly diagnosed non-small-cell lung cancer (Streamline L) and colorectal cancer (Streamline C).

Design: The design was a prospective multicentre cohort study.

Setting: The setting was 16 NHS hospitals.

Participants: Consecutive patients aged ≥ 18 years with histologically proven or suspected colorectal (Streamline C) or non-small-cell lung cancer (Streamline L).

Interventions: Whole-body magnetic resonance imaging. Standard staging investigations (e.g. computed tomography and positron emission tomography–computed tomography).

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Reference standard: Consensus panel decision using 12-month follow-up data.

Main outcome measures: The primary outcome was per-patient sensitivity difference between whole-body magnetic resonance imaging and standard staging pathways for metastasis. Secondary outcomes included differences in specificity, the nature of the first major treatment decision, time and number of tests to complete staging, patient experience and cost-effectiveness.

Results: Streamline C – 299 participants were included. Per-patient sensitivity for metastatic disease was 67% (95% confidence interval 56% to 78%) and 63% (95% confidence interval 51% to 74%) for whole-body magnetic resonance imaging and standard pathways, respectively, a difference in sensitivity of 4% (95% confidence interval –5% to 13%; $p = 0.51$). Specificity was 95% (95% confidence interval 92% to 97%) and 93% (95% confidence interval 90% to 96%) respectively, a difference of 2% (95% confidence interval –2% to 6%). Pathway treatment decisions agreed with the multidisciplinary team treatment decision in 96% and 95% of cases, respectively, a difference of 1% (95% confidence interval –2% to 4%). Time for staging was 8 days (95% confidence interval 6 to 9 days) and 13 days (95% confidence interval 11 to 15 days) for whole-body magnetic resonance imaging and standard pathways, respectively, a difference of 5 days (95% confidence interval 3 to 7 days). The whole-body magnetic resonance imaging pathway was cheaper than the standard staging pathway: £216 (95% confidence interval £211 to £221) versus £285 (95% confidence interval £260 to £310). Streamline L – 187 participants were included. Per-patient sensitivity for metastatic disease was 50% (95% confidence interval 37% to 63%) and 54% (95% confidence interval 41% to 67%) for whole-body magnetic resonance imaging and standard pathways, respectively, a difference in sensitivity of 4% (95% confidence interval –7% to 15%; $p = 0.73$). Specificity was 93% (95% confidence interval 88% to 96%) and 95% (95% confidence interval 91% to 98%), respectively, a difference of 2% (95% confidence interval –2% to 7%). Pathway treatment decisions agreed with the multidisciplinary team treatment decision in 98% and 99% of cases, respectively, a difference of 1% (95% confidence interval –2% to 4%). Time for staging was 13 days (95% confidence interval 12 to 14 days) and 19 days (95% confidence interval 17 to 21 days) for whole-body magnetic resonance imaging and standard pathways, respectively, a difference of 6 days (95% confidence interval 4 to 8 days). The whole-body magnetic resonance imaging pathway was cheaper than the standard staging pathway: £317 (95% confidence interval £273 to £361) versus £620 (95% confidence interval £574 to £666). Participants generally found whole-body magnetic resonance imaging more burdensome than standard imaging but most participants preferred the whole-body magnetic resonance imaging staging pathway if it reduced time to staging and/or number of tests.

Limitations: Whole-body magnetic resonance imaging was interpreted by practitioners blinded to other clinical data, which may not fully reflect how it is used in clinical practice.

Conclusions: In colorectal and non-small-cell lung cancer, the whole-body magnetic resonance imaging staging pathway has similar accuracy to standard staging pathways, is generally preferred by patients, improves staging efficiency and has lower staging costs. Future work should address the utility of whole-body magnetic resonance imaging for treatment response assessment.

Trial registration: Current Controlled Trials ISRCTN43958015 and ISRCTN50436483.

Funding: This project was funded by the NIHR Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 23, No. 66. See the NIHR Journals Library website for further project information.

Contents

List of tables xvii

List of figures xxv

List of boxes xxvii

List of supplementary material xxix

List of abbreviations xxxi

Plain English summary xxxiii

Scientific summary xxxv

Chapter 1 Introduction 1
Background 1
Existing literature on the diagnostic accuracy of WB-MRI in cancer staging 2
Objectives of the Streamline studies 7

Chapter 2 Methods 9
Study design 9
Patient and public involvement 10
Recruitment sites 11
Inclusion criteria: Streamline C 11
Exclusion criteria: Streamline C 12
Inclusion criteria: Streamline L 13
Exclusion criteria: Streamline L 13
Test methods 13
WB-MRI 13
Standard investigations 15
Revelation of WB-MRI findings and multidisciplinary team treatment decisions 15
Early release of WB-MRI findings 16
Recruitment 16
Data collection and follow-up 17
Reference standard 17
Outcomes 19
Sample size 19
Primary outcome 19
Analysis 20

Chapter 3 Results: Streamline C 23
Participants 23
Equivocal results 25
Streamline C: staging pathway tests 25
Pathway results and outcomes 26
Primary outcome 26
Secondary outcomes 28
Local staging 28

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Impact of staging pathways on primary treatment decision 31
Time to complete staging 32
Number of tests in staging pathways 33
WB-MRI as a stand-alone investigation 33
WB-MRI as an additional investigation 34
Perceptual errors 34
WB-MRI quality assurance: Streamline C 35

Chapter 4 Results: Streamline L **37**
Participants 37
Equivocal results 39
Streamline L: staging pathway tests 40
Pathway results and outcomes 41
Primary outcome 41
Secondary outcomes 42
Local staging 44
Impact of staging pathways on primary treatment decision 45
Time to complete staging 46
Number of tests in staging pathways 47
WB-MRI as a stand-alone investigation 47
WB-MRI as an additional investigation 48
Perceptual errors 48
WB-MRI quality assurance: Streamline L 48

Chapter 5 Economic analysis of WB-MRI staging pathway versus standard staging pathways in non-small-cell lung cancer and colorectal cancer **49**
Introduction 49
Overview of economic analysis 49
Treatment decision-making processes in Streamline 49
Effect of decision concordance on the economic analysis 50
Available treatment decisions in Streamline C and L 51
Concordance in treatment decisions in Streamline C and L 51
Cost of the treatment decision pathways 51
Results 53
Streamline C 53
Streamline L 54

Chapter 6 WB-MRI and standard staging pathways: comparison of participant experience, perceived burden and preferences **55**
Introduction 55
Participant experience study 1: participant interview study to examine WB-MRI burden and acceptability in comparison with standard staging scans 56
Methods 56
Participants 56
Interviews 56
Data coding and analysis 56
Results 56
Participant experience of WB-MRI in comparison to standard staging investigations 57
Challenges of WB-MRI 57
Explaining differences in people's experiences 61
Willingness to have another WB-MRI scan 62

Participant experience study 2: quantitative assessment of perceived WB-MRI scan burden and acceptability – comparison with CT/PET-CT 63
Methods 63
Questionnaires 63
Power calculation 63
Statistical analysis 64
Results 64
Post-scan participant recovery and scan acceptability 65
Scan burden 66
Predictors of WB-MRI scan burden 67
Participant experience study 3: participant preferences for whole body MRI or conventional staging pathways in lung and colorectal cancer – a discrete choice experiment 68
Discrete choice experiment questionnaire attributes and levels 68
Discrete choice experiment questionnaire design 69
Sample size calculation 70
Analysis 70
Study 3 results 71
Importance of attributes 71
Traders versus non-traders 72
Marginal rates of substitution 72

Chapter 7 Influence of WB-MRI sequences on diagnostic accuracy **77**
Introduction 77
Methods 77
Analysis 78
Results 78
Streamline C and Streamline L 78

Chapter 8 Interobserver agreement in WB-MRI interpretation **81**
Introduction 81
Methods 81
Data set selection 81
Analysis 82
Results 83

Chapter 9 Discussion **85**
Diagnostic accuracy 85
First major treatment decision 89
Staging pathway efficiency 90
Cost-effectiveness 91
Main trial strengths and limitations 91
Participant experience 92
Influence of WB-MRI sequences on diagnostic accuracy 95
Interobserver agreement in WB-MRI interpretation 96
Overall conclusions 96
Implications for practice 97
Recommendations for future research 97

Acknowledgements **99**

References **105**

Appendix 1 WB-MRI acquisition: minimum data set **111**

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 2 Quality assurance grading score definitions 113

Appendix 3 Flow chart of multidisciplinary team process recording the first major treatment decision based on standard and WB-MRI pathways 115

Appendix 4 Summary of primary and secondary outcomes 117

Appendix 5 Streamline C: reasons for non-inclusion of screened patients 119

Appendix 6 Streamline C: reasons for withdrawal of recruited participants 121

Appendix 7 Streamline C: primary tumour status at the time of consensus review 123

Appendix 8 Streamline C: final T and N stage based on consensus reference standard 125

Appendix 9 Streamline C: local tumour spread based on consensus reference 127

Appendix 10 Streamline C: detailed organ sites of metastatic disease 129

Appendix 11 Streamline C: number of metastatic deposits per organ 131

Appendix 12 Streamline C: maximum size of metastatic deposits according to organ 133

Appendix 13 Streamline C: equivocal results for metastatic disease 135

Appendix 14 Streamline C: short-interval follow-up for equivocal findings 137

Appendix 15 Streamline C: per-organ sensitivity and specificity for metastatic disease – standard staging pathways versus WB-MRI staging pathway against the consensus reference standard 139

Appendix 16 Streamline C: per-participant agreement for tumour T stage – standard staging pathways versus WB-MRI staging pathway against the consensus reference standard in participants with histological proof 141

Appendix 17 Streamline C: per-participants agreement for tumour N stage – standard staging pathways versus WB-MRI staging pathway against the consensus reference standard in participants with histological proof 143

Appendix 18 Streamline C: breakdown of the primary treatment decision according to the standard staging pathway compared with the retrospective consensus panel optimal treatment decision (non-rectal colon cancer) 145

Appendix 19 Streamline C: breakdown of the primary treatment decision according to the WB-MRI staging pathway compared with the retrospective consensus panel optimal treatment decision (non-rectal colon cancer) 147

Appendix 20 Streamline C: breakdown of the primary treatment decision according to the standard staging pathway compared with the retrospective consensus panel optimal treatment decision (rectal cancer) 149

Appendix 21 Streamline C: breakdown of the primary treatment decision according to the WB-MRI staging pathway compared with the retrospective consensus panel optimal treatment decision (rectal cancer) 151

Appendix 22 Streamline C: time to complete staging according to staging pathway: interquartile range 153

Appendix 23 Streamline C: per-participant sensitivity and specificity for metastatic disease – standard staging pathway versus WB-MRI alone against the consensus reference standard 155

Appendix 24 Streamline C: per-organ sensitivity and specificity for metastatic disease – standard staging pathways versus WB-MRI alone against the consensus reference standard 157

Appendix 25 Streamline C: per-participant agreement for tumour T stage – standard staging pathways versus WB-MRI alone against the consensus reference standard – all participants 159

Appendix 26 Streamline C: per-participant agreement for tumour T stage – standard staging pathways versus WB-MRI alone against the consensus reference standard in participants with histological proof 161

Appendix 27 Streamline C: per-participant agreement for tumour N stage – standard staging pathways versus WB-MRI alone against the consensus reference standard – all participants 163

Appendix 28 Streamline C: per-participant agreement for tumour N stage – standard staging pathways versus WB-MRI alone against the consensus reference standard in participants with histological proof 165

Appendix 29 Streamline C: per-participant sensitivity and specificity for metastatic disease: chest, abdomen and pelvis CT versus WB-MRI alone against the consensus reference standard 167

Appendix 30 Streamline C: per-participant sensitivity and specificity for metastatic disease – standard staging pathways versus standard staging pathways plus WB-MRI against the consensus reference standard 169

Appendix 31 Streamline C: perceptual errors for metastatic disease according to staging pathway 171

Appendix 32 Streamline C: perceptual errors for metastatic disease according to staging pathway and organ disease site 173

Appendix 33 WB-MRI image quality scoring 175

Appendix 34 Detailed image quality score categories for scans rated as overall suboptimal category 177

Appendix 35 Detailed image quality score categories for scans rated as overall degraded category 179

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 36 Streamline L: reasons for non-inclusion of screened patients 181

Appendix 37 Streamline L: reasons for withdrawal of recruited participants 183

Appendix 38 Streamline L: primary tumour status at the time of consensus review 185

Appendix 39 Streamline L: final T and N stage based on consensus reference standard 187

Appendix 40 Streamline L: final nodal disease sites based on the consensus reference standard 189

Appendix 41 Streamline L: detailed organs sites of metastatic disease 191

Appendix 42 Streamline L: number of metastatic deposits per organ 193

Appendix 43 Streamline L: maximum size of metastatic deposits according to organ 195

Appendix 44 Streamline L: equivocal results for metastatic disease 197

Appendix 45 Streamline L: short-interval follow-up for equivocal findings 199

Appendix 46 Streamline L: per-organ sensitivity and specificity for metastatic disease – standard staging pathways versus WB-MRI staging pathway against the consensus reference standard 201

Appendix 47 Streamline L: per-participant agreement for tumour T stage – standard staging pathways versus WB-MRI staging pathway against the consensus reference standard in participants with histological proof 203

Appendix 48 Streamline L: per-participant agreement for tumour N stage – standard staging pathways versus WB-MRI staging pathway against the consensus reference standard in participants with histological proof 205

Appendix 49 Streamline L: nodal site agreement – standard staging pathways versus WB-MRI staging pathway against the consensus reference standard in participants with histological proof 207

Appendix 50 Streamline L: breakdown of the primary treatment decision according to the standard staging pathway compared with the retrospective consensus panel optimal treatment decision 209

Appendix 51 Streamline L: breakdown of the primary treatment decision according to the WB-MRI staging pathway compared with the retrospective consensus panel optimal treatment decision 211

Appendix 52 Streamline L: time to complete staging according to staging pathway 213

Appendix 53 Streamline L: per-participant sensitivity and specificity for metastatic disease-standard staging pathway versus WB-MRI alone against the consensus reference standard 215

Appendix 54 Streamline L: per-organ sensitivity and specificity for metastatic disease – standard staging pathways versus WB-MRI alone against the consensus reference standard 217

Appendix 55 Streamline L: per-participant agreement for tumour T stage – standard staging pathways versus WB-MRI alone against the consensus reference standard 219

Appendix 56 Streamline L: per-participant agreement for tumour T stage – standard staging pathways versus WB-MRI alone against the consensus reference standard in participants with histological proof 221

Appendix 57 Streamline L: per-participant agreement for tumour N stage – standard staging pathways versus WB-MRI alone against the consensus reference standard – all participants 223

Appendix 58 Streamline L: per-participant agreement for tumour N stage – standard staging pathways versus WB-MRI alone against the consensus reference standard in participants with histological proof 225

Appendix 59 Streamline L: per-participant agreement for metastatic disease – PET-CT versus WB-MRI alone against the consensus reference standard 227

Appendix 60 Streamline L: per-participant sensitivity and specificity for metastatic disease-standard staging pathways versus standard staging pathways plus WB-MRI against the consensus reference standard 229

Appendix 61 Streamline L: perceptual errors for metastatic disease according to staging pathway 231

Appendix 62 Streamline L: perceptual errors for metastatic disease according to staging pathway and organ disease site 233

Appendix 63 Streamline C: concordance between treatment decisions based on standard and WB-MRI staging pathway (d1 and d2) 235

Appendix 64 Streamline L: concordance between treatment decisions based on standard and WB-MRI staging pathway (d1 and d2) 237

Appendix 65 Streamline C: discordance between treatment decisions based on standard staging (d1) and the consensus reference panel optimal decision (d4) 239

Appendix 66 Streamline C: discordance between treatment decisions based on WB-MRI staging pathway (d1) and the consensus reference panel optimal decision (d4) 241

Appendix 67 Streamline L: discordance between treatment decisions based on standard staging (d1) and the consensus reference panel optimal decision (d4) 243

Appendix 68 Streamline L: discordance between treatment decisions based on WB-MRI staging pathway (d1) and the consensus reference panel optimal decision (d4) 245

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 69 Interview topic guide 247

Appendix 70 Demographic characteristics of interview participants by trial 249

Appendix 71 Comparison of demographic characteristics of participants sent both questionnaires, who were or were not included in the final analysis (study 2) 251

Appendix 72 Demographic and characteristics of participants who completed the post-staging questionnaire 253

Appendix 73 Comparative experience of WB-MRI vs. CT/PET-CT: scan information, communication and facilities 255

Appendix 74 Demographic and psychological characteristics of the cohort completing the discrete choice experiment 257

Appendix 75 Results of conditional logit regression analysis: Streamline L participants stratified by home ownership 259

Appendix 76 Results of conditional logit regression analysis: Streamline L participants stratified by education 261

Appendix 77 Results of conditional logit regression analysis: Streamline L participants stratified by self-rated health 263

Appendix 78 Results of conditional logit regression analysis: Streamline L participants stratified by test preference 265

Appendix 79 Participants unwilling to trade attributes off against one another 267

Appendix 80 WB-MRI interobserver variability (not considering agreement with the consensus reference standard): Streamline C and Streamline L 269

List of tables

TABLE 1 Summary of recent meta-analyses and systematic reviews reporting the diagnostic accuracy of WB-MRI in cancer staging 3

TABLE 2 Recruitment sites and imaging hubs 12

TABLE 3 Streamline C: grouping of treatment decisions for analysis 20

TABLE 4 Streamline L: grouping of treatment decisions for analysis 21

TABLE 5 Demographics of the final study cohort 24

TABLE 6 Streamline C: organ sites of metastatic disease 25

TABLE 7 Streamline C: constituent tests of the standard staging pathway 26

TABLE 8 Streamline C: MDT-stated additional staging tests generated by WB-MRI 27

TABLE 9 Streamline C: per-participant sensitivity and specificity for metastatic disease – standard staging pathway vs. WB-MRI staging pathway against the consensus reference standard 27

TABLE 10 Streamline C: per-participant sensitivity and specificity for metastatic disease according to the size of metastasis – standard staging pathway vs. WB-MRI staging pathway against the consensus reference standard 30

TABLE 11 Streamline C: per-participant agreement for tumour T stage – standard staging pathway vs. WB-MRI staging pathway against the consensus reference standard 30

TABLE 12 Streamline C: per-participant agreement for N stage – standard staging pathway vs. WB-MRI staging pathway against the consensus reference standard 31

TABLE 13 Streamline C: primary treatment decision according to the WB-MRI staging pathway and standard staging pathway compared with the final MDT treatment decision 31

TABLE 14 Streamline C: primary treatment decision according to the WB-MRI staging pathway and standard staging pathway compared with the retrospective optimal consensus panel treatment decision 32

TABLE 15 Streamline C: time to complete staging according to treatment pathway 33

TABLE 16 Streamline C: total number of imaging/endoscopic tests before the first major treatment decision according to staging pathway 33

TABLE 17 Demographics of the final study: Streamline L cohort 38

TABLE 18 Streamline L: organ sites of metastatic disease 39

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 19 Streamline L: constituent tests of the standard staging pathway 40

TABLE 20 Streamline L: MDT-stated additional staging tests generated by the WB-MRI 41

TABLE 21 Streamline L: per-participant sensitivity and specificity for metastatic disease – standard staging pathway vs. WB-MRI staging pathway against the consensus reference standard 42

TABLE 22 Streamline L: per-participant sensitivity and specificity for metastatic disease according to the size of metastasis – standard staging pathway vs. WB-MRI staging pathway against the consensus reference standard 44

TABLE 23 Streamline L: per-participant agreement for tumour T stage – standard staging pathway vs. WB-MRI staging pathway against the consensus reference standard 44

TABLE 24 Streamline L: per-participant agreement for N stage – standard staging pathway vs. WB-MRI staging pathway against the consensus reference standard 45

TABLE 25 Streamline L: primary treatment decision according to the WB-MRI staging pathway and standard staging pathway compared with the final MDT treatment decision 46

TABLE 26 Streamline L: primary treatment decision according to the WB-MRI staging pathway and standard staging pathway compared with the retrospective optimal consensus panel treatment decision 46

TABLE 27 Streamline L: time to complete staging according to treatment pathway 47

TABLE 28 Streamline L: total number of imaging/endoscopic tests before the first major treatment decision according to staging pathway 47

TABLE 29 Unit costs from *NHS Reference Costs 2016/17* used for Streamline C and L 52

TABLE 30 Mean per-participant staging cost ($n = 299$ participants) according to staging pathway: Streamline C 53

TABLE 31 Mean per-participant staging cost ($n = 187$ participants) according to staging pathway: Streamline L 54

TABLE 32 Problems and the diversity of challenge experienced related to WB-MRI 58

TABLE 33 Comparative experience of WB-MRI vs. CT/PET-CT: recovery time and scan acceptability 66

TABLE 34 Comparative scan burden: WB-MRI vs. CT/PET-CT 67

TABLE 35 Attributes and attribute levels 68

TABLE 36 Results of conditional logit regression analysis by cohort 71

TABLE 37 Marginal rates of substitution across all attributes 73

TABLE 38 Streamline C: sensitivity and specificity for metastatic disease according to WB-MRI protocol sequences 78

TABLE 39 Streamline L: sensitivity and specificity for metastatic disease according to WB-MRI protocol sequences 79

TABLE 40 WB-MRI interobserver variability against the consensus reference standard: Streamline C and Streamline L 83

TABLE 41 Quality assurance grading score definitions 113

TABLE 42 Summary of primary and secondary outcomes 117

TABLE 43 Streamline C: reasons for non-inclusion of screened patients 119

TABLE 44 Streamline C: reasons for withdrawal of recruited participants 121

TABLE 45 Streamline C: primary tumour status at the time of consensus review 123

TABLE 46 Streamline C: final T and N stage based on consensus reference standard 125

TABLE 47 Streamline C: local tumour spread based on consensus reference standard 127

TABLE 48 Streamline C: detailed organ sites of metastatic disease 129

TABLE 49 Streamline C: number of metastatic deposits per organ 131

TABLE 50 Streamline C: maximum size of metastatic deposits according to organ 133

TABLE 51 Streamline C: equivocal results for metastatic disease 135

TABLE 52 Streamline C: short-interval follow-up for equivocal findings 137

TABLE 53 Streamline C: per-organ sensitivity and specificity for metastatic disease – standard staging pathways vs. WB-MRI staging pathway against the consensus reference standard 140

TABLE 54 Streamline C: per-participant agreement for tumour T stage – standard staging pathways vs. WB-MRI staging pathway against the consensus reference standard in participants with histological proof 141

TABLE 55 Streamline C: per-participant agreement for tumour N stage – standard staging pathways vs. WB-MRI staging pathway against the consensus reference standard in participants with histological proof 143

TABLE 56 Streamline C: breakdown of the primary treatment decision according to the standard staging pathway compared with the retrospective consensus panel optimal treatment decision (non-rectal colon cancer) 145

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 57 Streamline C: breakdown of the primary treatment decision according to the WB-MRI staging pathway compared with the retrospective consensus panel optimal treatment decision (non-rectal colon cancer) 147

TABLE 58 Streamline C: breakdown of the primary treatment decision according to the standard staging pathway compared with the retrospective consensus panel optimal treatment decision (rectal cancer) 149

TABLE 59 Streamline C: breakdown of the primary treatment decision according to the WB-MRI staging pathway compared with the retrospective consensus panel optimal treatment decision (rectal cancer) 151

TABLE 60 Time to complete staging according to staging pathway: interquartile range 153

TABLE 61 Streamline C: per-participant sensitivity and specificity for metastatic disease – standard staging pathway vs. WB-MRI alone against the consensus reference standard 156

TABLE 62 Streamline C: per-organ sensitivity and specificity for metastatic disease – standard staging pathways vs. WB-MRI alone against the consensus reference standard 158

TABLE 63 Streamline C: per-participant agreement for tumour T stage – standard staging pathways vs. WB-MRI alone against the consensus reference standard – all participants 159

TABLE 64 Streamline C: per-participant agreement for tumour T stage – standard staging pathways vs. WB-MRI alone against the consensus reference standard in participants with histological proof 161

TABLE 65 Streamline C: per-participant agreement for tumour N stage – standard staging pathways vs. WB-MRI alone against the consensus reference standard – all participants 163

TABLE 66 Streamline C: per-participant agreement for tumour N stage – standard staging pathways vs. WB-MRI alone against the consensus reference standard in participants with histological proof 165

TABLE 67 Streamline C: per-participant sensitivity and specificity for metastatic disease: chest, abdomen and pelvis CT vs. WB-MRI alone, against the consensus reference standard 168

TABLE 68 Streamline C: per-participant sensitivity and specificity for metastatic disease – standard staging pathways vs. standard staging pathways plus WB-MRI against the consensus reference standard 170

TABLE 69 Streamline C: perceptual errors for metastatic disease according to staging pathway 171

TABLE 70 Streamline C: perceptual errors for metastatic disease according to staging pathway and organ disease site 173

TABLE 71 WB-MRI image quality scoring 175

TABLE 72 Detailed image quality score categories for scans rated as overall suboptimal category 177

TABLE 73 Detailed image quality score categories for scans rated as overall degraded category 179

TABLE 74 Streamline L: reasons for non-inclusion of screened participants 181

TABLE 75 Streamline L: reasons for withdrawal of recruited participants 183

TABLE 76 Streamline L: primary tumour status at the time of consensus review 185

TABLE 77 Streamline L: final T and N stage based on consensus reference standard 187

TABLE 78 Streamline L: final nodal disease sites based on the consensus reference standard 189

TABLE 79 Streamline L: detailed organs sites of metastatic disease 191

TABLE 80 Streamline L: number of metastatic deposits per organ 193

TABLE 81 Streamline L: maximum size of metastatic deposits according to organ 195

TABLE 82 Streamline L: equivocal results for metastatic disease 198

TABLE 83 Streamline L: short-interval follow-up for equivocal findings 199

TABLE 84 Streamline L: per-organ sensitivity and specificity for metastatic disease – standard staging pathways vs. WB-MRI staging pathway against the consensus reference standard 202

TABLE 85 Streamline L: per-participant agreement for tumour T stage – standard staging pathways vs. WB-MRI staging pathway against the consensus reference standard in participants with histological proof 203

TABLE 86 Streamline L: per-participant agreement for tumour N stage – standard staging pathways vs. WB-MRI staging pathway against the consensus reference standard in participants with histological proof 205

TABLE 87 Streamline L: nodal site agreement – standard staging pathways vs. WB-MRI staging pathway against the consensus reference standard in participants with histological proof 207

TABLE 88 Streamline L: breakdown of the primary treatment decision according to the standard staging pathway compared with the retrospective consensus panel optimal treatment decision 209

TABLE 89 Streamline L: breakdown of the primary treatment decision according to the WB-MRI staging pathway compared with the retrospective consensus panel optimal treatment decision 211

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 90 Streamline L: time to complete staging according to staging pathway 213

TABLE 91 Streamline L: per-participant sensitivity and specificity for metastatic disease-standard staging pathway vs. WB-MRI alone, against the consensus reference standard 216

TABLE 92 Streamline L: per-organ sensitivity and specificity for metastatic disease-standard staging pathways vs. WB-MRI alone against the consensus reference standard 218

TABLE 93 Streamline L: per-participant agreement for tumour T stage – standard staging pathways vs. WB-MRI alone against the consensus reference standard 219

TABLE 94 Streamline L: per-participant agreement for tumour T stage – standard staging pathways vs. WB-MRI alone against the consensus reference standard in participants with histological proof 221

TABLE 95 Streamline L: per-participant agreement for tumour N stage – standard staging pathways vs. WB-MRI alone against the consensus reference standard – all participants 223

TABLE 96 Streamline L: per-participant agreement for tumour N stage – standard staging pathways vs. WB-MRI alone against the consensus reference standard in participants with histological proof 225

TABLE 97 Streamline L: per-participant agreement for metastatic disease: PET-CT versus WB-MRI alone against the consensus reference standard 228

TABLE 98 Streamline L: per-participant sensitivity and specificity for metastatic disease-standard staging pathways vs. standard staging pathways plus WB-MRI against the consensus reference standard 230

TABLE 99 Streamline L: perceptual errors for metastatic disease according to staging pathway 231

TABLE 100 Streamline L: perceptual errors for metastatic disease according to staging pathway and organ disease site 233

TABLE 101 Streamline C: concordance between treatment decisions based on standard and WB-MRI staging pathway (d1 and d2) 235

TABLE 102 Streamline L: concordance between treatment decisions based on standard and WB-MRI staging pathway (d1 and d2) 237

TABLE 103 Streamline C: discordance between treatment decisions based on standard staging (d1) and the consensus reference panel optimal decision (d4) 239

TABLE 104 Streamline C: discordance between treatment decisions based on WB-MRI staging pathway (d1) and the consensus reference panel optimal decision (d4) 241

TABLE 105 Streamline L: discordance between treatment decisions based on standard staging (d1) and the consensus reference panel optimal decision (d4) 243

TABLE 106 Streamline L: discordance between treatment decisions based on WB-MRI staging pathway (d1) and the consensus reference panel optimal decision (d4) 245

TABLE 107 Demographic characteristics of interview participants by trial 249

TABLE 108 Comparison of demographic characteristics of participants sent both questionnaires, who were or were not included in the final analysis (study 2) 251

TABLE 109 Demographic and characteristics of participants who completed the post-staging questionnaire 253

TABLE 110 Comparative experience of WB-MRI vs. CT/PET-CT: scan information, communication and facilities 255

TABLE 111 Demographic and psychological characteristics of the cohort completing the DCE 257

TABLE 112 Results of conditional logit regression analysis: Streamline L participants stratified by home ownership 260

TABLE 113 Results of conditional logit regression analysis: Streamline L participants stratified by education 262

TABLE 114 Results of conditional logit regression analysis: Streamline L participants stratified by self-rated health 264

TABLE 115 Results of conditional logit regression analysis: Streamline L participants stratified by test preference 266

TABLE 116 Participants unwilling to trade attributes off against one another 267

TABLE 117 WB-MRI interobserver variability (not considering agreement with the consensus reference standard): Streamline C and Streamline L 269

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

List of figures

FIGURE 1 Participant flow through the main study: Streamline C 10

FIGURE 2 Participant flow through the main study: Streamline L 11

FIGURE 3 Streamline C participant flow diagram 23

FIGURE 4 Potential impact of staging colorectal cancer with WB-MRI staging pathway or standard staging pathway in a theoretical 1000-patient cohort 29

FIGURE 5 Streamline L: participant flow diagram 37

FIGURE 6 Potential impact of staging non-small-cell lung cancer with WB-MRI or standard staging pathway in a theoretical 1000-participant cohort 43

FIGURE 7 Flow diagram of participants through the studies 1 and 2 65

FIGURE 8 Example of a choice set 69

FIGURE 9 Predicted probabilities of choosing an alternative staging pathway in comparison to a default staging pathway (PET-CT plus one additional scan): Streamline L 74

FIGURE 10 Predicted probabilities of choosing an alternative staging pathways in comparison to a default staging pathway (CT plus one additional scan): Streamline C 74

FIGURE 11 Flow chart of MDT process recording the first major treatment decision based on standard and WB-MRI pathways 115

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

List of boxes

BOX 1 Interview topic guide 247

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

List of supplementary material

Report Supplementary Material 1 Background to power calculations

Supplementary material can be found on the NIHR Journals Library report project page (www.journalslibrary.nihr.ac.uk/programmes/hta/106801/#/documentation).

Supplementary material has been provided by the authors to support the report and any files provided at submission will have been seen by peer reviewers, but not extensively reviewed. Any supplementary material provided at a later stage in the process may not have been peer reviewed.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

List of abbreviations

3D	three dimensional
CI	confidence interval
CRF	case report form
CT	computed tomography
DCE	discrete choice experiment
df	degrees of freedom
DWI	diffusion-weighted imaging
EBUS	endobronchial ultrasonography
FOV	field of view
GHQ-12	General Health Questionnaire-12 items
GRE	gradient echo
HTA	Health Technology Assessment
IDMC	Independent Data Monitoring Committee
IMD	Index of Multiple Deprivation
IT	information technology
MDT	multidisciplinary team
MeSH	Medical Subject Headings
MID	minimally important difference
MRI	magnetic resonance imaging
MRS	marginal rates of substitution
NICE	National Institute for Health and Care Excellence
PABAK	prevalence-adjusted bias-adjusted kappa
PACS	picture archiving and communications system
PC	personal computer
PET-CT	positron emission tomography–computed tomography
REC	Research Ethics Committee
SFOV	small field of view
TMG	Trial Management Group
TNM	Tumour Node Metastasis Classification of Malignant Tumors
TSC	Trial Steering Committee
UCL CTC	University College London Cancer Trials Centre
US	ultrasonography
VATS	video-assisted thoracoscopic surgery
WB-MRI	whole-body magnetic resonance imaging

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Plain English summary

Colorectal and lung cancer are the leading causes of cancer-related deaths in the UK. Optimal treatment depends on accurately defining (or 'staging') the extent of disease, particularly if it has spread to other parts of the body such as the liver. Current staging pathways are complex and rely on a variety of tests that use X-rays, such as computed tomography and positron emission tomography–computed tomography scans. Patients often undergo multiple tests before starting treatment. Alternatively, it is possible to scan the whole body using magnetic resonance imaging without X-rays, and this may be more accurate and reduce the time and number of tests needed before treatment can start. We compared the ability to detect cancer spread, efficiency, patient experience and cost-effectiveness of staging based on whole-body magnetic resonance imaging with the standard NHS pathways in participants newly diagnosed with either lung (187 participants) or colorectal (299 participants) cancer. We found that the whole-body magnetic resonance imaging pathway was as accurate as standard staging pathways and resulted in very similar treatment decisions made by the clinical teams. The whole-body magnetic resonance imaging pathway detected 67% and 50% of participants with cancer spread in colorectal and lung cancer, respectively, compared with 63% and 54%, respectively, for standard staging. However, staging was quicker using whole-body magnetic resonance imaging (by 5 days for colorectal cancer and 6 days for lung cancer) and needed on average one less test to stage colorectal cancer. The whole-body magnetic resonance imaging pathway was also cheaper (costing on average £216 and £317 for colorectal and lung cancer, respectively, compared with £285 and £620, respectively, for standard pathways). Participants generally found whole-body magnetic resonance imaging more burdensome than standard imaging but most preferred the whole-body magnetic resonance imaging pathway if it reduced the time to staging and/or the number of tests. Agreement between different radiology doctors interpreting the same whole-body magnetic resonance imaging scan was moderate for colon cancer and low for lung cancer, emphasising the need for training.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Scientific summary

Background

Colorectal and lung cancer are the leading causes of cancer-related deaths in the UK. Accurate staging is fundamental to the treatment and, in particular, the detection of metastatic disease. Cross-sectional imaging forms the bedrock of cancer staging in the NHS, notably computed tomography, positron emission tomography and magnetic resonance imaging. Ultrasonography, plain radiography and scintigraphy are also widely used. For both cancers, the National Institute for Health and Care Excellence provides guidance on staging pathways, detailing the sequential deployment of various tests before staging is deemed complete and treatment can begin. This stepwise deployment of cross-sectional imaging is resource intensive, exposes patients to ionising radiation and increases their anxiety. Whole-body magnetic resonance imaging has been proposed as an alternative to multimodality staging pathways and can image the body in 1 hour or less using magnetic resonance imaging scanners available throughout the NHS. It does not impart ionising radiation and could reduce the time and number of tests required before staging is complete and treatment can begin. Meta-analyses suggest that whole-body magnetic resonance imaging is equivalent to or may surpass the accuracy of standard tests for metastatic disease detection, although the primary literature is of questionable quality and prospective multicentre comparative data are lacking. High-quality evidence is needed to guide implementation.

Objectives

The primary objective of the Streamline trials was to compare the per-patient diagnostic accuracy for metastatic disease of staging pathways utilising initial whole-body magnetic resonance imaging (whole-body magnetic resonance imaging staging pathway) with those following current National Institute for Health and Care Excellence recommendations (standard pathway). We conducted two parallel but separate trials in colorectal cancer (Streamline C) and non-small-cell lung cancer (Streamline L). Additional objectives were a comparison between alternative pathways for the time and number of tests taken before staging completion and the nature of the first major treatment decision, accuracy for local tumour staging, cost-effectiveness, patient experience, interobserver variability of whole-body magnetic resonance imaging interpretation and diagnostic accuracy of time-efficient whole-body magnetic resonance imaging sequence protocols.

Design

We performed two parallel multicentre prospective cohort trials (Streamline C and Streamline L). For both trials, participants were recruited from 16 representative NHS teaching and general hospitals in England. Patients were eligible for Streamline C if they had histologically proven or suspected colorectal cancer (defined as a mass highly suspicious for cancer on endoscopy or imaging) and were referred for staging investigations. Polyp cancers were excluded. Patients were eligible for Streamline L if they had histologically proven non-small-cell lung cancer or this was suspected on chest computed tomography with sufficient confidence to trigger staging investigations/biopsy and deemed radically treatable (stage IIIB or less). Patients with unequivocal metastatic or N3 disease on diagnostic chest computed tomography were excluded. For both trials, patients were excluded if they were aged < 18 years, were pregnant or had contraindications to magnetic resonance imaging or if the final diagnosis was not lung or colorectal cancer as appropriate. Consecutive (i.e. unselected) patients underwent whole-body magnetic resonance imaging (including T1-weighted imaging pre and post gadolinium enhancement and T2- and diffusion-weighted imaging sequences) in addition to all standard staging investigations. Standard investigations were interpreted locally by the usual clinical care team blinded to the whole-body magnetic resonance imaging. Whole-body

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

magnetic resonance imaging was interpreted by 19 (Streamline C) and 16 (Streamline L) radiologists blinded to all standard staging investigations and clinical data other than the location of the primary tumour. Radiologists recorded the local stage of the primary tumour together with the presence, organ site and size of any metastatic deposits. The whole-body magnetic resonance imaging images and report were withheld from the multidisciplinary team meeting until all standard investigations had completed such that the first major treatment decision could be made (which was recorded). The whole-body magnetic resonance imaging was then revealed to the multidisciplinary team, which stated if additional tests would have been generated had it been the initial staging investigation. The multidisciplinary team then recorded its treatment decision, based on the whole-body magnetic resonance imaging and the results of generated additional tests (if any), the whole-body magnetic resonance imaging staging pathway, and the final treatment decision based on all available tests. We used the construct reference standard paradigm (multidisciplinary panel diagnosis). Participants' clinical courses were followed for 12 months. For each participant, the panel considered the results of all staging and follow-up imaging investigations and all additional information including surgical findings, histopathology and clinical course. Based on all available data, the panel adjudicated on the Tumour Node Metastasis Classification of Malignant Tumors stage of the participant at the time of recruitment and indicated the optimal retrospective treatment decision. For Streamline C, we estimated that a sample size of 290 participants would give 80% power to detect a 10% sensitivity difference between the whole-body magnetic resonance imaging pathway and standard pathways, assuming a 40% prevalence of metastasis. For Streamline L we estimated that a sample size of 200 participants would give 80% power to detect a 24% sensitivity difference between pathways, assuming a 25% prevalence of metastasis. Binary comparisons (sensitivity, specificity, treatment decision agreement) were calculated using paired proportions (population marginal). We calculated the time and number of tests taken to complete staging for each pathway (by adding times for staging tests [from request to performance] to median treatment decision multidisciplinary team wait times). The median difference in time and number of staging tests between pathways was compared for each participant with 95% confidence interval from 2.5 and 97.5 centiles of 1999 bootstrap samples with replacement. We compared the nature of the first major treatment decision with that made by the multidisciplinary team and also with the retrospective optimal treatment decision made by the consensus reference panel. We also tested interobserver variability in whole-body magnetic resonance imaging interpretation and the impact of whole-body magnetic resonance imaging sequences on radiologist accuracy (T1- and diffusion-weighted sequences vs. T1-, diffusion- and T2-weighted sequences vs. T1-, diffusion- and T2-weighted sequences and contrast-enhanced T1-weighted sequences), investigated participant experience staging pathways using interviews, questionnaires and a discrete choice experiment, and performed a cost-effectiveness analysis.

Results

Streamline C

A total of 299 participants completed the trial; 68 (23%) had metastasis. The whole-body magnetic resonance imaging pathway had 67% (95% confidence interval 56% to 78%) sensitivity for participants with metastasis, not significantly different from standard pathways [63% (95% confidence interval 51% to 74%)], a 4% (95% confidence interval –5% to 13%; $p = 0.508$) difference in sensitivity. Specificity was not significantly different [whole-body magnetic resonance imaging pathway 95% (95% confidence interval 92% to 97%) vs. standard pathways 93% (95% confidence interval 90% to 96%)]. The whole-body magnetic resonance imaging staging pathway had 86% (95% confidence interval 74% to 94%) sensitivity for participants in whom the largest metastasis was $\geq$ 1 cm in size, which was not significantly different from that of standard staging pathways [82% (95% confidence interval 69% to 91%)]. As a stand-alone investigation, whole-body magnetic resonance imaging had a comparable sensitivity of 70% (95% confidence interval 59% to 80%) to the standard staging pathway, but a significantly lower specificity [86% (95% confidence interval 81% to 90%) vs. 94% (95% confidence interval 90% to 96%), respectively]. The whole-body magnetic resonance imaging staging pathway had 54% agreement for T stage compared with 60% for the standard pathway, a difference of 6% (95% confidence interval 0% to 12%). There was no significant difference in agreement between the whole-body magnetic resonance imaging and standard

pathways for N staging (58% vs. 56%, respectively), a difference of 2% (95% confidence interval –4% to 7%). Agreement with the final multidisciplinary team treatment decision was 96% and 95% for whole-body magnetic resonance imaging and standard pathways, respectively, and 68% with the retrospective optimal treatment decision for both pathways. Time to complete staging was significantly shorter for whole-body magnetic resonance imaging [8 days (95% confidence interval 6 to 9 days) vs. 13 days (95% confidence interval 11 to 15 days)], a 5-day (95% confidence interval 3- to 7-day) difference. Across the cohort, a total of 558 and 320 investigations were performed as part of standard pathways and the whole-body magnetic resonance imaging staging pathway, respectively (whole-body magnetic resonance imaging would have generated an additional 21 tests). The median number of tests in the whole-body magnetic resonance imaging pathway was significantly lower [1 test (95% confidence interval 1 to 1 test) vs. 2 tests (95% confidence interval 2 to 2 tests)], a difference of 1 test (95% confidence interval 1 to 1 test).

The mean test costs per participant (bootstrapped 95% confidence intervals) were £285 (95% confidence interval £260 to £310) for standard pathways and £216 (95% confidence interval £211 to £221) for the whole-body magnetic resonance imaging pathway.

Streamline L

A total of 187 participants completed the trial; 52 (28%) had metastasis. The whole-body magnetic resonance imaging pathway had 50% (95% confidence interval 37% to 63%) sensitivity for participants with metastasis, not significantly different from standard pathways [54% (95% confidence interval 41% to 67%)], a 4% (95% confidence interval –7% to 15%; $p = 0.73$) difference. Specificity was not significantly different [93% (95% confidence interval 88% to 96%) vs. 95% (95% confidence interval 91% to 98%), respectively]. The whole-body magnetic resonance imaging staging pathway had 82% (95% confidence interval 64% to 92%) sensitivity for participants in whom the largest metastasis was ≥ 1 cm, which was not significantly different from that of standard staging pathways [75% (95% confidence interval 57% to 87%)]. As a stand-alone investigation, whole-body magnetic resonance imaging had comparable sensitivity [50% (95% confidence interval 37% to 63%)] to standard pathways but lower specificity [85% (95% confidence interval 78% to 90%)]. The whole-body magnetic resonance imaging staging pathway had 65% agreement for N stage compared with 75% for the standard pathway, a statistically significant difference of –10% (95% confidence interval –3% to –18%). There was no significant difference in agreement between the whole-body magnetic resonance imaging and standard pathways for T stage (54% vs. 55%, respectively), a difference of 1% (95% confidence interval –8% to 9%). Agreement with the final multidisciplinary team treatment decision was 98% and 99% for whole-body magnetic resonance imaging and standard pathways, respectively, and 83% and 82%, respectively, with the retrospective optimal treatment decision. Time to complete staging was significantly shorter for whole-body magnetic resonance imaging [13 days (95% confidence interval 12 to 14 days) vs. 19 days (95% confidence interval 17 to 21 days)], a 6-day (95% confidence interval 4- to 8-day) difference. Across the cohort a total of 302 and 232 investigations were performed as part of standard and whole-body magnetic resonance imaging staging pathways, respectively (whole-body magnetic resonance imaging would have generated an additional 45 tests). The median number of tests in the whole-body magnetic resonance imaging pathway were similar [1 test (95% confidence interval 1 to 1 test) vs. 1 test (95% confidence interval 1 to 2 tests)], a difference of 0 tests (95% confidence interval –1 to 0 tests). Mean test costs per participant (bootstrapped 95% confidence intervals) were £620 (95% confidence interval £574 to £666) for standard pathways and £317 (95% confidence interval £273 to £361) for the whole-body magnetic resonance imaging pathway.

Participant experience

In general, whole-body magnetic resonance imaging presented a greater challenge than standard scans. Key challenges were the enclosed space, noise and scan duration. Reduced participant tolerance was associated with claustrophobia, pulmonary symptoms and existing comorbidities. Coping strategies facilitated scan tolerance. Perceived whole-body magnetic resonance imaging burden was greater than for computed tomography (mean score 2.09 vs. 1.70, respectively; $p < 0.0001$) and positron emission tomography–computed tomography (2.33 vs. 2.05; $p = 0.003$). However, participants preferred the

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

whole-body magnetic resonance imaging-based pathway (probability: lung 0.64, colorectal 0.66) if it was equivalent in accuracy, total scan number and time to diagnosis to a standard staging pathway. Preference was stronger if whole-body magnetic resonance imaging reduced time to staging and/or number of tests.

Whole-body magnetic resonance imaging sequence selection

Streamline C

A combination of T1-weighted imaging pre and post gadolinium enhancement and T2- and diffusion-weighted sequences had significantly greater sensitivity for participants with metastatic disease than T1- and diffusion-weighted sequences alone [72% (95% confidence interval 60% to 81%) vs. 63% (95% confidence interval 51% to 73%), respectively], a difference of 9% (95% confidence interval 1% to 17%). Specificity was not significantly different [85% (95% confidence interval 80% to 89%) vs. 84% (95% confidence interval 78% to 88%), respectively]. A combination of non-contrast enhanced T1-, T2- and diffusion-weighted sequences had 70% (95% confidence interval 58% to 80%) sensitivity and specificity of 86% (81% to 90%).

Streamline L

A combination of T1-weighted imaging pre and post gadolinium enhancement and T2- and diffusion-weighted sequences had significantly greater sensitivity for participants with metastatic disease than T1- and diffusion-weighted sequences alone [52% (95% confidence interval 39% to 65%) vs. 42% (95% confidence interval 29% to 56%), respectively], difference of 10% (95% confidence interval 1% to 19%). Specificity was not significantly different [86% (95% confidence interval 79% to 81%) vs. 82% (95% confidence interval 74% to 87%), respectively]. A combination of non-contrast enhanced T1-, T2- and diffusion-weighted sequences had 48% (95% confidence interval 35% to 61%) sensitivity and 84% (95% confidence interval 76% to 89%) specificity.

Whole-body magnetic resonance imaging interobserver agreement

Interobserver agreement for interpretation of whole-body magnetic resonance imaging against the final reference standard was moderate for Streamline C (80% overall agreement against the consensus reference standard, $\kappa = 0.6$), but poor for Streamline L (44% overall agreement against the consensus reference standard, $\kappa = -0.12$). For the primary outcome, not considering the final participant metastatic status based on the consensus reference standard, two reads agreed in 35 out of 40 (88%) whole-body magnetic resonance imaging data sets (Streamline C) and 28 out of 43 (65%) whole-body magnetic resonance imaging data sets (Streamline L).

Conclusions

For both colorectal and non-small-cell lung cancer, the whole-body magnetic resonance imaging staging pathway is as accurate as current standard staging pathways for identifying patients with metastatic disease and results in the same treatment decisions. It is more efficient, reducing time to complete staging, the number of staging tests (in colorectal cancer) and staging costs.

Implications for health care

In a NHS setting, the whole-body magnetic resonance imaging staging pathway achieves similar accuracy as standard staging pathways for detecting patients with metastatic disease in both colorectal cancer and non-small-cell lung cancer. Agreement for local T and N stage is also similar, although in non-small-cell lung cancer, the whole-body magnetic resonance imaging pathway has lower agreement and sensitivity for N stage than standard pathways. However, for both cancers, agreement between the primary treatment decisions based on the whole-body magnetic resonance imaging pathway and both a contemporaneous multidisciplinary team treatment decision and a retrospective optimal treatment decision is nearly identical

to that for decisions based on standard staging pathways. For colorectal cancer, the whole-body magnetic resonance imaging pathway reduces the number of tests required to complete staging, and for both cancers whole-body magnetic resonance imaging significantly reduces the time to complete staging. Based on NHS reference costs, the whole-body magnetic resonance imaging staging pathway is cheaper than standard pathways in both colorectal and non-small-cell lung cancer. Although patients find whole-body magnetic resonance imaging a more burdensome test than standard staging investigations, in general they prefer the whole-body magnetic resonance imaging staging pathway if it at least matches the accuracy and efficiency (time and number of tests) of standard pathways. Time-efficient whole-body magnetic resonance imaging protocols using just T1- and diffusion-weighted sequences have lower sensitivity than protocols that also include T2- and post-gadolinium-enhanced T1-weighted sequences. Interobserver variation in whole-body magnetic resonance imaging interpretation is moderate, and lower for non-small-cell lung cancer than for colorectal cancer.

Recommendations for future research

Future research should investigate:

1. the diagnostic accuracy, patient acceptability and cost-effectiveness of whole-body magnetic resonance imaging for staging and treatment follow-up of other primary cancer sites, notably breast, prostate and myeloma, compared with standard investigations
2. the diagnostic accuracy, patient acceptability and cost-effectiveness of whole-body magnetic resonance imaging for investigating patients with clinically suspected recurrence of colorectal and non-small-cell lung cancer compared with standard investigations
3. the diagnostic accuracy, patient acceptability and cost-effectiveness of whole-body magnetic resonance imaging in routine post-cancer therapy surveillance in comparison with standard investigations
4. the impact on diagnostic accuracy and cost-effectiveness of adding liver-specific contrast agents to whole-body magnetic resonance imaging protocols
5. the impact of formalised training on radiologist performance and interobserver agreement.

Trial registration

These trials are registered as ISRCTN43958015 and ISRCTN50436483.

Funding

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

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Chapter 1 Introduction

Background

Accurate staging of cancer is fundamental to optimal treatment. Staging relates to the local tumour extent as well as the presence or otherwise of disease remote from the primary organ of origin – so-called metastatic disease. Conventionally, cancer stage is expressed using the Tumour Node Metastasis Classification of Malignant Tumors (TNM) system, with T denoting the characteristics of the primary tumour, based for example on size and invasion of local structures, N stage denoting the involvement or otherwise of lymph nodes and M stage denoting metastatic status.

Colorectal and lung cancer are the third and fourth most common cancers in the UK, accounting for 13% and 12% of all new cancers, respectively, and they are the leading causes of cancer-related deaths.[1] In both cancers, detection of metastatic disease is fundamental to treatment strategy. For example, after surgical removal of the primary tumour, up to 50% of patients with colon cancer relapse with undiagnosed metastatic disease, usually to the liver and/or lung.[2] Upfront detection of this disease would allow informed clinical decision-making and the appropriate use of chemotherapeutic, surgical and ablative therapies.[3] Similarly, prognosis in lung cancer is closely related to stage at diagnosis; correctly identified and surgically treated non-metastatic early-stage disease is associated with 5-year survival rates of between 54% and 80%, although at least 20% of patients undergoing lung surgery with curative intent relapse with undiagnosed metastatic disease – so-called futile thoracotomy.[4]

Cross-sectional imaging forms the bedrock of cancer staging in the NHS. A range of imaging tests are available including computed tomography (CT), positron emission tomography–computed tomography (PET-CT) and magnetic resonance imaging (MRI). Ultrasonography (US), plain radiography and scintigraphy are also widely used. The choice of imaging investigations is governed by the specific cancer type and predilection for metastatic spread, as imaging modalities differ in their diagnostic accuracy across the various organ sites.

For both lung and colorectal cancer, the National Institute for Health and Care Excellence (NICE) has provided guidance on staging pathways,[5,6] which details the deployment of imaging tests as part of standard staging pathways. In the case of colorectal cancer, CT of the chest, abdomen and pelvis is recommended, supplemented by pelvic MRI for local staging of rectal cancer. Although the NICE does not currently recommend additional imaging, in routine clinical practice it is not unusual for patients to undergo PET-CT and/or liver MRI if metastatic disease is suspected. Staging pathways in lung cancer are more complex, with CT, PET-CT, MRI, US and endobronchial/percutaneous biopsy all recommended at various points in the staging algorithm.

The stepwise deployment of multimodality cross-sectional imaging as part of the staging pathways in colorectal and lung cancer is resource intensive and exposes patients to relatively high doses of ionising radiation, which may increase their risk of subsequent malignancy.[7] Furthermore, patients have understandably very high anxiety levels when being investigated for suspected cancer, and this anxiety increases if the diagnostic and staging process is protracted.[8]

Whole-body magnetic resonance imaging (WB-MRI) has been proposed as an alternative to multimodality staging pathways. Using modern MRI scanners available throughout the NHS, it is possible to acquire images through the body (skull vertex to mid-thigh) in 60 minutes or less. Importantly, WB-MRI does not impart diagnostic ionising radiation to patients, and promising data support its ability to stage

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

malignancy, particularly in detection of metastatic disease. It could also increase staging efficiency, reducing the time to complete staging and the number of tests required. Overall, WB-MRI may therefore be a safer, more efficient and accurate alternative to complex multimodality staging pathways in colorectal and lung cancer.

Existing literature on the diagnostic accuracy of WB-MRI in cancer staging

We searched PubMed and Embase for articles between 1 January 1990 and 1 October 2018 without language restriction. We used MeSH (Medical Subject Headings) and a full-text search for 'cancer', 'neoplasm', 'staging', 'diagnostic accuracy', 'magnetic resonance imaging', 'whole body imaging', 'diffusion magnetic resonance imaging', 'colon', 'colorectal', 'bronchial' and 'lung'. Meta-analyses and systematic reviews were identified using appropriate search limits.

There have been a number of meta-analyses and systematic reviews reporting the accuracy of WB-MRI for cancer staging.[9–20] Many have combined multiple cancer types in one single meta-analysis.[9–11,13,14,16,17] Those meta-analyses considering the staging of only lung cancer have limited themselves to the detection of bone[12,15] or nodal[18–20] metastasis. No meta-analysis has considered colon cancer in isolation, and the largest study to date recruited just 20 participants.[21]

A variety of comparators have been selected but the majority compare WB-MRI with PET-CT and scintigraphy (in the case of bone metastasis). A summary of the main meta-analyses and systematic reviews is shown in *Table 1*.

Those meta-analyses that consider all organ sites (albeit in a range of primary tumour sites) have generally reported similar performance between WB-MRI and conventional imaging (notably PET-CT) for detection of metastasis, with sensitivity between 80% and 90% and specificity in excess of 95%. Studies reporting the accuracy for bone metastasis detection have been inconsistent. Those limited to lung cancer staging alone have generally reported lower sensitivity for WB-MRI. For example, Qu *et al.*[15] and Liu *et al.*[12] reported the sensitivity of WB-MRI at 77% and 80% compared with 92% and 95%, respectively, for PET-CT. Conversely, in a more recent meta-analysis of mixed primary tumour sites (for which lung cancer was the largest contributor), Duo *et al.*[10] estimated the sensitivity of WB-MRI to be 84% compared with 80% for PET-CT. Most studies contributing to the secondary literature are small, including < 50 participants. In the largest study in lung cancer to date, Ohno *et al.*[22] investigated WB-MRI staging compared with PET-CT in 203 non-small-cell lung cancer patients referred as potential surgical candidates to a single university hospital. All imaging was read in consensus by two experienced radiologists or nuclear medicine physicians. Of the 203 participants, 40 had metastatic disease. The sensitivity and specificity of WB-MRI for metastatic disease was 70% (95% CI 53% to 83%) and 92% (95% CI 87% to 96%), respectively, compared with 63% (95% CI 46% to 77%) and 95% (95% CI 90% to 97%), respectively, for PET-CT. The largest study to date investigating WB-MRI in colon cancer staging included just 20 participants and compared diagnostic performance to PET-CT.[21] WB-MRI detected four more liver metastases than PET-CT (27 vs. 24, respectively) but fewer lung metastases (19 vs. 25, respectively).

As well as deficiencies in the secondary literature resulting from grouping disparate primary tumour sites together and/or focusing on one specific organ site of metastatic spread, the majority of contributory studies are small single-site explanatory studies with imaging interpreted by a small number of experienced radiologists. Generalisability is also limited because most studies investigate single modality comparisons (e.g. WB-MRI vs. PET-CT) rather than comparing 'real-life' complex multimodality staging pathways.[23] Most studies also use a single optimised MRI platform, rather than a range of platforms that typifies NHS practice. Finally, there has been no study on the influence of alternative staging pathways on actual treatment decisions, which is crucial to the evaluation of any new technology.[24] Overall, the current evidence base is insufficiently mature to guide the implementation or otherwise of WB-MRI in colorectal and/or non-small-cell lung staging pathways within the NHS.

TABLE 1 Summary of recent meta-analyses and systematic reviews reporting the diagnostic accuracy of WB-MRI in cancer staging

Study	Year	Modalities considered	Cancer type	Number of included studies (number of participants)	Number of participants in largest contributing WB-MRI study	Main outcome	Main findings
Xu *et al.*[9]	2013	PET-CT, WB-MRI	Mixed (melanoma, lung, breast, naso/oropharynx, colorectal)	13 (1239)	203 (lung)	Diagnostic accuracy for all site metastatic disease	Per-patient sensitivity: WB-MRI 86% (95% CI 70% to 94%); PET-CT 85% (95% CI 68% to 94%) Per-patient specificity: WB-MRI 97% (95% CI 94% to 99%); PET-CT 96% (95% CI 95% to 97%) Per-lesion sensitivity: WB-MRI 89% (95% CI 81% to 94%); PET-CT 85% (95% CI 79% to 90%) Per-lesion specificity: WB-MRI 89% (95% CI 81% to 94%); PET-CT 90% (95% CI 82% to 94%)
Duo *et al.*[10]	2013	PET-CT, gadolinium-enhanced WB-MRI	Mixed (melanoma, lung, breast, naso/oropharynx)	9 (1116)	203 (lung)	Diagnostic accuracy for bone metastasis	Per-patient sensitivity: WB-MRI 84% (95% CI 63% to 94%); PET-CT 80% (95% CI 64% to 90%) Per-patient specificity: WB-MRI 98% (95% CI 95% to 99%); PET-CT 99% (95% CI 97% to 94%)
Li *et al.*[11]	2014	PET-CT, diffusion-weighted WB-MRI	Mixed (colorectal, myeloma, breast, lymphoma, lung, kidney, thyroid, prostate, liver)	8 (584)	203 (lung)	Diagnostic accuracy for primary lesion and metastatic disease	Per-patient/-lesion sensitivity: WB-MRI 90% (95% CI 88% to 92%); PET-CT 90% (95% CI 87% to 92%) Per-patient/-lesion specificity: WB-MRI 95% (95% CI 94% to 96%); PET-CT 98% (95% CI 97% to 98%)

continued

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 1 Summary of recent meta-analyses and systematic reviews reporting the diagnostic accuracy of WB-MRI in cancer staging (*continued*)

Study	Year	Modalities considered	Cancer type	Number of included studies (number of participants)	Number of participants in largest contributing WB-MRI study	Main outcome	Main findings
Liu *et al.*[12]	2011	PET-CT, PET, WB-MRI, bone scintigraphy	Lung	14 [4241 (PET, 2446; WB-MRI 258; scintigraphy, 1537)]	115	Diagnostic accuracy for bone metastasis	Per-patient sensitivity: WB-MRI 80% (95% CI 67% to 90%); PET 87% (95% CI 81% to 92%); PET-CT 95% (95% CI 91% to 97%); scintigraphy 92% (95% CI 89% to 94%) Per-patient specificity: WB-MRI 91% (95% CI 86% to 94%); PET 95% (95% CI 93% to 97%); PET-CT 98% (95% CI 97% to 98%); scintigraphy 69% (95% CI 66% to 72%)
Wu *et al.*[13]	2011	WB-MRI	Mixed (breast, lung, renal, prostate)	11 (495)	115	Diagnostic accuracy for bone metastasis	Per-patient sensitivity: WB-MRI 90% (95% CI 85% to 94%) Per-patient specificity: WB-MRI 92% (95% CI 88% to 95%)
Yang *et al.*[14]	2011	PET-CT, CT, WB-MRI, bone scintigraphy	Mixed	67 (10,760 [PET, 4367; CT, 723; WB-MRI, 1032; scintigraphy, 4638])	115	Diagnostic accuracy for bone metastasis	Per-patient sensitivity: PET 87% (95% CI 84% to 90%); CT 73% (95% CI 67% to 79%); PET-CT 94% (95% CI 91% to 96%); MRI 91% (95% CI 87% to 94%); scintigraphy 86% (95% CI 84% to 88%) Per-patient specificity: PET 96% (95% CI 96% to 97%); CT 95% (95% CI 94% to 97%); PET-CT 97% (95% CI 96% to 98%); MRI 95% (95% CI 94% to 97%); scintigraphy 81% (95% CI 80% to 83%)

Study	Year	Modalities considered	Cancer type	Number of included studies (number of participants)	Number of participants in largest contributing WB-MRI study	Main outcome	Main findings
Qu *et al.*[15]	2012	PET-CT, PET, WB-MRI, bone scintigraphy	Lung	17 (2940 [WB-MRI, 252])	115	Diagnostic accuracy for bone metastasis	Per-patient sensitivity: WB-MRI 77% (95% CI 65% to 87%); PET 87% (95% CI 81% to 92%); PET-CT 92% (95% CI 88% to 95%); scintigraphy 86% (95% CI 82% to 89%) Per-patient specificity: WB-MRI 92% (95% CI 88% to 95%); PET 94% (95% CI 92% to 96%); PET-CT 98% (95% CI 97% to 98%); scintigraphy 88% (95% CI 86% to 89%)
Wu *et al.*[16]	2013	WB-MRI, bone scintigraphy	Mixed (lung, breast, renal, prostate)	7 (328)	115	Diagnostic accuracy for bone metastasis	Per-patient sensitivity: WB-MRI 84% (95% CI 72% to 91%); scintigraphy 83% (95% CI 73% to 89%) Per-patient specificity: WB-MRI 96% (95% CI 81% to 99%); scintigraphy 94% (95% CI 68% to 99%)
Smets *et al.*[17]	2018	WB-MRI	Paediatric solid tumours (neuroblastoma, sarcoma, hepatoblastoma, renal)	5 (132)	39	Diagnostic accuracy for bone metastasis	Per-lesion sensitivity: MRI range 82%–100%
Peerlings *et al.*[18]	2016	MRI	Lung	12 (1122)	250	Diagnostic accuracy for lymph node status	Per-patient sensitivity: MRI 87% (95% CI 78% to 92%) Per-patient specificity: MRI 88% (95% CI 77% to 94%) Per-nodal sensitivity: MRI 88% (95% CI 78% to 94%) Per-nodal specificity: MRI 95% (95% CI 87% to 98%)

continued

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

DOI: 10.3310/hta23660

TABLE 1 Summary of recent meta-analyses and systematic reviews reporting the diagnostic accuracy of WB-MRI in cancer staging (*continued*)

Study	Year	Modalities considered	Cancer type	Number of included studies (number of participants)	Number of participants in largest contributing WB-MRI study	Main outcome	Main findings
Shen *et al.*[19]	2016	MRI	Lung	18 (1116)	250	Diagnostic accuracy for lymph node status	Per-patient sensitivity: MRI 68% (95% CI 63% to 73%) Per-patient specificity: MRI 92% (95% CI 90% to 94%) Per-nodal sensitivity: MRI 72% (95% CI 69% to 75%) Per-nodal specificity: MRI 96% (95% CI 96% to 96%)
Wu *et al.*[20]	2012	MRI, PET-CT	Lung	19 (2845)	250	Diagnostic accuracy for lymph node status	Per-patient sensitivity: MRI 72% (95% CI 63% to 80%); PET-CT 75% (95% CI 68% to 81%) Per-patient specificity: MRI 95% (95% CI 85% to 98%); PET-CT 89% (95% CI 85% to 91%)

CI, confidence interval.

Objectives of the Streamline studies

The primary aim of the Streamline studies was to compare the per-patient diagnostic accuracy for metastatic disease of staging pathways utilising initial WB-MRI with those following the current NICE recommendations (the 'standard pathway').[23] We conducted two parallel but separate trials in colorectal cancer (Streamline C) and non-small-cell lung cancer (Streamline L). Secondary objectives were the time and number of tests taken to complete staging and the nature of the first major treatment decision based on WB-MRI or standard pathway, diagnostic accuracy for local tumour staging, cost-effectiveness of staging pathway, patients' experiences of staging using WB-MRI compared with standard staging pathways, the interobserver variability of WB-MRI diagnosis of metastatic disease by different radiologists and the diagnostic accuracy of WB-MRI protocols limited to certain sequence combinations. Additional objectives included the diagnostic accuracy of standard pathways with WB-MRI as an additional test and of WB-MRI as a single stand-alone replacement test.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Chapter 2 Methods

Parts of this chapter are reproduced from Taylor *et al.*[25] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Parts of this chapter are reproduced from Taylor *et al.*[26] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

The trial protocol has been previously published.[23]

Study design

The Streamlines studies were multicentred, non-randomised, single-arm, prospective cohort studies comparing the diagnostic accuracy for metastatic disease of staging pathways based on initial WB-MRI with NICE-recommended standard staging pathways in colorectal cancer (Streamline C) and non-small-cell lung cancer (Streamline L). Ethics permission was granted by Camden and King's Cross Research Ethics Committee [Streamline C: Research Ethics Committee (REC) reference number 12/LO/1176; Streamline L: REC reference number 12/LO/1177] in October 2012 and the trials were conducted in accordance with the principles of Good Clinical Practice. The trial was co-ordinated by Cancer Research UK, University College London Cancer Trials Centre (UCL CTC) and an independent Data Monitoring Committee and Trial Steering Committee (TSC). All participants gave written informed consent prior to participation.

The two trials (Streamline C and Streamline L) were separate but conducted in parallel. Consecutive (i.e. unselected) eligible patients with known or highly suspected colorectal or lung cancer underwent WB-MRI in addition to standard staging investigations performed as part of their usual care. Multidisciplinary team (MDT) treatment decisions based on the standard staging investigations were documented prior to the revelation of the WB-MRI findings, after which the MDT documented its theoretical treatment decision based on the WB-MRI (and any additional tests it generated). A final treatment decision was then made based on all available tests, including WB-MRI. Participants' clinical course was followed for a period of 12 months. A multidisciplinary consensus panel derived the reference standard for the tumour TNM stage, including the sites of metastatic disease (if present) using all available clinical, imaging, biochemical and histological data over the 12-month follow-up period. A summary of patient flow in the main trial is shown in *Figures 1* and *2*.

Agreement between radiologists' interpretation of WB-MRI was tested in a sample of participants and the contribution of specific MRE sequences on radiologists' accuracy investigated. Participants' experiences of staging using WB-MRI compared with standard staging pathways, and the priorities placed on differing pathway attributes, was investigated using interviews, structured questionnaires and a discrete choice experiment (DCE). The cost-effectiveness of WBI-MRI and standard pathways was assessed in an economic evaluation. The full study protocols for Streamline C and Streamline L are available on the project web page [URL: www.journalslibrary.nihr.ac.uk/programmes/hta/106801 (accessed 12 September 2019)].

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

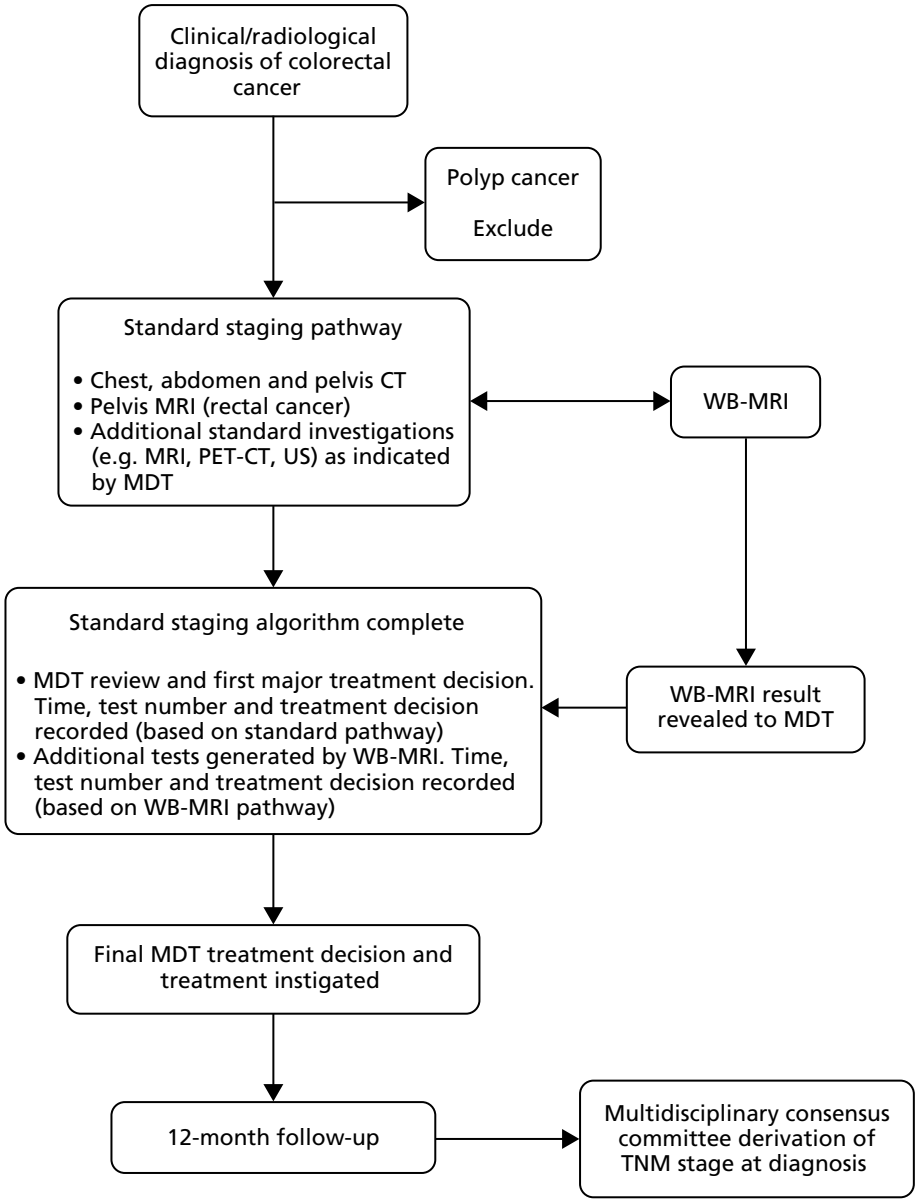


FIGURE 1 Participant flow through the main study: Streamline C.

Patient and public involvement

The Streamline studies were developed in collaboration with patient representatives who joined the trial team at the inception of the project. The patient representatives helped refine the research questions, devise the protocol and successfully apply for the funding. By way of example, their advice and guidance was fundamental to the assessment of patient experience. All patient-facing materials were designed with the patient representatives (and were in general very well received by participants). The representative sat on the trial management and steering committees, providing ongoing guidance throughout the running of the trial and subsequent write-up, for example helping to refine recruitment strategies and advising on dissemination. This collaboration has been very productive and patient representatives will aid dissemination of the work via patient forums.

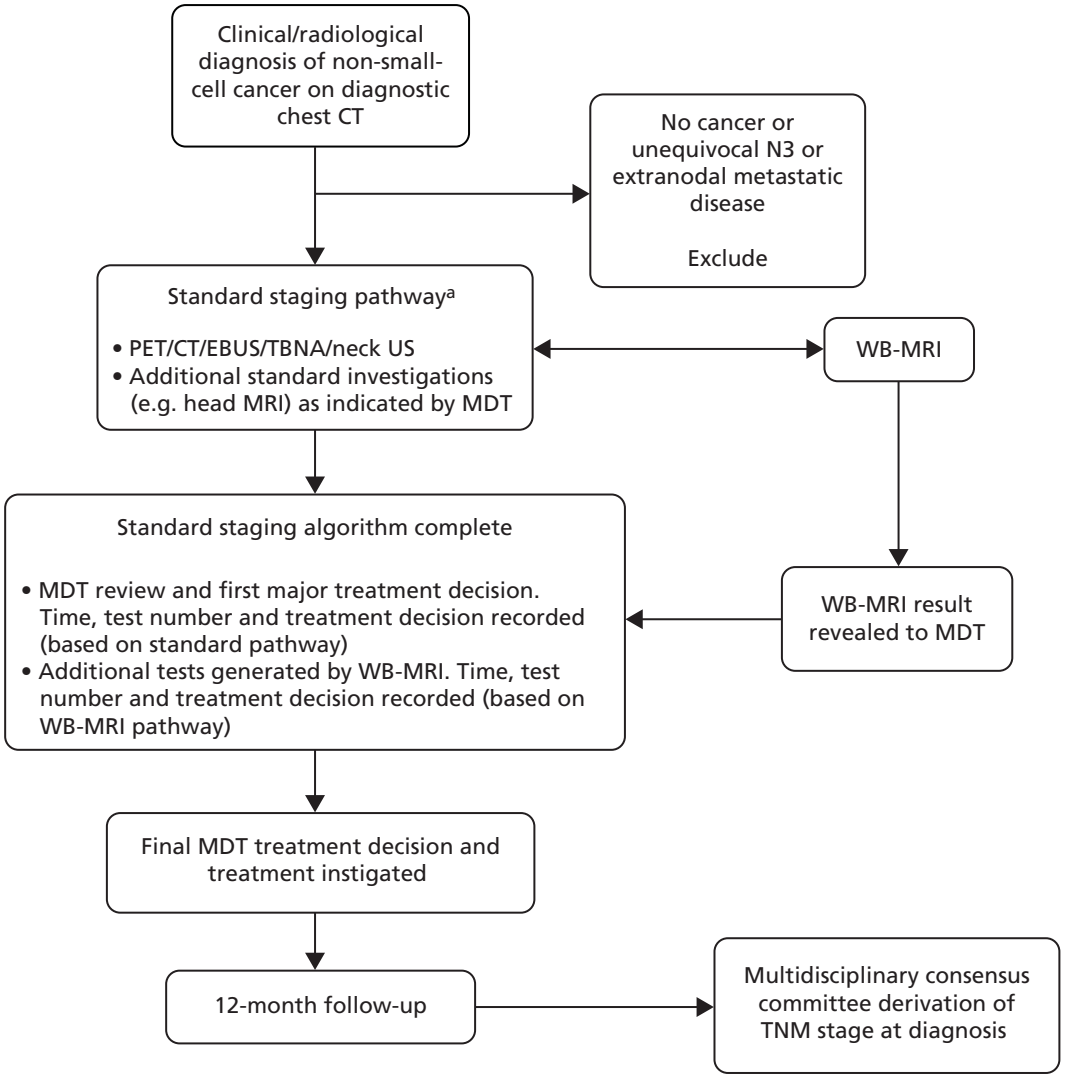


FIGURE 2 Participant flow through the main study: Streamline L. a, Specific choice of investigations dependent on CT findings (notably tumour location and mediastinal nodal size). EBUS, endobronchial ultrasonography; TBNA, transbronchial needle aspiration; VATS, video-assisted thoracoscopic surgery.

Recruitment sites

Streamline L recruited participants from 16 NHS hospitals in England and Streamline C also recruited from 16 NHS hospitals in England. Participants were recruited to both trials from 11 hospitals. Recruitment sites were a mixture of district general hospitals and teaching hospitals. Recruited participants from sites without the infrastructure to perform WB-MRI locally underwent WB-MRI at nearby imaging hubs (in total eight and 11 imaging hubs for Streamline C and L, respectively). Some sites acted as both recruitment sites and imaging hubs. A summary of recruitment sites and their associated imaging hub is shown in *Table 2*.

Inclusion criteria: Streamline C

- Adult patients (aged ≥ 18 years) with histologically proven or suspected colorectal cancer referred for staging.
- Suspicion of colorectal cancer defined as presence of a mass highly suspicious for colorectal cancer on endoscopy, barium enema, CT colonography or other imaging that triggers staging investigations.
- Patient must have given written informed consent and be willing to comply with the protocol intervention and follow-up.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 2 Recruitment sites and imaging hubs

Streamline C		Streamline L	
Recruitment site	**Imaging hub**	**Recruitment site**	**Imaging hub**
UCLH	UCLH	UCLH	UCLH
Barnet and Chase Farm NHS Hospitals Trust[a]		Barnet and Chase Farm NHS Hospitals Trust	
Princess Alexandra Hospital		Princess Alexandra Hospital	
North Middlesex University Hospital		North Middlesex University Hospital	
St Mark's Hospital		Whittington Hospital	
West Middlesex University Hospital		Queen's Hospital	
Whittington Hospital		Lister Hospital	Paul Strickland Scanner Centre, Mount Vernon Hospital
Queen's Hospital		Croydon University Hospital	Royal Marsden Hospital
Homerton University Hospital[b]	Homerton University Hospital	Kingston Hospital	
St Bartholomew's Hospital/ Royal London Hospital	St Bartholomew's Hospital/ Royal London Hospital	Charing Cross Hospital	Charing Cross Hospital
Charing Cross Hospital	Charing Cross Hospital	Homerton Hospital[c]	Homerton Hospital
University Hospital Southampton NHS Foundation Trust	University Hospital Southampton NHS Foundation Trust	Guy's and St Thomas' NHS Foundation Trust	Guy's and St Thomas' Hospital NHS Foundation Trust
Queen Alexandra Hospital	Queen Alexandra Hospital	Lewisham Hospital	
Guy's and St Thomas' NHS Foundation Trust	Guy's and St Thomas' NHS Foundation Trust	St Bartholomew's Hospital	St Bartholomew's Hospital/ Royal London Hospital
Lewisham Hospital		Whipps Cross University Hospital	
Bradford Royal Infirmary	Bradford Royal Infirmary	Newham University Hospital	

UCLH, University College London Hospital.
a One participant underwent WB-MRI at the Paul Strickland Scanner Centre, Mount Vernon Hospital.
b 10 participants underwent WB-MRI at St Bartholomew's Hospital.
c One participant underwent WB-MRI at St Bartholomew's Hospital and one at UCLH.

Exclusion criteria: Streamline C

- Any psychiatric or other disorder likely to have an impact on informed consent.
- Evidence of severe or uncontrolled systemic disease that makes it undesirable for the patient to participate in the trial.
- Pregnancy.
- Contraindications to MRI (e.g. cardiac pacemaker, severe claustrophobia, inability to lie flat).
- Polyp cancer.

Inclusion criteria: Streamline L

- Adult patients (aged ≥ 18 years) with suspected primary non-small-cell lung cancer on chest CT with sufficient confidence to trigger staging investigations/biopsy or with already histologically proven primary non-small-cell lung cancer.
- Disease is potentially radically treatable as defined as stage IIIB or less on diagnostic chest CT (i.e. T1–4, N0–2, M0).
- Performance status 0–2 (fit to undergo radical treatment if indicated).
- Patient must have given written informed consent and be willing to comply with the protocol intervention and follow-up.

Exclusion criteria: Streamline L

- Unequivocal metastatic or N3 disease on diagnostic chest and abdomen CT (including M1a disease, malignant pleural effusion).
- Further staging work-up not indicated in the opinion of the MDT or clinician owing to poor performance status or patient choice.
- Histologies other than non-small-cell lung cancer.
- Any psychiatric or other disorder likely to have an impact on informed consent.
- Evidence of severe or uncontrolled systemic disease that makes it undesirable for the patient to participate in the trial.
- Pregnancy.
- Contraindications to MRI (e.g. cardiac pacemaker, severe claustrophobia, inability to lie flat).

Test methods

WB-MRI

In both trials, recruited participants underwent WB-MRI performed at their recruitment site/designated imaging hub in addition to all standard staging investigations performed as part of their usual clinical care. The choice of MRI platform (i.e. manufacturer and tesla strength) was decided by the local hub lead radiologist according to scanner availability and their usual practice. A minimum data set of sequences was acquired including whole-body axial diffusion-weighted imaging (DWI), whole-body axial T2-weighted imaging, whole-body axial or coronal pre-contrast T1-weighted imaging and axial T1-weighted imaging post intravenous gadolinium contrast through at least the liver, lungs and head. It was stipulated that, in general, scan acquisition should not take longer than 1 hour. A full description of the WB-MRI acquisition minimum data set is given in *Appendix 1*. The WB-MRI was performed concurrently with the standard staging investigations and no later than 3 weeks after the final standard staging investigation.

WB-MRI blinding

To maintain the integrity of the trial, WB-MRI scans were reported by radiologists blinded to the standard imaging tests and other clinical information (other than the cancer diagnosis and anatomical location – either colonic segment for Streamline C or lung lobe for Streamline L). In addition, to ensure that radiologists reporting the standard staging investigations and other MDT members were not unblinded to the WB-MRI images and/or report until the designated time of revelation in the MDT (see *Revelation of WB-MRI findings and MDT treatment decisions*), unanonymised WB-MRI images were not immediately sent to the picture archiving and communications system (PACS) at either the imaging hub or the recruitment sites. Instead, images were uploaded to a secure central imaging server, 3Dnet™ (Biotronics3D, London, UK). This solution allowed easy upload of MRI data sets via a standard internet connection. A personal computer (PC)-based internet gateway was installed in each imaging hub to facilitate automated transfer of WB-MRI from the scanner or workstation to 3Dnet, and thereafter automatically back to PACS at the appropriate time point after MDT revelation (see *Revelation of WB-MRI findings and MDT treatment decisions*).

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

WB-MRI reporting and radiologist competence/training

Whole-body magnetic resonance imaging was interpreted by designated radiologists at each imaging hub who were experienced in interpretation based on previous experience of reporting WB-MRI in cancer staging and were post fellowship of the Royal College of Radiologists. Specifically, all had interpreted at least 20 validated WB-MRI cases in patients with lung or colorectal cancer. Thereafter, radiologists with experience of < 100 WB-MRI data sets underwent a period of 'buddy' reporting with a more experienced radiologist (experience of > 100 WB-MRI data sets) and were only permitted to report alone once deemed competent by the more experienced radiologist.[25,26] This process was specifically designed to mirror how WB-MRI would be reported in standard NHS practice were it to become disseminated. We specifically avoided using a small number of highly experienced subspecialty radiologists who would not be representative of the NHS radiological workforce. As noted above, the radiologists were blinded to the standard staging investigations performed on recruited participants. In total, 19 radiologists interpreted WB-MRI for Streamline C and 16 radiologists interpreted WB-MRI for Streamline L.

Interpretation was performed using the 3Dnet visualisation software, or a stand-alone workstation after data set download, according to the preference of the radiologist.

Images were analysed in three blocks in the following order:

1. diffusion-weighted and non-contrast-enhanced T1-weighted images
2. diffusion-weighted and non-contrast-enhanced T1- and T2-weighted images
3. diffusion-weighted, non-contrast enhanced T1- and T2-weighted images and gadolinium-enhanced T1-weighted images.

After viewing each sequence block (and before reviewing the next block) the radiologist completed a WB-MRI imaging booklet case report form (CRF) – available on the project web page [URL: www.journalslibrary.nihr.ac.uk/programmes/hta/106801 (accessed 12 September 2019)] – documenting their findings. Items recorded on the CRFs included the availability of WB-MRI sequences, location and T and N stage of the local tumour (TNM 7th edition)[27] and the presence, location and size of metastatic disease. The presence of metastatic disease was recorded for a range of soft tissue and bony anatomical sites using six confidence levels grouped into normal (levels 1 and 2), equivocal (levels 3 and 4) and abnormal (levels 5 and 6). If disease presence was recorded as equivocal or abnormal (i.e. confidence level 3 or above), the size of the largest two metastatic deposits per organ was recoded, along with the additional number of deposits < 6 mm, 6–9 mm and ≥ 10 mm. Reporting radiologists were instructed to interpret the WB-MRI as they would in their routine clinical practice, taking into account the known morphology and characteristics of potential metastatic disease across the various sequences.[22,25,26] Following completion of the CRF, the reporting radiologist produced a free-text clinical report as per their usual clinical practice (using all available sequences and based on the TNM 7th edition)[27] for subsequent release to the clinical team. This report contained information relating to the local T and N stage of the tumour, together with the presence, location, number and size of metastatic deposits as well as important 'incidental' findings, for example a second malignancy. The radiologists were instructed to express their level of confidence in their report as they would in their normal clinical practice and to reflect their confidence scores entered on the CRF. If the radiologist would usually recommend additional tests for equivocal findings, this was included in the report to also mirror routine clinical practice and guide the MDT. The free-text report was uploaded onto the 3Dnet software.

Quality assurance

Whole-body magnetic resonance imaging scans were scored centrally by a radiographer with 8 years of experience in WB-MRI using a predefined scoring system based on technical quality (1 to 5) and anatomical coverage (1 to 4) (see *Appendix 2*, *Table 41*). Scans were then classified as optimal (maximum score for both technical quality and anatomical coverage), suboptimal but fully diagnostic (technical quality score ≥ 3, anatomical coverage score ≥ 2) or degraded (one or more sequences substantially degraded, technical quality score < 3, anatomical coverage score < 2). The radiographer visited all imaging hubs during the first months of their opening to feedback on their WB-MRI imaging quality and advise on improvements (if required).

Standard investigations

Recruited participants underwent all the standard staging investigations employed at their recruiting institution according to local protocols and the requirements of their clinical care team. All standard investigations were performed and interpreted by the usual radiologists and clinicians employed at the site where they were performed [at the local recruitment site or usual tertiary referral site for more complex investigations such as PET-CT and endobronchial ultrasonography (EBUS)] as per usual clinical care. The images and findings of the WB-MRI were not available to the local clinical care teams until the point of revelation to the MDT. The date of the first staging investigation and all subsequent investigations was recorded. Clinical reports for the standard investigations (and their results) were freely available on hospital PACSs, radiology information systems and Clinical Data Repositories as per usual clinical practice. The type and date of all standard investigations (e.g. CT, PET-CT, organ-specific MRI, biopsy) was recorded on CRFs, along with the presence and location of metastatic disease based on the radiological reports. Standard imaging and standard imaging disease CRFs are available on the project web page [URL: www.journalslibrary.nihr.ac.uk/programmes/hta/106801 (accessed 12 September 2019)].

Revelation of WB-MRI findings and multidisciplinary team treatment decisions

Whole-body magnetic resonance images and reports were withheld from clinical care teams until participants had completed all their standard staging investigations such that their recruitment site MDT was able to make the first major treatment decision. Each MDT compromised key personnel from a range of hospital specialties and as a minimum included a radiologist, oncologist, histopathologist and respiratory physician and/or surgeon (Streamline L) or colorectal surgeon (Streamline C). The first major treatment decision was defined as:

- referral for surgical excision of the primary tumour and/or a metastatic site
- instigation of definitive treatment using chemotherapy, radiotherapy or a combination of the two
- decision to offer palliative/supportive care only
- request for a highly invasive surgical staging procedure such as surgical mediastinal lymph node sampling (mediastinoscopy), video-assisted thoracoscopic surgery (VATS) or laparoscopy.

Recruited participants were discussed in the relevant MDT as per their usual care pathways. After reviewing all standard imaging and all available clinical data, the MDT stated if it had sufficient information to make the first major treatment decision. If not, and the participant needed further investigations, these were performed and the participant rediscussed at the next available MDT. Once sufficient staging information was available to make the first major treatment decision, the MDT documented this, along with the TNM stage (based on the standard imaging performed) on a dedicated MDT CRF – available on the project web page [URL: www.journalslibrary.nihr.ac.uk/programmes/hta/106801 (accessed 12 September 2019)]. The MDT then recorded the first major treatment decision based on the standard imaging pathway from a range of options on the CRF. If the participant was referred to another MDT for the final treatment decision (for example to a specialised liver MDT), the CRF was transferred to and completed by this MDT.

The WB-MRI report was then revealed to the same MDT. Usually this was by electronically accessing 3Dnet in the MDT room via internet-enabled computer such that the WB-MRI report and key images (saved by the reporting radiologist) could be projected to the whole MDT. Presentation of a paper copy of the report was permissible in cases of information technology (IT) failure. The MDT reviewed the WB-MRI report (and key images) and stated whether or not WB-MRI would have generated additional tests before the first major treatment decision if it had been the initial stand-alone staging investigation, for example a bone scan for equivocal bone lesion or PET-CT for an indeterminant lymph node. If additional tests would have been generated but these had already been performed as part of the standard care pathway, the nature and results of these additional tests was recorded on the MDT CRF. However, if these had not already been performed and the MDT considered them essential for patient care (e.g. if WB-MRI identified a new potentially important site of disease or incidental abnormality), these were requested and performed

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

and the participant rediscussed in the next available MDT once complete. The MDT then recorded the TNM stage (and sites and number of metastasis) based on the WB-MRI and the results of additional tests it generated (if any) – the 'WB-MRI staging pathway' – and stated what its theoretical first major treatment decision would have been had only the investigations in the WB-MRI staging pathway been available to them. The MDT then stated its final treatment decision based on all available tests (i.e. standard pathway, WB-MRI and generated tests). In this way (and to assuage ethics concerns), the findings of WB-MRI could be considered when deciding the final patient treatment. Occasionally, and for logistical reasons, the treatment decision based on the WB-MRI pathway was documented after the MDT. In summary, the following were therefore recorded:

- stage and treatment decision based on standard investigations (and the number, timing, nature and findings of these investigations)
- stage and theoretical treatment decision based on the WB-MRI staging pathway (and the number, timing, nature and findings of any additional tests generated)
- final treatment decision incorporating all available tests
- a summary flow chart of the MDT's first major treatment decision-making process (see *Appendix 3*, *Figure 11*).

Finally, the MDT recorded on the CRF if any pathway would have resulted in short-interval follow-up imaging, for example for indeterminate findings such as small lung nodules.

If a participant was on an accelerated treatment pathway such that they were due to commence treatment before a MDT where the WB-MRI could be revealed, an ad hoc MDT was convened consisting of all specialties relevant to the participant's clinical care and the MDT process followed as described above.[23]

Early release of WB-MRI findings

The WB-MRI report and images were released in an unanonymised form to the hospital PACS and electronic patient record after revelation during the MDT process as described in *Revelation of WB-MRI findings and multidisciplinary team treatment decisions*. However, it was permitted to release the WB-MRI findings to the clinical teams before this time point if WB-MRI revealed a serious finding that could have an immediate impact on direct patient care, for example impending spinal cord compression, deep-vein thrombosis, pulmonary embolism or brain metastasis with significant mass effect requiring immediate treatment. Reporting radiologists were instructed to contact the participant's clinician to discuss the finding and a decision was made about whether or not the results should be revealed early to all members of the clinical team (based on review of standard tests already performed that may also have detected the finding).

Recruitment

Suitable patients were identified from outpatient clinics, MDT meetings, inpatient wards, imaging requests and endoscopy lists by members of the local research team, who established whether or not the patient met the trial entry criteria (for Streamline C or L as appropriate). A screening log recorded the details of all patients approached to take part in the trials and reasons for non-participation if applicable. All patients were handed or posted a participant information sheet detailing the study (applicable to Streamline C or L as appropriate) – available on the project web page [URL: www.journalslibrary.nihr.ac.uk/programmes/hta/106801 (accessed 12 September 2019)] – and the contact details of study team. The study purpose and requirements were also explained face to face to patients by an appropriately trained member of the research team. All participants gave written consent prior to participation in the Streamline trials. Participants gave additional written consent if they agreed to participate in the patient experience studies (see *Chapter 6*). Participants retained a copy of their consent form and participant information sheet and were informed that they could withdraw from the study at any time.

Data collection and follow-up

Data collation was co-ordinated by UCL CTC. Participant age, performance status and sex were collected, together with the request date of the first staging investigation [defined as the date of request of the first staging investigation following a proven or assumed diagnosis of colorectal (Streamline C) or non-small-cell lung (Streamline L) cancer, for example the date of request for chest, abdomen and pelvis CT after colonoscopic diagnosis of probable malignant tumour or PET-CT following chest CT diagnosis of probable lung tumour]. The completion date of staging was also recorded (defined as the date of the final staging investigation). The time between this and the MDT's first major treatment decision was also recorded, as was the date of WB-MRI. The type and date of all standard investigations performed prior to the first major treatment decision (e.g. CT, PET-CT, organ-specific MRI, biopsy) were recorded on CRFs, along with the presence and location of metastatic disease based on the radiological reports – available on the project web page [URL: www.journalslibrary.nihr.ac.uk/programmes/hta/106801 (accessed 12 September 2019)]. Complications related to the WB-MRI were recorded on a specific CRF.

Recruited participants were followed for a period of 12 months (or until the date of death if sooner) to inform the consensus panel review process (see *Reference standard*) and collate data for the health economic analysis (see *Chapter 5*). During this time, details of any imaging investigations, surgical interventions (and biopsies) and cancer treatments were recorded on CRFs, along with outpatient visits, hospital day visits and inpatient stays. Histological findings from any surgical resections or biopsies were also recorded on CRFs. All CRFs were collated by local site research nurses/practitioners and sent to the clinical trial unit by post or fax. Forms were entered on to a bespoke study database and any missing fields or apparent data inaccuracies queried with the centre to optimise data collection.

Reference standard

Multidisciplinary consensus panel review is the standard methodology for diagnostic test accuracy studies where an independent reference standard does not exist or is impossible because of incorporation bias.[23,25,26,28] The Streamline trials applied this methodology. Specifically, each recruitment site convened a series of consensus panels to derive the reference standard for the cancer TNM stage at recruitment for their participants. The panel consisted of one radiologist external to the recruitment site and at least one radiologist internal to the recruitment site or associated imaging hub. A second internal radiologist attended if required to ensure sufficient expertise in cross-sectional imaging and nuclear medicine techniques such as PET-CT was available to the panel. In addition, an oncologist and/or a colorectal surgeon (Streamline C) or respiratory physician/thoracic surgeon (Streamline L) attended the vast majority of meetings. Occasionally, such an individual was not available to attend and reviewed the panel decisions after the meeting. The panel had access to a histopathologist if required. A member of the UCL CTC attended each meeting to ensure full data collection and uniformity of the consensus process. Typically, each consensus meeting considered around 10–20 participants in a 2- to 3-hour session. The panel considered all available clinical information over the 12 months follow-up period, including the images and results of staging investigations (including WB-MRI, CT, PET-CT, MRI), surgical findings (if applicable), histopathology (surgical resection and biopsies), all follow-up imaging and the participants' clinical course. Panels were given a summary of pertinent data extracted from trial CRFs and had full access to clinical results, records and letters via the patient notes and/or electronic patient record. Panels also had access to the hospital PACS so that they could review all imaging during the meeting.

Based on all available data, the panel adjudicated on the TNM stage of the participant at the time of recruitment. Emphasis was placed on the available histological data, for example following surgical resections of the primary tumour and/or biopsy data of remote disease sites. In the absence of histological proof of metastasis, metastatic disease was assumed if new lesions appeared during the 12-month follow-up period with imaging characteristics compatible with metastasis and no alternative explanation, or if already present lesions with characteristics compatible with metastasis either grew or shrank (on therapy). Lesions that

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

remained stable over the 12-month follow-up period were assumed not to be metastatic unless there were specific circumstances considered by the panel that indicated malignancy (e.g. change in lesion morphology with treatment).

Given the variable follow-up period (e.g. owing to participant death) and the potential for primary tumours left in situ to seed new metastatic disease after cancer diagnosis (which would unfairly penalise diagnostic accuracy estimates of initial staging investigations), the following criteria were applied by the panel when opining on metastatic status (defined upfront as part of the trial protocols):[25,26]

- For participants in whom the primary tumour was completely removed within 3 months of diagnosis, all new metastatic sites identified over the follow-up period were be assumed to have been present at diagnosis.
- If the primary tumour was left in situ for > 3 months after diagnosis (or there was incomplete removal), new metastatic sites were assumed to have been present at diagnosis if they were identified within 6 months of diagnosis. If new metastatic sites were diagnosed > 6 months after diagnosis, and there was no evidence of their presence on retrospective review of all staging investigations, they were assumed to be new disease and not present at diagnosis.
- If participants with tumours left in situ did not undergo any imaging capable of detecting metastatic disease within 6 months of diagnosis of the primary tumour, and new metastatic sites were apparent beyond 6 months but not visible in retrospect on any trial imaging, the consensus panel decided if the disease was probably present at diagnosis based on its location, size and imaging characteristics.
- If a participant died before the 12-month follow-up, the panel reviewed all available imaging, histopathology and clinical course prior to death and in consensus decided if a confident diagnosis of the presence or absence of metastatic disease could be made [e.g. the presence of imaging characteristics compatible with metastasis and no alternative explanation or if lesions with characteristics compatible with metastasis either grew or shrank (on therapy)]. If this judgement could not be made with confidence (e.g. if the participant had equivocal lesions on staging investigations and no further follow-up), the participant was excluded.

Using a specific CRF – available on the project web page [URL: www.journalslibrary.nihr.ac.uk/programmes/hta/106801 (accessed 12 September 2019)] – the panel recorded for each participant the location of the primary tumour, TNM stage and the availability of histopathological proof for T and N stage. For colorectal cancer (Streamline C), the panel also recorded whether or not the circumferential resection margin was involved and whether or not there was presence of extra mural vascular invasion. For non-small-cell lung cancer (Streamline L), a detailed breakdown of nodal involvement across the 14 TNM stations was recorded. The presence of metastatic disease was recorded overall, and then on a per-organ basis. For each involved organ the rational for the diagnosis of metastatic disease was given (histological proof, characteristic imaging appearances, growth on follow-up and/or response to therapy). The size of the largest two organ deposits at the time of staging was recorded, measured using the most appropriate staging investigation if pathological measurement from a surgical resection was not available, together with the number of deposits in the organ as a whole (split into < 6 mm, 6–9 mm and ≥ 10 mm). If the metastatic deposit had appeared after the initial staging investigation and was not visible in retrospect on any staging investigation yet still fulfilled the panel criteria for presence at diagnosis, it was marked as not visible. Diffuse metastatic disease such as peritoneal or pleural disease was marked as non-measurable as appropriate. In addition, the consensus panel reviewed all imaging tests (including WB-MRI) performed during staging and indicated if there were perceptual errors (i.e. unreported metastatic disease visible on retrospective review of the standard staging pathway or WB-MRI alone or the WB-MRI staging pathway).

Finally, the panel considered the final allocated TNM stage, all follow-up data (including the WB-MRI) and the patient outcomes and indicated the optimal retrospective treatment decision using the same options as those provided to the initial staging MDT.

Outcomes

A summary of the primary and secondary outcomes is shown in *Appendix 4*, *Table 42*. The primary outcome was the difference in per-patient sensitivity for metastatic disease between standard staging pathways and the WB-MRI staging pathway against the consensus reference standard, that is the ability of the pathways to identify patients with metastatic disease at the time of diagnosis. Differences in specificity was a secondary outcome. Disease reported as equivocal was generally treated as positive for disease presence given the potential clinical implications of an equivocal result on patient management and the need for further investigations. However, for rectal cancer only, on advice of the Independent Data Monitoring Committee (IDMC), disease reported as equivocal was treated as disease negative (rectal cancer patients often undergo chemoradiation treatment prior to surgery, allowing additional time for equivocal lesion characterisation). A sensitivity analysis was also performed, treating equivocal results as negative for disease presence (or positive for rectal cancer). Subanalyses divided participants into those whose largest metastatic deposit was $\geq$ 1 cm and those for whom it was < 1 cm to assess the impact of lesion size on diagnostic accuracy based on a size cut-off commonly reported in the literature and pathway differences in sensitivity and specificity according to specific organ sites of metastatic disease. The number of participants equivocal for metastatic disease and those with radiologist perceptual errors was also noted.

Secondary outcomes also included the difference in per-patient sensitivity and specificity for metastatic disease between standard staging pathways and WB-MRI as a stand-alone test (based on the radiologist WB-MRI report alone and to allow comparison with the existing literature, which usually considers WB-MRI as a stand-alone examination),and between standard staging pathways and the combination of standard tests with WB-MRI in addition, both against the consensus reference standard. Additional secondary outcomes were differences between standard staging pathways and the WB-MRI staging pathway for the nature of the first major treatment decision (compared with the final MDT treatment decision and consensus panel retrospective optimal treatment decision), time taken to complete staging and the number of tests needed to complete staging. Secondary outcomes also included differences in per-patient agreement for local T and N stage disease between standard staging pathways and the WB-MRI staging pathway, WB-MRI as a stand-alone test and the combination of standard tests with WB-MRI, all against the consensus reference standard, and as an exploratory outcome against a histopathological standard of reference when available.

Additional secondary outcomes pertaining to the lifetime incremental cost and cost-effectiveness of standard staging pathways and the WB-MRI staging pathway, comparative patient experience of staging pathways, diagnostic impact of WB-MRI sequence blocks and interobserver variation in the evaluation WB-MRI are described in the relevant chapters (see *Chapters 5–8*).

Sample size

Primary outcome

The sample-size calculation was based on the primary outcome stipulated by the Health Technology Assessment (HTA) commissioning brief: diagnostic accuracy for metastatic disease.

Streamline C

Based on the available literature at the time of study design (see *Report Supplementary Material 1*) it was assumed that WB-MRI would achieve 84.8% sensitivity for liver metastasis (by far the most common site for metastatic disease) compared with 74.9% for contrast-enhanced CT. Using methods for comparative studies,[29] a sample size of 290 participants would give 80% power to detect this 10% sensitivity difference between the WB-MRI pathway and standard pathways, assuming a 40% prevalence of metastasis, and 73% concordance between pathways. Allowing for a 10% drop-out rate (loss to follow-up) at 1 year gave a target sample size of 322 participants. This was revised by the IDMC to 360 participants owing to a higher than expected withdrawal rate of 19%.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Streamline L

Based on the available literature at the time of study design (see *Report Supplementary Material 1*) it was assumed that WB-MRI would achieve greater sensitivity, particularly for bone and brain metastasis, than standard pathways, notably PET-CT. Overall it was assumed that WB-MRI would have 79% per-patient sensitivity for metastatic disease compared with 55% per-patient sensitivity for conventional staging (PET-CT). Using methods for comparative studies,[29] a sample size of 200 participants would give 80% power to detect this 24% sensitivity difference between the WB-MRI pathway and standard pathways assuming a 25% prevalence of metastasis (12% isolated brain, 5% isolated bone and 8% at other sites) and 53% concordance between pathways. Allowing for a 20% drop-out rate (loss to follow-up) at 1 year gave a target sample size of 250 participants. This was revised by the IDMC to 353 participants owing to a higher than expected withdrawal rate of 43%.

Analysis

We report our prespecified primary and secondary outcomes and additional sensitivity analyses. Sensitivity (per patient) for each imaging pathway was defined as the percentage of participants identified as having metastasis by the reference standard who had a positive finding for metastasis in the imaging pathway. Binary comparisons (sensitivity, specificity, treatment decision agreement) were calculated using paired proportions (population marginal) in Stata® 14.2 [StataCorp LP, College Station, TX, USA]. For the Streamline L primary outcome, equivocal disease was considered positive for metastasis. Sensitivity analysis treated equivocal results as negative. For the Streamline C primary outcome, disease reported as equivocal for rectal cancer was considered disease-negative. Sensitivity analysis treated equivocal results as all negative or all positive. There were no missing data for the primary outcome so imputation was not required. Statistical significance was based on 95% confidence intervals (CIs) using the Newcombe paired proportion method[30] (McNemar's test *p*-values are reported). Pathway treatment decisions were grouped for analysis (see *Tables 3* and *4*) and compared with the final MDT and consensus panel treatment decisions (as a sensitivity analysis).

TABLE 3 Streamline C: grouping of treatment decisions for analysis

Summary treatment decision category	Component treatment decisions
Surgery for the primary tumour but no chemotherapy	Surgical removal of primary tumour alone
Surgery for the primary tumour and chemotherapy (and/or radiotherapy)	Surgery for primary tumour followed by planned adjuvant chemotherapy
	Neoadjuvant chemotherapy (and/or radiotherapy) alone followed by anticipated planned surgical removal of primary tumour
Chemotherapy (and/or radiotherapy) without surgery	Neoadjuvant chemotherapy (and/or radiotherapy) alone
	Palliative care
Surgical metastectomy with or without chemotherapy	Surgery for primary tumour followed by planned chemotherapy followed by surgical removal of metastasis
	Surgical removal of primary tumour and metastatic site(s) alone
	Surgery for primary tumour and metastatic site(s) followed by anticipated planned adjuvant chemotherapy
	Neoadjuvant chemotherapy (and/or radiotherapy) alone followed by anticipated planned surgical removal of primary tumour and metastatic site(s)

Reproduced from Taylor *et al.*[25] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

TABLE 4 Streamline L: grouping of treatment decisions for analysis

Summary treatment decision category	Component treatment decisions
Treatment with curative intent	Surgical removal of primary tumour alone
	Radical radiotherapy alone
	Combination chemoradiotherapy
Treatment with non-curative intent	Chemotherapy alone
	Non-radical dose radiotherapy alone
	Supportive/palliative care

Reproduced from Taylor *et al.*[26] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Time to complete staging pathways (excluding initial diagnostic tests) was calculated in days by adding times for staging tests (from request to performance) to median treatment decision MDT wait times, calculated across all participants. In the case of missing data, median times from the same or similar tests were used. The median difference in time and number of staging tests between pathways was compared for each participant with 95% CI from the 2.5th and 97.5th percentiles of 1999 bootstrap samples with replacement.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Chapter 3 Results: Streamline C

Parts of this chapter are reproduced from Taylor *et al.*[25] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Participants

Recruitment commenced in March 2013 and was completed in August 2016. Overall, 1020 patients were assessed for eligibility, of whom 650 were excluded (*Figure 3* and *Appendix 5, Table 43*).

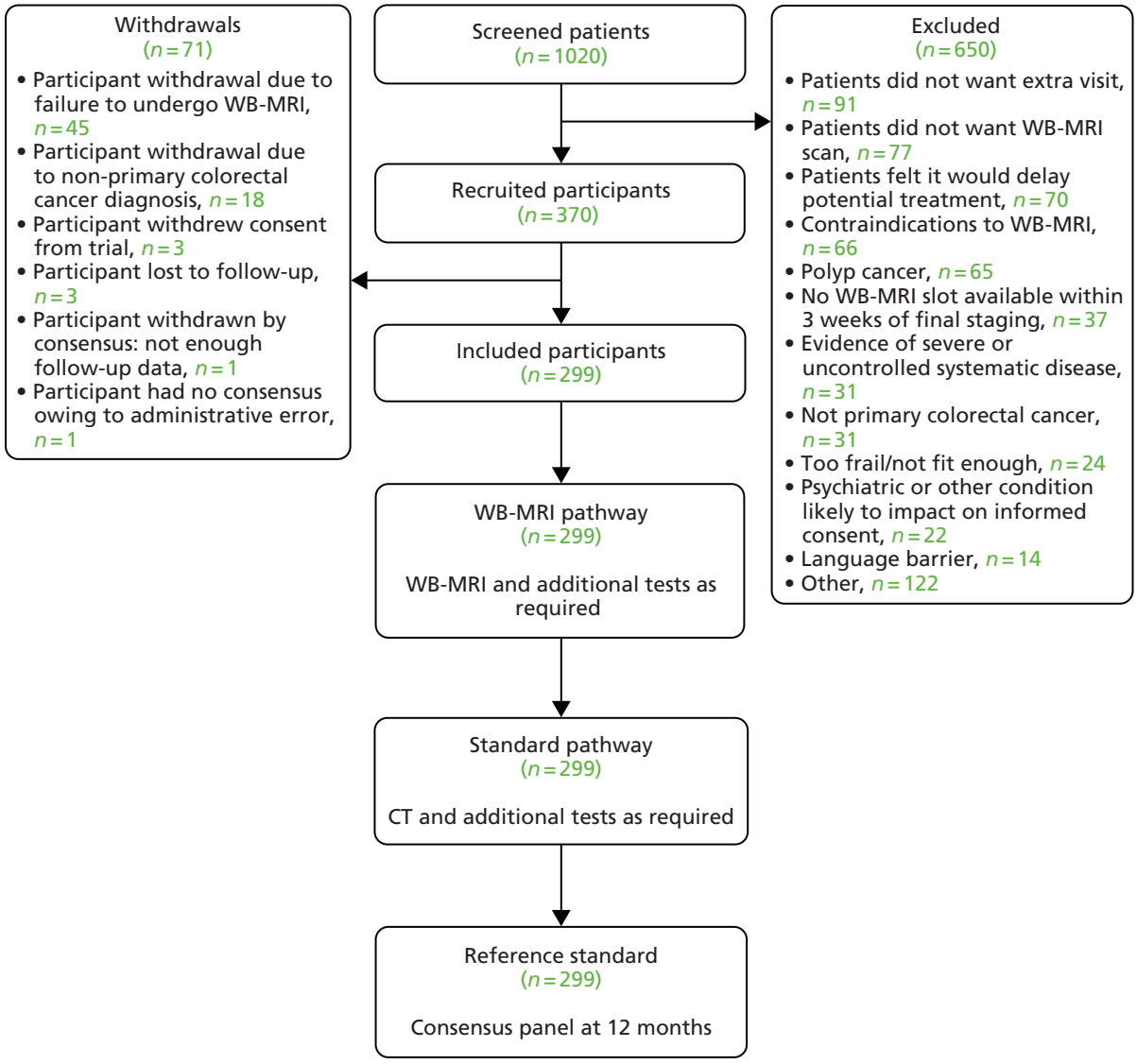


FIGURE 3 Streamline C participant flow diagram. Reproduced from Taylor *et al.*[25] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Of the 370 participants entering the trial, 71 were withdrawn, mainly owing to a final diagnosis other than colorectal cancer ($n = 18$) or failure to undergo MRI, usually due to scheduling issues ($n = 45$) (see *Figure 3* and *Appendix 6, Table 44*). The final cohort consisted of 299 participants, including 106 (35%) women (*Table 5* and see *Figure 3*). There we no reported series adverse reactions.

Overall, 17 participants (6%) died within the 12-month follow-up period with enough collected data for the consensus panels to be confident in assigning a TNM stage.

The status of the primary tumour and follow-up imaging available to the panels is shown in *Appendix 7*, *Table 45*.

Of the 299-participant cohort, as defined by the consensus panel, 288 participants (96%) were stage T2 or above and 166 participants (56%) were node positive (See *Appendix 8, Table 46*).

In total, 24 participants (8%) had a threatened resection margin and 118 participants (40%) had extra mural venous invasion (See *Appendix 9, Table 47*).

Overall, 68 participants (23%) had metastatic disease and 231 participants (77%) did not. The organ location of the metastatic disease and the basis for diagnosis is shown in *Table 6*. A more detailed

TABLE 5 Demographics of the final study cohort

Characteristic	*n* (%)
Sex	
Male	193 (65)
Female	106 (35)
Age (years)	
Mean (SD)	64 (12)
Range	30–90
Performance status	
Fully active	199 (67)
Ambulatory (able to work)	31 (10)
Ambulatory (not able to work)	3 (1)
Not recorded	66 (22)
Tumour location[a]	
Rectum	130 (43)
Sigmoid	86 (29)
Descending	11 (4)
Transverse[b]	24 (8)
Ascending	29 (10)
Caecum	43 (14)

SD, standard deviation.
a By consensus reference standard. Some participants have multiple tumour locations.
b Flexures tumours were combined and categorised as transverse colon.
Reproduced from Taylor *et al.*[25] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

TABLE 6 Streamline C: organ sites of metastatic disease

Organ site	Number with metastatic disease[a]	Histological proof, *n* (%)	Imaging diagnosis without histological proof,[b] *n* (%)		
			Characteristic imaging appearances	Growth on follow-up	Response to therapy
Liver	48	12 (25)	33 (92)	26 (72)	19 (53)
Lung	20	0 (0)	19 (95)	13 (65)	11 (55)
Bone	3	0 (0)	3 (100)	2 (67)	1 (33)
Mesentery/peritoneum	7	2 (29)	4 (80)	4 (80)	1 (20)
Nodal (metastatic)	11	2 (28)	8 (89)	7 (78)	3 (33)
Other[c]	2	1 (33)	2 (100)	2 (100)	0 (0)

a Participants may have more than one site of metastatic disease.
b Metastases may fulfil more than one criterion.
c Two participants with metastasis in three other sites (skin, ovary and vagina).

Reproduced from Taylor *et al.*[25] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

breakdown is shown in *Appendix 10*, *Table 48*. The majority of participants with metastatic disease had liver and/or lung metastasis, with three having metastatic spread to the bone.

The number of deposits per organ site is shown in *Appendix 11*, *Table 49*. The maximum size of the metastasis at the time of staging is shown in *Appendix 12*, *Table 50*. Of the 48 participants with liver metastasis, 37 participants' largest deposit was ≥ 1 cm, and of the 20 participants with lung metastasis, only five participants' largest deposit was ≥ 1 cm. There were six participants in whom metastasis was not visible on any staging investigation, even in retrospect, and became apparent only during the follow-up period (but fulfilled the a priori definitions for metastatic disease at the time of staging).

A total of 10 participants had a second malignancy reported.

Equivocal results

The number of equivocal results according to the two staging pathways and WB-MRI alone is shown in *Appendix 13*, *Table 51*. On a per-participant basis, four of the 68 participants with metastatic disease had equivocal results based on the WB-MRI pathway, compared with seven participants based on the standard pathway and nine participants based on WB-MRI alone. Of the 231 participants without metastases, seven participants had equivocal results based on the WB-MRI pathway, compared with 14 participants based on the standard pathway and 25 participants based on WB-MRI alone.

Streamline C: staging pathway tests

The type and number of standard staging tests performed prior to the first major treatment decision across the whole cohort is shown in *Table 7*.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 7 Streamline C: constituent tests of the standard staging pathway

Test	Number of tests, *n* (%)[a]
Chest, abdomen and pelvis CT	243 (81)
Pelvis/rectum MRI[b]	120 (40)
Chest CT	44 (15)
Abdomen and pelvis CT	27 (9)
Liver MRI	35 (12)
Colonography CT	10 (3)
PET-CT	43 (14)
Radiography[c]	9 (3)
US[d]	12 (4)
Bone scan	2 (1)
Liver CT	3 (1)
Abdomen and pelvis MRI	4 (1)
Rectal US	4 (1)
Other[e]	2 (1)

a Participants may undergo more than one imaging test.
b One participant with two pelvis/rectum MRI.
c Two participants had two types of radiography (abdominal and chest).
d One participant had two types of US (abdominal and liver).
e Other tests include sigmoidoscopy and non-specified CT.

Reproduced from Taylor *et al.*[25] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

The additional tests the MDT stated would have been generated by the WB-MRI is shown in *Table 8*.

The number of short-interval follow-up scans generated for equivocal findings by the alternative staging pathways is shown in *Appendix 14*, *Table 52*. Overall, the standard imaging pathway generated 13 short-interval follow-up tests and the WB-MRI staging pathway generated 16 short-interval follow-up tests.

Pathway results and outcomes

Primary outcome

The per-participant sensitivity and specificity for metastatic disease according to staging pathway is shown in *Table 9*. The sensitivity analysis treating all equivocal results as either all positive or all negative is also shown.

Overall, there was no significant difference in sensitivity for participants with metastasis between the standard staging pathway [63% (95% CI 51% to 74%)] and the WB-MRI staging pathway [67% (95% CI 56% to 78%)], a difference in sensitivity of 4% (95% CI –5% to 13%; $p = 0.508$).

TABLE 8 Streamline C: MDT-stated additional staging tests generated by WB-MRI

Test	Number of tests, *n* (%)[a]
Chest, abdomen and pelvis CT	5 (2)
Pelvis/rectum MRI	6 (2)
Chest CT	1 (1)
Abdomen and pelvis CT	0 (0)
Liver MRI	0 (0)
Colonography CT	0 (0)
PET-CT	0 (0)
Radiography	0 (0)
US	2 (1)
Bone scan	1 (1)
Liver CT	1 (1)
Abdomen and pelvis MRI	0 (0)
Rectal US	0 (0)
US-guided biopsy	1 (1)
Other[b]	4 (1)

a Participants may undergo more than one imaging test.
b Four other tests include one spine MRI and three brain MRIs (two in same participant).
Reproduced from Taylor *et al.*[25] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

TABLE 9 Streamline C: per-participant sensitivity and specificity for metastatic disease – standard staging pathway vs. WB-MRI staging pathway against the consensus reference standard

	Sensitivity, % (95% CI; *p*-value)				Specificity, % (95% CI; *p*-value)			
Outcome	Number with metastatic disease[a]	WB-MRI staging pathway[b]	Standard staging pathway	Difference: WB-MRI[b] vs. standard	Number without metastatic disease[a]	WB-MRI staging pathway[b]	Standard staging pathway	Difference: WB-MRI[b] vs. standard
Diagnostic accuracy[c]	68	67 (56 to 78)	63 (51 to 74)	4 (–5 to 13; 0.508)	231	95 (92 to 97)	93 (90 to 96)	2 (–2 to 6; 0.481)
Equivocal lesions considered positive[d]	68	71 (59 to 80)	68 (56 to 78)	3 (–6 to 12; 0.727)	231	95 (91 to 97)	92 (88 to 95)	3 (–2 to 7; 0.286)
Equivocal lesions considered negative[d]	68	65 (53 to 75)	58 (46 to 68)	7 (–2 to 17; 0.227)	231	98 (94 to 99)	98 (95 to 99)	0 (–3 to 2; > 0.999)

a Participants by consensus reference standard.
b WB-MRI plus additional generated tests.
c Equivocal results considered positive for colonic tumours and negative for rectal tumours.
d Sensitivity analysis for imaging findings reported as equivocal.
Reproduced from Taylor *et al.*[25] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Secondary outcomes

There was also no significant difference in specificity between the standard pathway and WB-MRI pathway [93% (95% CI 90% to 96%) vs. 95% (95% CI 92% to 97%), respectively], a difference of 2% (95% CI –2% to 6%). As would be expected, the sensitivity of both pathways increased if all equivocal findings were assumed positive and decreased if all equivocal findings were considered negative, but there remained no significant difference in sensitivity or specificity between the pathways (see *Table 9*).

Results in context

To place the results in context, in a hypothetical population of 1000 patients with newly diagnosed colon cancer with a prevalence of 23% for metastatic disease, the number of patients with metastases correctly identified with metastatic disease does not differ between the WB-MRI pathway and the standard staging pathway. Of the 227 patients with metastatic disease, the results suggest that 10 more patients would be identified with disease using the WB-MRI pathway (with 73 missed) compared with the standard pathway (with 84 missed). However, this number could vary from 14 fewer patients to 36 more patients identified with the WB-MRI pathway. On average, 154 and 144 patients with metastatic disease would be identified correctly using the WB-MRI pathway and the standard pathway, respectively.

The number of patients without metastases diagnosed correctly is also not different in patients staged using the WB-MRI pathway and standard pathway. In 773 patients without metastatic disease, 14 more patients would be identified without metastatic disease using the WB-MRI pathway than using the standard pathway, but this number could vary from 18 fewer to 44 more patients. On average, 736 and 722 patients without metastatic disease would be identified correctly using the WB-MRI pathway and the standard pathway, respectively, with metastatic spread overdiagnosed in an average of 37 and 50 patients, respectively.

Figure 4 presents these data graphically.

The per-participant sensitivity and specificity for metastatic disease according to staging pathway and size of the largest metastatic deposit is shown in *Table 10*.

Sensitivity of both pathways was much higher when the largest metastasis was $\geq$ 1 cm than when it was < 1 cm. There was no significant difference between pathways when analysed according to the size of the largest metastatic deposit.

The per-participant sensitivity and specificity for metastatic disease in individual organ sites according to staging pathway is shown in *Appendix 15*, *Table 53*. Overall, there were no significant differences between pathways according to the site of metastatic deposit. Sensitivity in the liver was almost identical [72% (95% CI 57% to 83%) for standard imaging and 74% (95% CI 60 %to 84%) for the WB-MRI staging pathway]. The standard staging pathway detected 65% (95% CI 43% to 82%) of participants with lung metastasis compared with 55% (95% CI 34% to 74%) for the WB-MRI staging pathway.

Local staging

The agreement of the staging pathways for local T stage based on the consensus reference standard is shown in *Table 11* and in participants with histological confirmation in *Appendix 16*, *Table 54.*

Overall, the WB-MRI staging pathway had 54% agreement for T stage compared with 60% for the standard pathway, a non-significant difference of 6% (95% CI 0% to 12%). However, the differences were statistically significant for both T3 and T4 stage, with the WB-MRI pathway achieving lower agreement. When restricted to participants with histological proof of T stage, there was no overall significant difference between the standard pathway and the WB-MRI pathway (51% vs. 47%), a difference of 4% (95% CI –3% to 10%) (see *Appendix 16*, *Table 54*). However, agreement for

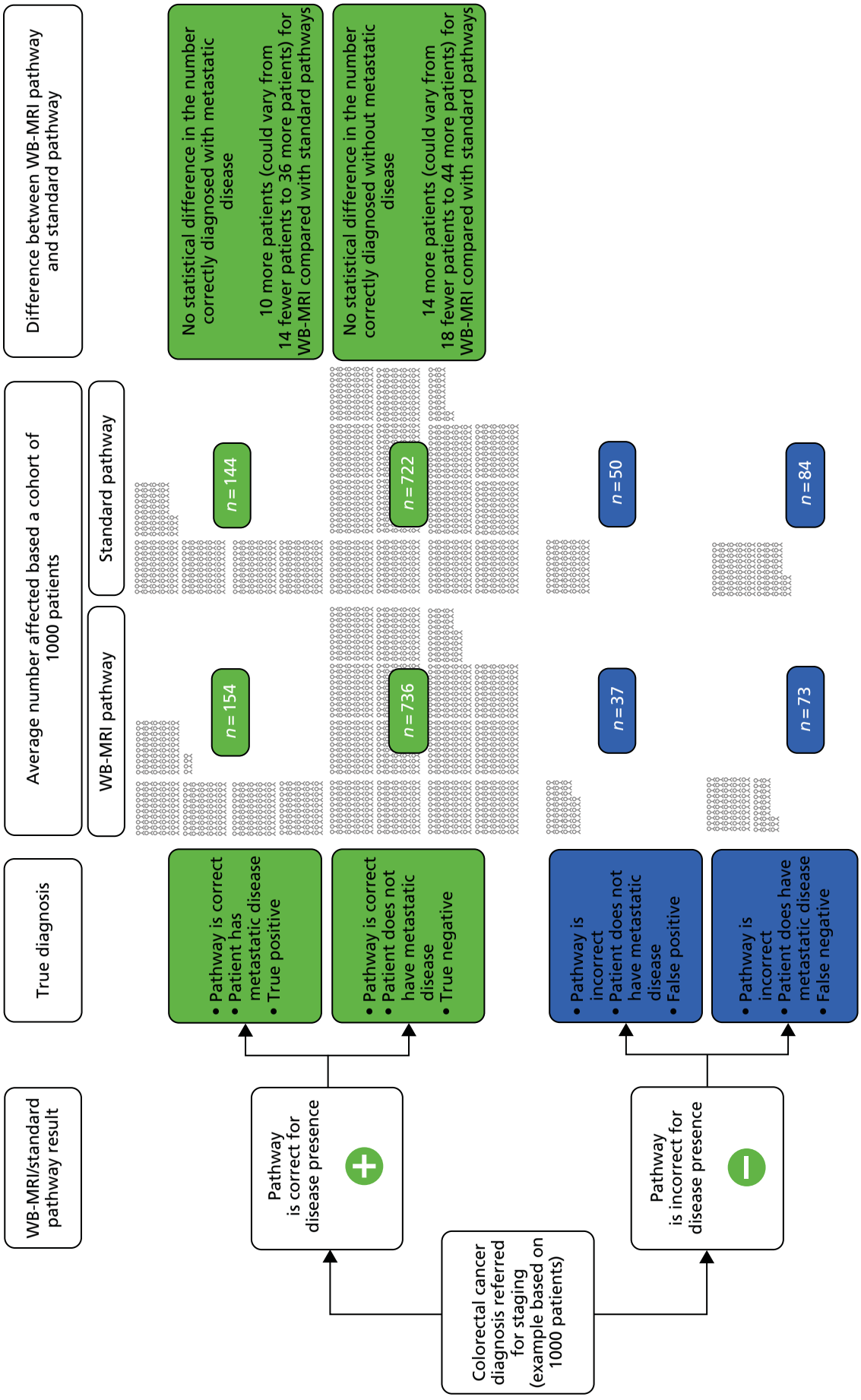


FIGURE 4 Potential impact of staging colorectal cancer with WB-MRI staging pathway or standard staging pathway in a theoretical 1000-patient cohort.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 10 Streamline C: per-participant sensitivity and specificity for metastatic disease according to the size of metastasis – standard staging pathway vs. WB-MRI staging pathway against the consensus reference standard

	Sensitivity, % (95% CI; *p*-value)				Specificity, % (95% CI; *p*-value)			
Maximum metastatic deposit size	Number with metastatic disease[a]	WB-MRI staging pathway[b]	Standard staging pathway	Difference: WB-MRI[b] vs. standard	Number without metastatic disease[a]	WB-MRI staging pathway[b]	Standard staging pathway	Difference: WB-MRI[b] vs. standard
≥ 1cm	45	86 (74 to 94)	82 (69 to 91)	4 (–8 to 17; 0.688)	231	95 (92 to 97)	93 (90 to 96)	2 (–2 to 6; 0.481)
< 1cm	20	35 (18 to 57)	30 (15 to 52)	5 (–13 to 23; > 0.999)				

a Participants by consensus reference standard. Three participants had missing metastatic disease size.
b WB-MRI plus additional generated tests.
Reproduced from Taylor *et al.*[25] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

TABLE 11 Streamline C: per-participant agreement for tumour T stage – standard staging pathway vs. WB-MRI staging pathway against the consensus reference standard

T stage	Number of participants[a]	Staging pathway, *n* (%)		Difference: WB-MRI[b] vs. standard, *n* (95% CI)
		WB-MRI[b,c]	Standard[c]	
T1	11	3 (27)	3 (27)	0 (–19 to 19)
T2	52	37 (71)	29 (56)	15 (2 to 29)
T3	174	100 (58)	120 (69)	–11 (–20 to –3)
T4	58	19 (33)	25 (43)	–10 (–20 to –1)
Total	295	159 (54)	177 (60)	–6 (–12 to 0)

a Four participants had missing data.
b WB-MRI plus additional generated tests.
c 11 and six participants staged at Tx by WB-MRI staging pathway and standard staging pathway, respectively.
Reproduced from Taylor *et al.*[25] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

histologically confirmed T3 stage was significantly lower for the WB-MRI staging pathway (52%) than for the standard staging pathway (62%), a difference of 10% (95% CI 1% to 20%). Conversely, the WB-MRI pathway had higher agreement for T2 stage than standard staging [69% vs 52%, a difference of 17% (95% CI 3% to 31%)].

The agreement of the staging pathways for local N stage based on the consensus reference standard is shown in *Table 12,* and in participants with histological confirmation in *Appendix 17*, *Table 55*. Sensitivity to detect nodal disease was 76% (95% CI 69% to 82%) for the WB-MRI staging pathway and 76% (95% CI 69% to 82%) for the standard staging pathway, not significantly different. Specificity was 60% and 61%, respectively, also not significantly different. Overall, there were no significant differences between N staging across the whole cohort against the consensus reference standard or when the analysis was restricted to participants with histological confirmation.

TABLE 12 Streamline C: per-participant agreement for N stage – standard staging pathway vs. WB-MRI staging pathway against the consensus reference standard

N stage	Number of participants[a]	Staging pathway, *n* (%)		Difference: WB-MRI[b] vs. standard, *n* (95% CI)
		WB-MRI[b]	Standard	
N0	132	80 (60)	81 (61)	–1 (–9 to 7)
N1	88	45 (51)	45 (51)	0 (–10 to 10)
N2	74	46 (62)	40 (54)	8 (–3 to 19)
Total	294	171 (58)	166 (56)	2 (–4 to 7)

a Five participants were missing data.
b WB-MRI plus additional generated tests.
Reproduced from Taylor *et al.*[25] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Impact of staging pathways on primary treatment decision

The groupings of treatment decisions for analysis are shown in *Table 3*. The agreement between the primary treatment decision based on the two staging pathways and the final MDT decision is shown in *Table 13*, and agreement compared with the retrospective optimal treatment decision made by the consensus reference standard panel is shown in *Table 14*.

TABLE 13 Streamline C: primary treatment decision according to the WB-MRI staging pathway and standard staging pathway compared with the final MDT treatment decision

Participant group	Total, *n*[b]	Staging pathway, *n* (%)				Difference in agreement: WB-MRI[b] vs. standard, % (95% CI)
		WB-MRI[a]		Standard		
		Agreement	Disagreement	Agreement	Disagreement	
Colorectal cancer						
All participants	296	284 (96)	12 (4)	282 (95)	14 (5)	1 (–2 to 4)
Colon cancer						
All participants	168	166 (99)	2 (1)	165 (98)	3 (2)	1 (–3 to 4)
Participants with metastatic disease	33	33 (100)	0 (0)	32 (97)	1 (3)	3 (–6 to 12)
Participants without metastatic disease	135	133 (99)	2 (1)	133 (99)	2 (1)	0 (–4 to 4)
Rectal cancer						
All participants	128	118 (92)	10 (8)	117 (91)	11 (9)	1 (–5 to 7)
Participants with metastatic disease	32	28 (88)	4 (12)	28 (88)	4 (12)	0 (–10 to 10)
Participants without metastatic disease	96	90 (94)	6 (6)	89 (93)	7 (7)	1 (–7 to 9)

a Three participants were missing at least one type of participant management decision.
b WB-MRI plus additional generated tests.
Reproduced from Taylor *et al.*[25] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 14 Streamline C: primary treatment decision according to the WB-MRI staging pathway and standard staging pathway compared with the retrospective optimal consensus panel treatment decision

Participant group	Total *n*[b]	Staging pathway, *n* (%)				Difference in agreement: WB-MRI[b] vs. standard, % (95% CI)
		WB-MRI[a]		Standard		
		Agreement	Disagreement	Agreement	Disagreement	
Colorectal cancer						
All participants	296	201 (68)	95 (32)	201 (68)	95 (32)	0 (–3 to 3)
Colon cancer						
All participants	168	120 (71)	48 (29)	119 (70)	49 (30)	1 (–2 to 3)
Participants with metastatic disease	33	18 (55)	15 (45)	17 (52)	16 (48)	3 (–5 to 11)
Participants without metastatic disease	135	102 (76)	33 (24)	102 (76)	33 (24)	0 (–3 to 3)
Rectal cancer						
All participants	128	81 (63)	47 (37)	82 (64)	46 (36)	–1 (–6 to 5)
Participants with metastatic disease	32	11 (34)	21 (66)	11 (34)	21 (66)	0 (–7 to 7)
Participants without metastatic disease	96	70 (73)	26 (27)	71 (74)	25 (26)	–1 (–8 to 6)

a WB-MRI plus additional generated tests.
b Three participants missing at least one type of participant management decision.
Reproduced from Taylor *et al.*[25] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

There was no significant difference between the WB-MRI and standard pathways in terms of primary treatment decision compared with the MDT final decision or with the retrospective consensus panel decision for either rectal or non-rectal colon cancer. For both pathways, agreement with the optimal retrospective treatment decision (which considered 12 months of participant follow-up) was lower than with the contemporaneous final MDT treatment decision. A more detailed breakdown of agreement between staging pathways and the retrospective consensus panel decision for rectal and non-rectal colon cancer for specific treatment decision categories is shown in *Appendices 18–21*, *Tables 56–59*.

Time to complete staging

The time taken to complete staging for the standard pathway and that modelled for the WB-MRI pathway is shown in *Table 15 and Appendix 22*, *Table 60*.

Overall, the WB-MRI staging pathway was significantly shorter than the standard staging pathway [by 5 days (95% CI 3 to 7 days)], and in participants with [10 days (95% CI 3 to 17 days)] and without [4 days (95% CI 2 to 7 days)] metastatic disease. For the WB-MRI staging pathway the interquartile range was 4–13 days, compared with 8–25 days for the standard staging pathway.

TABLE 15 Streamline C: time to complete staging according to treatment pathway

Participant group	Staging pathway, median days (95% CI)		Difference: WB-MRI[a] vs. standard (95% CI)
	WB-MRI[a]	Standard	
All participants	8 (6 to 9)	13 (11 to 15)	–5 (–7 to –3)
Participants with metastatic disease	8 (6 to 11)	18 (12 to 25)	–10 (–17 to –3)
Participants without metastatic disease	8 (6 to 9)	12 (11 to 14)	–4 (–7 to –2)

a WB-MRI plus additional generated tests.

Reproduced from Taylor *et al.*[25] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Number of tests in staging pathways

The median number of imaging or endoscopic tests in the staging pathways until the MDT was able to make the first major treatment decision is shown in *Table 16*. Across the whole cohort, the standard pathway utilised 558 tests (see *Table 7*) and the WB-MRI pathway 320 tests (including the 21 additional tests generated by the 299 WB-MRI scans; see *Table 8*). WB-MRI pathways required significantly fewer tests that standard pathways (median 1 [1 to 1]).

WB-MRI as a stand-alone investigation

The diagnostic accuracy of WB-MRI as a stand-alone investigation (based on the original radiologist report) is shown in *Appendix 23*, *Table 61*.

As a stand-alone investigation, WB-MRI had a sensitivity of 70% (95% CI 59% to 80%) for participants with metastasis, compared with 63% (95% CI 51% to 74%) for standard pathways, a difference of 7% (95% CI –3% to 18%), which was not statistically significant.

The per-participant sensitivity and specificity for metastatic disease in individual organ sites for WB-MRI as a stand-alone investigation is shown in *Appendix 24*, *Table 62*. There was no significant individual organ difference between pathways.

TABLE 16 Streamline C: total number of imaging/endoscopic tests before the first major treatment decision according to staging pathway

Participant group	Staging pathway, median number of tests (95% CI)		Difference: WB-MRI[a] vs. standard (95% CI)
	WB-MRI[a]	Standard	
All participants	1 (1 to 1)	2 (2 to 2)	–1 (–1 to –1)
Participants with metastatic disease	1 (1 to 1)	2 (2 to 3)	–1 (–2 to –1)
Participants without metastatic disease	1 (1 to 1)	2 (2 to 2)	–1 (–1 to –1)

a WB-MRI plus additional generated tests.

Reproduced from Taylor *et al.*[25] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

The agreement of WB-MRI alone for local T stage based on the consensus reference standard is shown in *Appendix 25*, *Table 63* and in participants with histological confirmation in *Appendix 26*, *Table 64*.

Overall, and against the consensus, WB-MRI alone had lower agreement than the standard pathway for T stage (46% vs. 64%, respectively), a difference of 18% (95% CI 10% to 25%). The differences were also statistically significant for T3-stage tumours. When restricted to participants with histological proof of T stage, there remained an overall significant difference between WB-MRI alone and the standard pathway (40% vs. 56%, respectively), a difference of 16% (95% CI 7% to 25%) (see *Appendix 26*, *Table 64*).

The agreement of WB-MRI alone for local N stage based on the consensus reference standard is shown in *Appendix 27*, *Table 65* and in participants with histological confirmation in *Appendix 28*, *Table 66*.

Overall, WB-MRI alone had significantly less agreement for overall N stage than the standard pathway (45% vs. 55%, respectively), a difference of 10% (95% CI 4% to 17%). This was also the case when the analysis was restricted to participants with histological proof of N stage (41% vs. 51%, respectively), a difference of 10% (95% CI 3% to 17%) (see *Appendix 28*, *Table 66*). WB-MRI had 64% (95% CI 56% to 71%) sensitivity for nodal disease, compared with 75% (95% CI 68% to 81%) for the standard pathway, a significant difference of 11% (95% CI –19% to –3%). Specificity of WB-MRI alone was also significantly lower than the standard pathway [47% (95% CI 39% to 56%) vs. 60% (95% CI 51% to 69%), respectively, a difference of 13% (95% CI –22% to –4%)].

The per-participant sensitivity and specificity for metastatic disease of WB-MRI alone compared with standard chest, abdomen and pelvis CT is shown in *Appendix 29*, *Table 67*. WB-MRI had a significantly higher sensitivity than the standard pathway [76% (95% CI 63% to 86%) vs. 50% (95% CI 37% to 63%), respectively, a difference of 26% (95% CI 13 to 39)], but lower specificity [86% (95% CI 81% to 91%) vs. 99% (95% CI 97% to 100%), respectively, a difference of 13% (95% CI 9% to 18%)].

WB-MRI as an additional investigation

The diagnostic accuracy of WB-MRI when added as an additional test to the standard staging pathway is shown in *Appendix 30*, *Table 68*.

When WB-MRI was added to the standard staging pathway, sensitivity increased from 63% (95% CI 51% to 73%) to 71% (95% CI 59% to 80%), a difference of 7% (95% CI 0% to 14%). Specificity was almost identical [94% (95% CI 90% to 96%) for standard staging and 95% (95% CI 91% to 97%) for standard staging plus WB-MRI].

Perceptual errors

The number of participants with metastases for whom radiologists made at least one perceptual error (i.e. seen in retrospect by the consensus panel) according to the staging pathways is shown in *Appendix 31*, *Table 69* and per organ in *Appendix 32*, *Table 70*.

The number of participants with at least one perceptual error was very similar for the WB-MRI staging pathway (16 participants), standard staging pathway (15 participants) and WB-MRI as a stand-alone investigation (16 participants). In many of these participants at least one metastatic deposit was also correctly detected. Overall, for the primary outcome (i.e. per-participant detection of metastasis) there were three perceptual errors in the WB-MRI pathway and six in the standard pathway.

WB-MRI quality assurance: Streamline C

The quality assurance scoring system is described in *Chapter 2* (see *Appendix 2*, *Table 41*). The summary WB-MRI quality assurance scores for Streamline C participants is shown in *Appendix 33*, *Table 71*, including Streamline L scores for comparison. A more detailed breakdown of scores for those scans judged as suboptimal or degraded is shown in *Appendix 34*, *Table 72*, *and Appendix 35*, *Table 73*. This includes some participants ultimately excluded from the final trial cohorts. Overall, 36 Streamline C WB-MRI scans were scored as degraded (including those of two participants ultimately withdrawn from the trial). Of the remaining 34 WB-MRI scans, just four attracted a technical quality score of 1 (more than one sequence with substantial degradation of the images). Of the 17 WB-MRI scans attracting an anatomical coverage score of 1, 10 received this score because gadolinium was not administered.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Chapter 4 Results: Streamline L

Parts of this chapter are reproduced from Taylor *et al.*[26] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Participants

Recruitment commenced in March 2013 and was completed in August 2016. Overall, 976 patients were screened for eligibility, of whom 623 were excluded (*Figure 5* and *Appendix 36*, *Table 74*). Of the 353 participants entering the trial, 166 were withdrawn, mainly owing to a final diagnosis other than

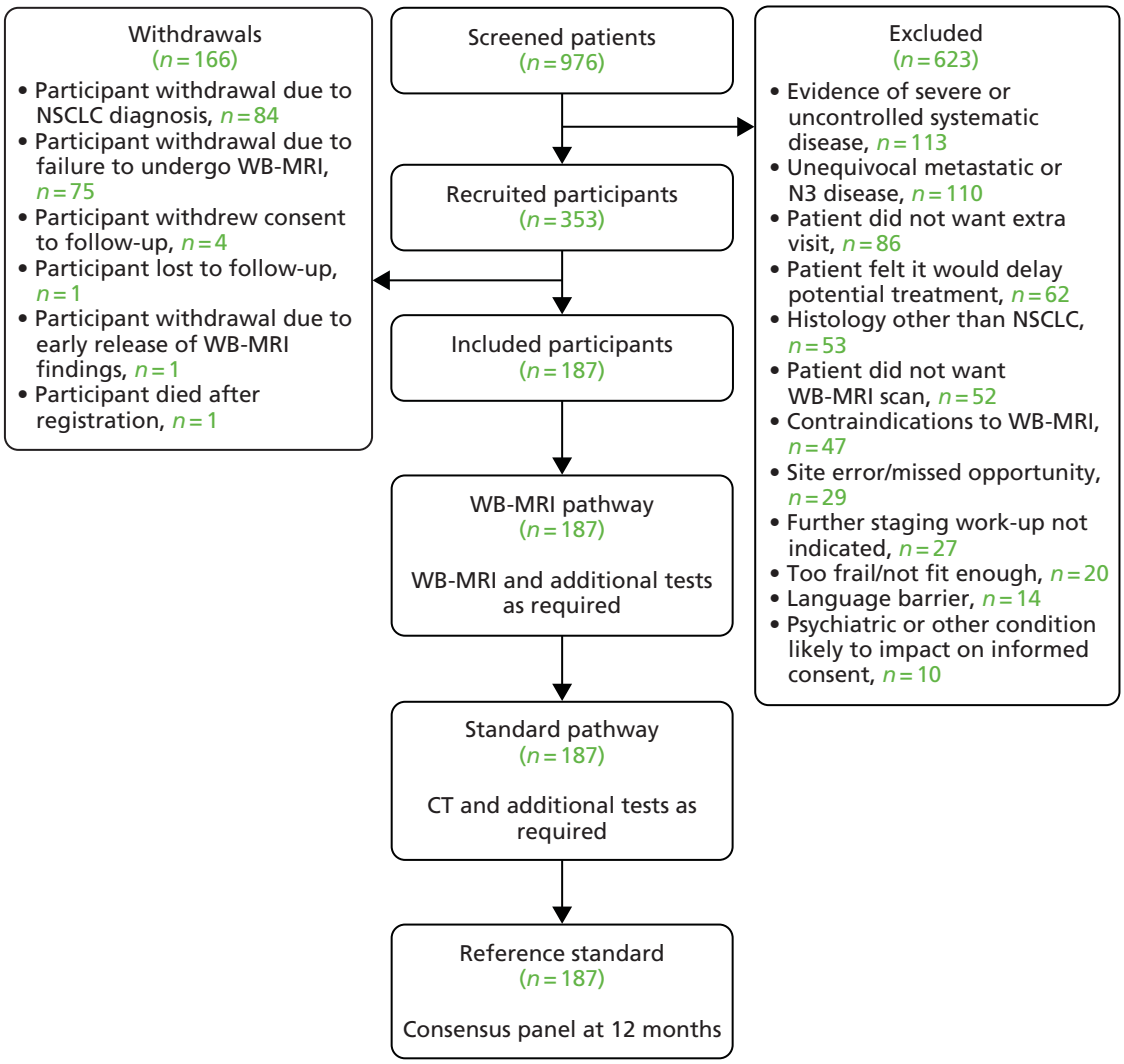


FIGURE 5 Streamline L: participant flow diagram. NSCLC, non-small-cell lung cancer. Reproduced from Taylor *et al.*[26] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

non-small-cell lung cancer ($n = 84$) or failure to undergo MRI, usually owing to scheduling issues ($n = 75$) (see *Appendix 37*, *Table 75*). The final cohort consisted of 187 participants, including 70 women (37%) (*Table 17* and see *Figure 5*). There were no reported series adverse reactions.

Overall, 31 (17%) participants died within the 12-month follow-up period with enough collected data for the consensus panels to be confident in assigning a TNM stage.

The status of the primary tumour and follow-up imaging available to the panels is shown in *Appendix 38*, *Table 76*.

TABLE 17 Demographics of the final study: Streamline L cohort

Characteristic	*n* (%)
Sex	
Male	117 (63)
Female	70 (37)
Age (years)	
Mean (SD)	68 (10)
Range	37 to 96
Performance status	
Fully active	86 (46)
Ambulatory (able to work)	75 (40)
Ambulatory (not able to work)	8 (4)
Not recorded	18 (10)
Tumour location[a]	
Right upper lobe	73 (39)
Right middle lobe	14 (7)
Right lower lobe	24 (13)
Left upper lobe (including lingula)	54 (29)
Left lower lobe	28 (15)
Histological subtype	
Adenocarcinoma	115 (62)
Large cell	4 (2)
Squamous	42 (22)
Adenosquamous	1 (1)
Other	13 (7)
No histology (clinical diagnosis of NSCLC)/missing	12 (6)

NSCLC, non-small-cell lung cancer; SD, standard deviation.
a By consensus reference standard. Some participants have multiple tumour locations.
Reproduced from Taylor *et al.*[26] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Of the 187 participants in the cohort, as defined by the consensus panel, 137 participants (73%) were stage T2a or above and 77 participants (41%) were node positive (see *Appendix 39*, *Table 77*). The detailed sites of nodal involvement are shown in *Appendix 40*, *Table 78*.

Overall, 52 participants (28%) had metastatic disease [8 M1a (7%) and 44 M1b (24%)] and 135 participants did not have metastasis. The organ location of the metastatic disease and the basis for diagnosis is shown in *Table 18* and a more detailed breakdown in *Appendix 41*, *Table 79*. The majority of participants with metastasis had at least liver ($n = 9$), bone ($n = 15$), brain ($n = 14$) and/or lung ($n = 14$) metastasis.

The number of deposits per organ site is shown in *Appendix 42*, *Table 80*. The maximum size of the metastasis at the time of staging is shown in *Appendix 43*, *Table 81*. In eight participants, the metastasis was not visible on any staging investigation, even in retrospect, and only became apparent during the 12-month follow-up period.

A total of seven participants had a second malignancy reported.

Equivocal results

The number of equivocal results according to the two staging pathways and WB-MRI alone is shown in *Appendix 44*, *Table 82*.

On a per-participant basis, one of the 52 participants with metastatic disease had equivocal results based on the WB-MRI pathway, compared with four participants based on the standard pathway and two participants based on WB-MRI alone. Of the 135 participants without metastases, one participant had equivocal results based on the WB-MRI pathway, compared with two participants based on the standard pathway and 17 participants based on WB-MRI alone.

TABLE 18 Streamline L: organ sites of metastatic disease

Site	Number of participants[a]	Histological proof, *n* (%)	If imaging diagnosis if no histology,[b] *n* (%)		
			Characteristic imaging appearances	Growth on follow-up	Response to therapy
Liver	9	2 (22)	6 (86)	4 (57)	1 (14)
Lung	14	0 (0)	14 (100)	11 (79)	3 (21)
Pleura	5	1 (20)	4 (100)	2 (50)	1 (25)
Brain	14	1 (7)	13 (100)	8 (62)	4 (31)
Adrenal	7	0 (0)	6 (86)	5 (71)	1 (14)
Bone	15	1 (7)	14 (100)	7 (50)	2 (14)
Other[c]	9	2 (15)	11 (100)	5 (45)	0 (0)

a Participants may have more than one site of metastatic disease.
b Metastasis may fulfil more than one criterion for imaging diagnosis.
c Nine participants with metastasis in 11 other sites (pancreas, limbs soft tissue, para-aortic node, abdomen/pelvis soft tissue, right thyroid, left axilla, neck/chest soft tissue, right kidney, leptomeningeal, right axillary node and mesenteric node).

Reproduced from Taylor *et al.*[26] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Streamline L: staging pathway tests

The type and number of standard staging tests performed prior to the first major treatment decision across the whole cohort is shown in *Table 19*.

The additional tests the MDT stated would have been generated by the WB-MRI are shown in *Table 20*.

TABLE 19 Streamline L: constituent tests of the standard staging pathway

Test	Number of tests (%)[a]
Neck chest and abdomen CT	3 (2)
Chest CT[b]	8 (4)
Abdomen and pelvis CT	11 (6)
Chest and abdomen CT	5 (3)
Chest abdomen and pelvis CT	2 (1)
PET-CT[c]	181 (97)
Head CT[d]	25 (13)
Head MRI[e]	16 (9)
TBNA	0 (0)
EBUS	6 (3)
EUS	2 (1)
Bone scan	0 (0)
CT-guided biopsy	5 (3)
Liver CT[f]	6 (3)
US-guided biopsy	6 (3)
US[g]	8 (4)
Adrenal MRI	2 (1)
MSK-MRI	2 (1)
Other MRI	1 (1)
Radiography[h]	9 (5)
Other[i]	4 (2)

EUS, endoscopic ultrasonography; MSK-MRI, musculoskeletal magnetic resonance imaging; TBNA, transbronchial needle aspiration.
a Participants may undergo more than one imaging test.
b One participant had two chest CTs.
c Two participants had two PET-CTs.
d One participant had two head CTs.
e One participant had two head MRIs.
f One participant had two liver CTs.
g One participant had two ultrasounds (transvaginal and abdominal and renal tract).
h One participant had three radiographs.
i Other tests include three bronchial washings and one surgical biopsy.

Reproduced from Taylor *et al.*[26] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

TABLE 20 Streamline L: MDT-stated additional staging tests generated by the WB-MRI

Test	Number of tests (%)[a]
Neck chest and abdomen CT	0 (0)
Chest CT	0 (0)
Abdomen and pelvis CT	0 (0)
Chest and abdomen CT	0 (0)
Chest abdomen and pelvis CT	1 (1)
PET-CT	26 (14)
Head CT	0 (0)
Head MRI	6 (3)
TBNA	0 (0)
EBUS/TBNA	4 (2)
EUS	1 (1)
Bone scan	0 (0)
CT-guided biopsy	0 (0)
Liver CT	0 (0)
US-guided biopsy	1 (1)
US	4 (2)
Adrenal MRI	0 (0)
MSK-MRI	0 (0)
Other MRI	2 (1)
Radiography	0 (0)
Other	0 (0)

EUS, endoscopic ultrasonography; MSK-MRI, musculoskeletal magnetic resonance imaging; TBNA, transbronchial needle aspiration.
a Participants may undergo more than one imaging test.
Reproduced from Taylor *et al.*[26] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

The number of short-interval follow-up scans generated for equivocal findings by the alternative staging pathways is shown in *Appendix 45*, *Table 83*. Overall, the standard imaging pathway generated four short-interval follow-up tests and the WB-MRI staging pathway generated two.

Pathway results and outcomes

Primary outcome

The per-participant sensitivity and specificity for metastatic disease according to staging pathway is shown in *Table 21*. The sensitivity analysis treating all equivocal results as positive or negative is also shown.

Overall, there was no significant difference in sensitivity for participants with metastasis between the standard staging pathway [54% (95% CI 41% to 67%)] and the WB-MRI staging pathway [50% (95% CI 37% to 63%)], a difference of 4% (95% CI –7% to 15%; $p = 0.727$).

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 21 Streamline L: per-participant sensitivity and specificity for metastatic disease – standard staging pathway vs. WB-MRI staging pathway against the consensus reference standard

Outcome	Sensitivity, % (95% CI; *p*-value)				Specificity, % (95% CI; *p*-value)			
	Number with metastatic disease[a]	WB-MRI staging pathway[b]	Standard staging pathway	Difference: WB-MRI vs. standard	Number without metastatic disease[a]	WB-MRI staging pathway[b]	Standard staging pathway	Difference: WB-MRI vs. standard pathway
Diagnostic accuracy[c]	52	50 (37 to 63)	54 (41 to 67)	–4 (–15 to 7; 0.727)	135	93 (88 to 96)	95 (91 to 98)	–2 (–7 to 2; 0.453)
Sensitivity analysis[d]	52	48 (35 to 61)	46 (33 to 59)	2 (–11 to 14; > 0.999)	135	94 (89 to 97)	97 (93 to 99)	–3 (–6 to 1; 0.125)

a Participants by consensus reference standard.
b WB-MRI plus additional generated tests.
c Equivocal lesions considered positive.
d Equivocal lesions considered negative.

Reproduced from Taylor *et al.*[26] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Secondary outcomes

There was no significant difference in specificity between the standard and WB-MRI pathways [95% (95% CI 91% to 98%) vs. 93% [95% CI 88% to 96%], respectively, a difference of 2% (95% CI –2% to 7%)]. As expected, sensitivity of both pathways decreased if equivocal lesions were considered negative, but there remained no significant difference in sensitivity or specificity between the pathways (see *Table 21*).

Results in context

To place the results in context, in a hypothetical population of 1000 patients with newly diagnosed non-small lung cancer with a prevalence of 28% for metastatic disease, the number of patients with metastases diagnosed correctly does not differ between the WB-MRI pathway and the standard staging pathway. Of the 278 patients with metastatic disease, the results suggest that 11 more patients would be identified with metastatic disease using the standard staging pathway (with 128 patients missed) than using the WB-MRI pathway (with 139 patients missed). However, this number could vary from 20 fewer patients to 42 more patients.

The number of patients without metastases diagnosed correctly is also not different in patients staged using the WB-MRI and standard pathways. In 722 patients without metastatic disease, 16 more patients would be identified without metastatic disease using the standard pathway than using the WB-MRI pathway, but this number could vary from 16 fewer patients to 48 more patients. On average, 674 and 690 patients without metastatic disease would be identified correctly using the WB-MRI pathway and standard pathway, respectively, with metastatic spread overdiagnosed in an average of 48 and 32 patients, respectively. *Figure 6* presents these data graphically.

The per-participant sensitivity and specificity for metastatic disease according to staging pathway and size of the largest metastatic deposit is shown in *Table 22*.

Sensitivity of both pathways was much higher when the largest metastasis was ≥ 1 cm than when it was < 1 cm. There was no significant difference between pathways when analysed according to the size of the largest metastatic deposit.

The per-participant sensitivity and specificity for metastatic disease in individual organ sites according to staging pathway is shown in *Appendix 46*, *Table 84*.

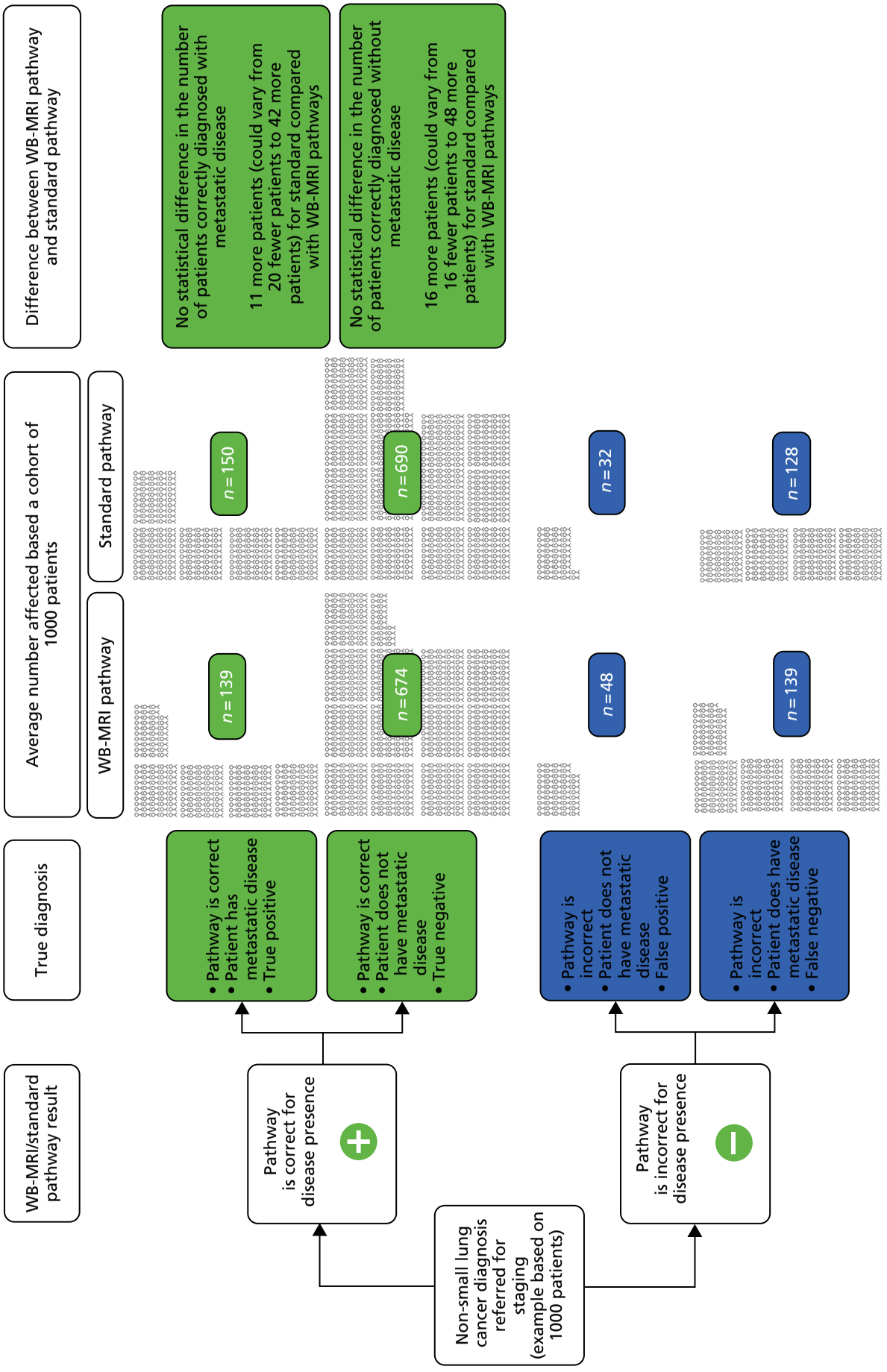


FIGURE 6 Potential impact of staging non-small-cell lung cancer with WB-MRI or standard staging pathway in a theoretical 1000-patient cohort.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 22 Streamline L: per-participant sensitivity and specificity for metastatic disease according to the size of metastasis – standard staging pathway vs. WB-MRI staging pathway against the consensus reference standard

Maximum metastatic deposit size	Number with metastatic disease[a,b]	Sensitivity, % (95% CI; *p*-value)		
		Staging pathway		
		WB-MRI[c]	Standard	Difference: WB-MRI[c] vs. standard
≥ 1cm	22	82 (64 to 92)	75 (57 to 87)	7 (–9 to 23; 0.625)
< 1cm	28	9 (3 to 28)	27 (13 to 48)	–18 (–37 to 1; 0.125)
Number without metastatic disease[a,b]		**Specificity, % (95% CI; *p*-value)**		
		Staging pathway		
		WB-MRI[c]	Standard	Difference: WB-MRI[c] vs. standard
135		93 (88 to 96)	95 (91 to 98)	–2 (–7 to 2; 0.453)

a Participants by consensus reference standard.
b Two participants were missing the largest metastatic disease size.
c WB-MRI plus additional generated tests.
Reproduced from Taylor *et al.*[26] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Overall, there were no significant differences between pathways according to the site of metastasis deposit. Sensitivity in the lung was identical [36% (95% CI 16% to 61%)]. The standard staging pathway detected six (44% [95% CI 21 to 67]) out of 14 participants with brain metastasis compared with eight [57% (95% CI 33% to 79%)] participants with brain metastasis for the WB-MRI staging pathway.

Local staging

The agreement of the staging pathways for local T stage based on the consensus reference standard is shown in *Table 23* and in participants with histological confirmation in *Appendix 47*, *Table 85*.

TABLE 23 Streamline L: per-participant agreement for tumour T stage – standard staging pathway vs. WB-MRI staging pathway against the consensus reference standard

T stage	Number of participants[a]	Staging pathway, *n* (%)		Difference in agreement: WB-MRI[b] vs. standard pathway, % (95% CI; *p*-value)
		WB-MRI[b,c]	Standard[c]	
T1a	29	28 (97)	26 (90)	7 (–9 to 22)
T1b	18	9 (50)	8 (44)	6 (–18 to 29)
T2a	51	20 (39)	23 (45)	–6 (–20 to 8)
T2b	23	8 (35)	9 (39)	–4 (–30 to 21)
T3	33	21 (64)	17 (52)	12 (–9 to 33)
T4	29	13 (45)	17 (59)	–14 (–39 to 11)
Overall T stage	183	99 (54)	100 (55)	–1 (–9 to 8; 1.00)

a Four participants were missing data.
b WB-MRI plus additional generated tests.
c One participant each staged Tx by WB-MRI and standard staging pathways.
Reproduced from Taylor *et al.*[26] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Overall, and against the consensus, there was no significant difference in T stage agreement between pathways. When restricted to participants with histological proof of T stage, there remained no significant difference between the two pathways.

The sensitivity of the agreement pathways for local N stage based on the consensus reference standard is shown in *Table 24* and in participants with histological confirmation in *Appendix 48*, *Table 86*. A more detailed nodal-site breakdown is shown in *Appendix 49*, *Table 87*.

Overall, the WB-MRI staging pathway had lower agreement for N stage (65%) than the standard staging pathway (75%), a difference of 10% (95% CI 3% to 18%). When restricted to participants with histological proof of N stage, there remained a statistically significant 10% (95% CI 1% to 19%) difference between the WB-MRI staging pathway and the standard staging pathway (72% vs. 62%, respectively). Sensitivity to detect nodal disease was 71% (95% CI 61% to 80%) for WB-MRI staging pathway and 84% (95% CI 75% to 91%) for the standard staging pathway, a significant difference of 13% (95% CI 3% to 22%). Specificity was 80% (95% CI 72% to 87%) and 86% (95% CI 78% to 91%), respectively, not significantly different.

Impact of staging pathways on primary treatment decision

The groupings of treatment decisions for analysis are shown in *Table 4*. The agreement between the primary treatment decision based on the two staging pathways and the final MDT decision is shown in *Table 25*, and agreement compared with the retrospective optimal treatment decision made by the consensus reference standard panel is shown in *Table 26*.

There was no significant difference between the WB-MRI pathway and standard pathway in terms of primary treatment decision compared with either the MDT final decision or retrospective consensus panel decision. For both pathways, agreement with the optimal retrospective treatment decision was lower than with the contemporaneous final MDT treatment decision. A more detailed breakdown of agreement between staging pathways is shown in *Appendix 50*, *Table 88*, and agreement with the retrospective consensus panel decision in *Appendix 51*, *Table 89*.

TABLE 24 Streamline L: per-participant agreement for N stage – standard staging pathway vs. WB-MRI staging pathway against the consensus reference standard

N stage	Number of participants[a]	Staging pathway, *n* (%)		Difference in agreement: WB-MRI[b] vs. standard pathway, % (95% CI)
		WB-MRI[b]	Standard[c]	
N0	107	86 (80)	92 (86)	–6 (–14 to 3)
N1	21	3 (14)	8 (38)	–24 (–49 to 2)
N2	36	20 (56)	26 (73)	–17 (–39 to 5)
N3	20	10 (50)	12 (60)	–10 (–34 to 14)
Overall N stage	184	119 (65)	138 (75)	–10 (–18 to –3)

a Three participants were missing data.
b WB-MRI plus additional generated tests.
c One participant staged Nx by standard staging pathway.

Reproduced from Taylor *et al.*[26] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 25 Streamline L: primary treatment decision according to the WB-MRI staging pathway and standard staging pathway compared with the final MDT treatment decision

		Staging pathway, *n* (%)				Difference in agreement: WB-MRI[b] vs. standard pathway, % (95% CI)
		WB-MRI[a]		Standard		
Participant group	Total, *n*[b]	Agreement	Disagreement	Agreement	Disagreement	
All participants	183	180 (98)	3 (2)	181 (99)	2 (1)	–1 (–4 to 2)
Participants with metastatic disease	52	51 (98)	1 (2)	50 (96)	2 (4)	2 (–7 to 11)
Participants without metastatic disease	131	129 (98)	2 (2)	131 (100)	0 (0)	–2 (–4 to 1)

a WB-MRI plus additional generated tests.
b Four participants were missing at least one type of participant treatment decision.
Reproduced from Taylor *et al.*[26] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

TABLE 26 Streamline L: primary treatment decision according to the WB-MRI staging pathway and standard staging pathway compared with the retrospective optimal consensus panel treatment decision

		Staging pathway, *n* (%)				Difference in agreement: WB-MRI[b] vs. standard pathway, % (95% CI)
		WB-MRI[a]		Standard		
Participant group	Total, *n*[b]	Agreement	Disagreement	Agreement	Disagreement	
All participants	183	152 (83)	31 (17)	151 (82)	32 (17)	1 (–2 to 3)
Participants with metastatic disease	52	29 (56)	23 (44)	28 (54)	24 (46)	2 (–5 to 9)
Participants without metastatic disease	131	123 (94)	8 (6)	123 (94)	8 (6)	0 (–4 to 4)

a WB-MRI plus additional generated tests.
b Four participants were missing at least one type of participant management decision.
Reproduced from Taylor *et al.*[26] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Time to complete staging

The time taken to complete staging for the standard pathway and that modelled for the WB-MRI pathway is show in *Table 27* and in *Appendix 52*, *Table 90*.

The time taken to stage the participant was significantly less with the WB-MRI pathway than with the standard pathway, with a median difference of 6 days (95% CI 4 to 8 days), increasing to 7 days (95% CI 3 to 11 days) in those with metastatic disease. For the WB-MRI staging pathway the interquartile range was 7–19 days, compared with 13–31 days for the standard staging pathway.

TABLE 27 Streamline L: time to complete staging according to treatment pathway

Participant group	Staging pathway, days (95% CI)		Difference: WB-MRI vs. standard pathway, days (95% CI)
	WB-MRI	Standard	
All participants	13 (12 to 14)	19 (17 to 21)	–6 (–8 to –4)
Participants with metastatic disease	13 (11 to 15)	20 (16 to 23)	–7 (–11 to –3)
Participants without metastatic disease	13 (11 to 14)	19 (16 to 21)	–6 (–9 to –3)

Reproduced from Taylor *et al.*[26] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Number of tests in staging pathways

The median number of imaging or endoscopic tests in the staging pathways prior to the first major treatment decision is shown in *Table 28*. Across the whole cohort the standard pathway utilised 302 tests (see *Table 19*) and the WB-MRI pathway utilised 232 tests (including the 45 additional tests generated by the 187 WB-MRI scans; see *Table 20*). Overall, however, there was no significant difference in the number of tests in each pathway.

WB-MRI as a stand-alone investigation

The diagnostic accuracy of WB-MRI as a stand-alone investigation (based on the original radiologist report) is shown in *Appendix 53*, *Table 91*.

As a stand-alone investigation, WB-MRI had a sensitivity of 50% (95% CI 37% to 63%) for participants with metastasis, compared with 54% (95% CI 41% to 67%) for standard pathways, a difference of 4% (95% CI –16% to 8%), which was not statistically significant. However, WB-MRI had lower specificity compared with the standard pathway [85% (95% CI 78% to 90%) vs. 95% (95% CI 91% to 98%), a difference of –10% (95% CI –17% to –4%)].

The per-participant sensitivity and specificity for metastatic disease in individual organ sites for WB-MRI as a stand-alone investigation is shown in *Appendix 54*, *Table 92*. There was no significant individual organ differences between pathways.

TABLE 28 Streamline L: total number of imaging/endoscopic tests before the first major treatment decision according to staging pathway

Participant group	Staging pathway, number of tests (95% CI)		Difference: WB-MRI vs. standard pathway, number of tests (95% CI)
	WB-MRI	Standard	
All participants	1 (1 to 1)	1 (1 to 2)	0 (–1 to 0)
Participants with metastatic disease	1 (1 to 1)	2 (1 to 2)	–1 (–1 to 0)
Participants without metastatic disease	1 (1 to 1)	1 (1 to 1)	0 (0 to 0)

Reproduced from Taylor *et al.*[26] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

The agreement of WB-MRI alone for local T stage based on the consensus reference standard is shown in *Appendix 55*, *Table 93* and in participants with histological confirmation in *Appendix 56*, *Table 94*. There was no significant difference between WB-MRI alone and standard staging pathways for T stage agreement against the consensus reference standard [9% difference (95% CI 0% to 17%)]. However, in those with histological proof, standard pathways had significantly higher agreement than WB-MRI alone [55% vs. 42%, respectively, a 13% (95% CI 3% to 23%) difference].

The agreement of WB-MRI alone for local N stage based on the consensus reference standard is shown in *Appendix 57*, *Table 95* and in participants with histological confirmation in *Appendix 58*, *Table 96*. WB-MRI alone had significantly lower agreement for overall N stage compared with standard pathways [60% vs. 74%, respectively, a difference of –14% (95% CI –22% to –6%)]. However, when the analysis was restricted to participants with histological proof, there was a non-significant difference of 10% (95% CI 0% to 20%) (see *Appendix 58*, *Table 96*). The sensitivity of WB-MRI alone for nodal disease was 67% (95% CI 56% to 77%), which was significantly lower than for standard pathways [84% (95% CI 75% to 91%), a difference of –17% (95% CI –29% to –5%)]. Specificity was 76% (95% CI 67% to 83%) for WB-MRI alone and 85% (95% CI 76% to 90%) for standard pathways, a difference of –9% (95% CI –17% to 0%).

The per-participant sensitivity and specificity for metastatic disease of WB-MRI alone compared with standard PET-CT is shown in *Appendix 59*, *Table 97*. There was no significant difference in sensitivity between WB-MRI and PET-CT [47% (95% CI 33% to 61%) vs. 53% (95% CI 39% to 67%), respectively]. Specificity was also not significantly different [85% (95% CI 78% to 91%) vs. 90% (95% CI 84% to 94%), respectively].

WB-MRI as an additional investigation

The diagnostic accuracy of WB-MRI when added as an additional test to the standard staging pathway is shown in *Appendix 60*, *Table 98*. When WB-MRI was added to the standard staging pathway, sensitivity increased from 54% (95% CI 41% to 67%) to 60% (95% CI 46% to 72%), a difference of 6% (95% CI –3 to 14%). Specificity was identical 96% (95% CI 91% to 98%).

Perceptual errors

The number of participants with metastases in whom radiologists made at least one perceptual error (i.e. seen in retrospect by the consensus panel) according to the staging pathways is shown in *Appendix 61*, *Table 99*, and per organ in *Appendix 62*, *Table 100*.

The number of participants with at least one perceptual error was greater for the WB-MRI staging pathway (18 participants) compared with the standard staging pathway (seven participants). For WB-MRI as a stand-alone test there was perpetual errors for 19 participants. In many of these participants at least one metastatic deposit was also correctly detected. Overall, for the primary outcome (i.e. per-participant detection of metastasis) there were seven perceptual errors in the WB-MRI pathway and three in the standard pathway.

WB-MRI quality assurance: Streamline L

The summary WB-MRI quality assurance scores for Streamline L participants are shown in *Appendix 33*, *Table 71*, including Streamline C scores for comparison. A more detailed breakdown of scores for those scans judged as suboptimal or degraded is shown in *Appendix 34*, *Table 72*, and *Appendix 35*, *Table 73*. This includes some participants ultimately excluded from the final trial cohorts. Overall, 33 of Streamline L WB-MRI scans were scored as degraded (including those of 13 participants ultimately withdrawn from the trial). Of the remaining 20 WB-MRI scans, just four attracted a technical quality score of 1 (more than one sequence with substantial degradation of the images). Of 11 scans attracting an anatomical coverage score of 1, seven received this score because gadolinium was not administered.

Chapter 5 Economic analysis of WB-MRI staging pathway versus standard staging pathways in non-small-cell lung cancer and colorectal cancer

Introduction

The aim of this substudy was to examine the cost-effectiveness of WB-MRI compared with standard pathways for staging of non-small-cell lung cancer and colorectal cancer in the UK. There have been no previous studies in this area. Huppertz *et al.*[31] examined the cost of WB-MRI versus a sequential multimodal diagnostic algorithm for staging participants with rectal cancer in Germany and found that substantial savings were achievable with WB-MRI, but this study did not consider cost-effectiveness and was not UK based.

Overview of economic analysis

We undertook two separate cost analyses to compare the costs of WB-MRI versus standard staging pathways for non-small-cell lung cancer and colorectal cancer based on data collected in Streamline L and C, respectively. We did not undertake lifetime horizon cost-effectiveness analysis because of the concordance between WB-MRI and standard staging pathways in informing treatment decisions for both types of cancer (see *Chapter 5, Concordance in treatment decisions in Streamline C and L*). This meant that we only needed to compare costs during the decision timeline, so did not consider costs or effects past that point. The cost analyses took a UK NHS perspective. Costs were calculated in 2016–17 Great British pounds (£) and inflated where necessary. The time horizon was the time from initial diagnosis to treatment decision by the MDT, reflecting the fact that beyond this point costs and outcomes would have been the same for both standard staging and WB-MRI given the concordance in treatment decisions (see *Chapter 5, Concordance in treatment decisions in Streamline C and L*). Given the time horizon for the cost analyses that we have performed, which was < 1 year, discounting was not applied. We completed a Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement – available on the project web page [URL: www.journalslibrary.nihr.ac.uk/programmes/hta/106801 (accessed 12 September 2019)].

Treatment decision-making processes in Streamline

All participants in each of the two Streamline trials followed the standard staging pathway, consisting of a battery of standard tests to stage their disease, according to local interpretations of the standard pathways for the two cancer types as recommended by the NICE and according to their particular presentation. As described in *Chapter 1*, in the case of colorectal cancer, CT of the chest, abdomen and pelvis is recommended, supplemented by pelvic MRI for local staging of rectal cancer. Although the NICE does not currently recommend additional imaging, in routine clinical practice it is not unusual for participants to undergo PET-CT and/or liver MRI if metastatic disease is suspected. Staging pathways in lung cancer are more complex, with CT, PET-CT, MRI, US and endobronchial/percutaneous biopsy all recommended at various points in the staging algorithm. All participants underwent a WB-MRI scan, followed by other indicated tests when requested by the clinician on the basis of the results of the WB-MRI. (Hypothetical) treatment decisions were subsequently made and recorded on the basis of these two staging methods (standard staging pathway and WB-MRI staging pathway; see *Chapter 2*). The two decisions could be the same or different.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

A third treatment decision was then made by the MDT using all of the available evidence from the two staging pathways (i.e. based on all available information from both the standard staging pathway and the WB-MRI staging pathway combined) and participants were treated empirically on the basis of this third decision. The treatment decision made by the MDT could be the same or different from the decisions made on the basis of standard staging alone or WB-MRI staging pathway alone.

As described in *Chapter 2*, participants were followed for a period of 12 months, or until death if sooner, and a consensus panel adjudicated on the TNM stage at diagnosis based on all the available clinical, pathological and imaging data. A fourth theoretical optimal treatment decision was made by the panel on the basis of all the previous evidence, including the three previous decisions plus all of the follow-up data over the previous year. The four decisions made in both of the Streamline trials were therefore as follows.

d1: standard pathway treatment decision, based on standard battery of tests alone
d2: WB-MRI pathway treatment decision, based on WB-MRI alone (plus generated extra tests)
d3: MDT treatment decision, based on the above information combined, as decided by the MDT – study participants were treated empirically according to this decision
d4: final reference decision, after 12-month follow-up at the consensus meeting – the final reference decision is the optimal treatment decision based on the ultimate participant outcome at 12 months.

Effect of decision concordance on the economic analysis

The care pathway for Streamline participants can be divided into two stages: the treatment decision pathway and the subsequent disease management pathway. The former includes the time from initial diagnosis to treatment decision by the MDT; the latter includes the time period following the treatment decision.

If there is no difference in treatment decisions made with the two different staging pathways (d1 and d2), the only difference in costs assigned to the two staging pathways is the differential costs of the two sets of staging tests ('standard' vs. 'WB-MRI plus generated tests'), and there will be no difference in treatment pathways or outcomes on the basis of the experimental staging tests used. If there is no difference between the treatment decisions made with the two different staging pathways (d1 and d2) and the MDT/consensus reference standard treatment decisions (d3 and d4), then, as well as there being concordance between the two staging pathways, the agreement with d4 suggests that both d1 and d2 are perfect decisions. Alternatively, if there is no difference between the treatment decisions made with the two different staging pathways (d1 and d2) but the MDT/consensus reference standard treatment decisions (d3 and d4) are different, this suggests that there is concordance between standard staging and WB-MRI pathways but that both pathways are imperfect.

As detailed in the study protocol,[23] concordance between the standard staging pathway and the WB-MRI staging pathway was defined as > 90% of treatment decisions being the same, or < 10% of treatment decisions being different, using both pathways. Discordance was defined as > 10% of treatment decisions being different, or < 90% of treatment decisions being the same, using both pathways. In the case of concordance, the economic analysis can focus on the cost of the treatment decision pathways only, because the disease management pathways will be no different. In this case the cost-effectiveness of WB-MRI versus standard staging pathways depends only on the incremental cost (positive or negative) of WB-MRI versus standard staging pathways in the treatment decision pathway.

Conversely, if there is discordance between the treatment decisions, suggesting that participants would have received different treatment depending on which of the two staging pathways was used, then the economic analysis ought to include both the treatment decision pathways and the subsequent disease management pathways because both of these will vary between the WB-MRI and the standard staging

pathway. In this case the cost-effectiveness of the WB-MRI pathway depends on the incremental cost of the WB-MRI versus standard staging algorithms in the treatment decision pathway plus the incremental costs and health benefits of the disease management pathway.

The precise nature of the economic analysis therefore depends on the degree of concordance between treatment decisions provoked by WB-MRI versus standard staging pathways.

Available treatment decisions in Streamline C and L

For colorectal cancer participants, the four available decisions are the following summary treatment decision categories (see *Table 3*):

(A) chemotherapy (and/or radiotherapy) without surgery
(B) surgery for the primary tumour and chemotherapy (and/or radiotherapy)
(C) surgery for the primary tumour but no chemotherapy
(D) surgical metastectomy with or without chemotherapy.

For lung cancer participants, the two available treatment decisions are the following summary treatment decision categories (see *Table 4*):

(X) treatments with curative intent
(Y) treatments with non-curative intent.

Concordance in treatment decisions in Streamline C and L

As can be seen from *Table 13*, *Table 25* and *Appendix 63*, *Table 101*, and *Appendix 64*, *Table 102*, there is > 90% agreement (i.e. concordance) between d1 and d2 in both Streamline trials. Specifically, in *Appendix 62*, *Table 101*, and *Appendix 63*, *Table 102*, the sums of the off-diagonal elements are < 10% (1.69 + 0.34 + 3.04 = 5.07 and 2.19 + 0.55 = 2.74 for Streamline C and L, respectively). *Appendix 65*, *Table 103,* and *Appendix 67*, *Table 104*, as well as *Appendix 67*, *Table 105*, and *Appendix 68*, *Table 106*, show that there is discordance between d1 and d2 and the panel consensus optimal treatment decision (d4), but with similar patterns in the magnitude of the off-diagonal elements seen for each of d1 and d2, in each cancer. These results suggest, first, that the two staging pathways are concordant in terms of treatment decisions made. Second, the results suggest that d1 and d2 are imperfect, that is, the optimal treatment decisions that were made with the benefit of 12 months of follow-up and clinical outcomes are similarly different for each pathway. Overall, d1 and d2 are equivalently imperfect when considering the subsequent 12 months of participant follow-up.

As the decision based on standard staging pathways (d1) and that based on the WB-MRI staging pathway (d2) have been shown to be equivalent in both cancers (non-small-cell lung and colorectal), the economic analyses presented here therefore focuses only on the incremental costs of the two staging pathways for each cancer; this means that a cost-effectiveness analysis using a lifetime time horizon was not required, because the only difference between the two options was the cost of obtaining the treatment decision.

Cost of the treatment decision pathways

We calculated the mean cost per participant of investigations performed when undergoing standard staging pathways only and when undergoing the WB-MRI staging pathway (including additional staging tests generated as a result of the WB-MRI) for Streamline L and Streamline C. We only included the cost of the investigations performed; the costs of the MDT were not included, as this cost was incurred irrespective of the type of staging investigation performed. The analysis plan also required inclusion of costs for any adverse events related to imaging, but there were no such events reported in either trial. The type and

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

number of investigations performed were recorded for every participant in Streamline L and Streamline C. Unit costs for each test were taken from *NHS Reference Costs 2016/17*[32] and are summarised in *Table 29*. Decisions about which reference costs should be assigned to each investigation were taken by clinical colleagues in consultation with the health economists. We report the total number of each type of investigation in each trial, the mean number of each type of investigation per participant and the mean cost of that test per participant. For the mean cost of test per participant, 95% CIs are reported, derived from 1000 bootstrapped replications of the mean.

TABLE 29 Unit costs from *NHS Reference Costs 2016/17*[32] used for Streamline C and L

Currency code	Currency description	Investigations applied in Streamline C	Investigations applied in Streamline L	Mean unit cost (£)
RD03Z	MRI scan of one area, with pre and post contrast	Liver MRI, spine MRI, brain MRI	Adrenal MRI, head MRI	180.35
RD01A	MRI scan of one area, without contrast, ≥ 19 years of age	Pelvis MRI, pelvis/rectum MRI	Other MRI, MSK-MRI	139.30
RD05Z	MRI scan of two or three areas, with contrast	WB-MRI, abdomen and pelvis MRI	WB-MRI	206.51
RN03A	PET-CT of more than three areas, ≥ 19 years of age	PET-CT	PET-CT	484.17
RD21A	CT scan of one area, post contrast only, ≥ 19 years of age	chest CT, other CT	n/a	97.39
RD26Z	CT scan of three areas, with contrast	Colonography CT	n/a	122.51
YD03Z	Percutaneous biopsy of lesion of lung or mediastinum	Surgical biopsy	CT-guided biopsy	791.50
RD22Z	CT scan of one area, pre and post contrast	n/a	Head CT	120.07
RD21A	CT scan of one area, post contrast only, ≥ 19 years of age	n/a	Liver CT	97.39
RD24Z	CT scan of two areas, with contrast	Abdomen and pelvis CT	Abdomen and pelvis CT, chest and abdomen CT	112.33
RD26Z	CT scan of three areas, with contrast	Chest, abdomen and pelvis CT	Chest, abdomen and pelvis CT, neck, chest and abdomen CT	122.51
RN15A	Nuclear bone scan of two or three phases, ≥ 19 years of age	n/a	Bone scan	292.40
EBUS	Endobronchial ultrasound examination of mediastinum and percutaneous biopsy of lesion of lung or mediastinum	n/a	EBUS and EBUS/TBNA	1441.17
RD40Z	Ultrasound scan with duration of < 20 minutes, without contrast	n/a	US	51.78
YJ04Z	Core needle biopsy of axillary lymph nodes	n/a	US-guided biopsy	92.43
RD97Z	Admission or attendance for diagnostic imaging	n/a	Radiography	18.71
FE35Z	Diagnostic flexible sigmoidoscopy, ≥ 19 years of age	n/a	Sigmoidoscopy	169.43
RD42Z	Ultrasound scan with duration of ≥ 20 minutes, without contrast	n/a	Rectal US	64.95
DZ70Z	Endobronchial ultrasound examination of mediastinum	n/a	EUS	649.67
DZ69A	Diagnostic bronchoscopy, ≥ 19 years of age	n/a	Bronchial washing	686.15

EUS, endoscopic ultrasonography; MSK-MRI, musculoskeletal magnetic resonance imaging; n/a, not applicable; TBNA, transbronchial needle aspiration.

Results

Streamline C

The most frequent investigations in the standard staging pathway were chest, abdomen and pelvis CT, pelvis/rectum MRI and chest CT (*Table 30*). Twenty-one additional staging tests were generated following WB-MRI across all 299 participants (mean 0.07 extra tests per participant).

Mean test costs per participant (bootstrapped 95% CIs) were £285 (95% CI £260 to £310) for standard pathways and £216 (95% CI £211 to £221) for the WB-MRI pathway.

TABLE 30 Mean per-participant staging cost ($n = 299$ participants) according to staging pathway: Streamline C

Investigation	Frequency (*n*)		Mean number of tests per participant		Unit cost (£)	Mean cost per participant, £ (95% CI)	
	Standard staging pathway	WB-MRI staging pathway	Standard staging pathway	WB-MRI staging pathway		Standard staging pathway	WB-MRI staging pathway
WB-MRI	0	299	0.00	1.00	207	0 (0 to 0)	207 (207 to 207)
Chest, abdomen and pelvis CT	243	5	0.81	0.02	123	100 (94 to 105)	2 (0 to 4)
Pelvis/rectum MRI	120	6	0.40	0.02	139	56 (48 to 64)	3 (1 to 5)
Chest CT	44	1	0.15	0.00	97	14 (10 to 18)	0 (0 to 1)
PET-CT	43	0	0.14	0.00	484	70 (51 to 88)	0 (0 to 0)
Liver MRI	35	0	0.12	0.00	180	21 (15 to 28)	0 (0 to 0)
Abdomen and pelvis CT	27	0	0.09	0.00	112	10 (6 to 14)	0 (0 to 0)
US	12	2	0.04	0.01	52	2 (1 to 3)	0 (0 to 1)
Colonography CT	10	0	0.03	0.00	123	4 (2 to 7)	0 (0 to 0)
Radiography	9	0	0.03	0.00	19	1 (0 to 1)	0 (0 to 0)
Rectal US	4	0	0.01	0.00	52	1 (0 to 2)	0 (0 to 0)
Liver CT	3	1	0.01	0.00	97	1 (0 to 2)	0 (0 to 1)
Abdomen and pelvis MRI	4	0	0.01	0.00	207	3 (0 to 5)	0 (0 to 0)
Bone scan	2	1	0.01	0.00	292	2 (0 to 5)	1 (0 to 3)
Brain MRI	0	3	0.00	0.01	180	0 (0 to 0)	2 (0 to 4)
Other CT	1	0	0.00	0.00	123	0 (0 to 1)	0 (0 to 0)
Sigmoidoscopy	1	0	0.00	0.00	169	1 (0 to 2)	0 (0 to 0)
US-guided biopsy	0	1	0.00	0.00	92	0 (0 to 0)	0 (0 to 1)
Spine MRI	0	1	0.00	0.00	180	0 (0 to 0)	1 (0 to 2)
Total			1.87	1.07		285 (260 to 310)	216 (211 to 221)

Participants may undergo more than one of the same imaging test. Tests ranked by the most frequent across both arms. All costs are 2016/17 Great British pounds, 95% CIs are derived from 1000 bootstrapped replications of the mean. Reproduced from Taylor *et al.*[26] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Streamline L

The most frequent tests with conventional imaging were PET-CT followed by head MRI and head CT (*Table 31*). WB-MRI generated 45 additional investigations overall across all 187 participants (mean 0.24 extra tests per participant). Twenty-six participants underwent PET-CT following WB-MRI.

Mean test costs per participant (bootstrapped 95% CIs) were £620 (95% CI £574 to £666) for conventional imaging and £317 (95% CI £273 to £361) for WB-MRI.

TABLE 31 Mean per-participant staging cost ($n = 187$ participants) according to staging pathway: Streamline L

Investigation	Frequency (n)		Mean number of tests per participant		Unit cost (£)	Mean cost (£) per participant (95% CI)	
	Standard staging pathway	WB-MRI staging pathway	Standard staging pathway	WB-MRI staging pathway		Standard staging pathway	WB-MRI staging pathway
PET-CT	181	26	0.97	0.14	484	469 (449 to 488)	67 (43 to 91)
WB-MRI	0	187	0.00	1.00	207	0 (0 to 0)	207 (207 to 207)
Head CT	25	0	0.13	0.00	120	16 (10 to 22)	0 (0 to 0)
Head MRI	16	6	0.09	0.03	180	15 (8 to 23)	6 (1 to 10)
US	8	4	0.04	0.02	52	2 (1 to 4)	1 (0 to 2)
Abdomen and pelvis CT	11	0	0.06	0.00	112	7 (3 to 10)	0 (0 to 0)
EBUS/TBNA	6	4	0.03	0.02	1441	46 (10 to 83)	31 (0 to 61)
Radiography	9	0	0.05	0.00	19	1 (0 to 2)	0 (0 to 0)
Chest CT	8	0	0.04	0.00	97	4 (1 to 7)	0 (0 to 0)
US-guided biopsy	6	1	0.03	0.01	92	3 (1 to 5)	0 (0 to 1)
Liver CT	6	0	0.03	0.00	97	3 (0 to 6)	0 (0 to 0)
Chest and abdomen CT	5	0	0.03	0.00	112	3 (0 to 6)	0 (0 to 0)
CT-guided biopsy	5	0	0.03	0.00	792	21 (3 to 39)	0 (0 to 0)
Neck, chest and abdomen CT	3	0	0.02	0.00	123	2 (0 to 4)	0 (0 to 0)
Chest, abdomen and pelvis CT	2	1	0.01	0.01	123	1 (0 to 3)	1 (0 to 2)
Other MRI	1	2	0.01	0.01	139	1 (0 to 2)	1 (0 to 2)
EUS	2	1	0.01	0.01	650	7 (0 to 17)	3 (0 to 10)
Bronchial washing	3	0	0.02	0.00	686	11 (0 to 23)	0 (0 to 0)
Adrenal MRI	2	0	0.01	0.00	180	2 (0 to 5)	0 (0 to 0)
MSK-MRI	2	0	0.01	0.00	139	1 (0 to 4)	0 (0 to 0)
Surgical biopsy	1	0	0.01	0.00	792	4 (0 to 12)	0 (0 to 0)
Total			1.61	1.24		620 (574 to 666)	317 (273 to 361)

MSK-MRI, musculoskeletal magnetic resonance imaging; TBNA, transbronchial needle aspiration.
Participants may undergo more than one of the same imaging test. Tests ranked by the most frequent across both arms.
All costs are 2016/17 Great British pounds, 95% CIs are derived from 1000 bootstrapped replications of the mean.
Reproduced from Taylor *et al.*[26] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Chapter 6 WB-MRI and standard staging pathways: comparison of participant experience, perceived burden and preferences

Parts of this chapter are reproduced from Evans *et al.*[33] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

Parts of this chapter are reproduced from Evans *et al.*[34] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

Parts of this chapter are reproduced from Miles *et al.*[35] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

Introduction

Whole-body magnetic resonance imaging has the potential to replace current multimodality staging strategies and improve patient outcomes by better triaging to optimal therapy and decreasing the time between diagnosis and treatment. A single WB-MRI scan has the added advantage of reducing exposure to ionising radiation, in theory reducing the risk of subsequent radiation-induced malignancies, particularly in those diagnosed at a younger age. However, acquisition times for WB-MRI (between 45 minutes and 1 hour) are much longer than most standard imaging techniques, which take a matter of minutes or even seconds. MRI machines are also noisier than standard imaging machines and WB-MRI requires full body and head immersion into a relatively narrow 'tube'. Between 5% and 30% of patients experience distress associated with the anticipated and actual experience of undergoing MRI[36–40] and severe claustrophobia can lead to premature scan termination in 1–15% of attempts.[41,42] Patient acceptability is crucial to the successful uptake of any new technology, as poor tolerability could reduce patient adherence to WB-MRI protocols, thereby reducing the impact, even if diagnostic accuracy is equivalent to or even better than existing tests.

The aims of the three participant experience studies were therefore to assess the psychological burden and acceptability of WB-MRI compared with standard staging scans (notably CT and PET-CT) and ascertain drivers of participant preferences for one pathway over another. Although there has been some qualitative work on patient experience during general MRI,[43,44] there has been very little work on WB-MRI specifically, particularly in cancer patients who have increased levels of psychological stress around the time of diagnosis.[45] Study 1 was therefore a qualitative study that aimed to ascertain the challenges faced by participants undergoing WB-MRI and standard tests via a series of detailed patient interviews. Study 2 then quantified the level of burden experienced by participants undergoing staging investigations using a questionnaire design and examined predictors of scan tolerance. Study 3 built on the findings of studies 1 and 2 and took the form of a DCE to examine the relative importance of different scan attributes to participants, and the trade-offs they were prepared to make between different attributes such as diagnostic accuracy and scan

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

duration. The attributes required by a staging pathway based on initial WB-MRI to be favoured by patients over standard staging were then identified.

Participant experience study 1: participant interview study to examine WB-MRI burden and acceptability in comparison with standard staging scans

Methods

A qualitative methodology was chosen given the lack of pre-existing research on patient experiences of WB-MRI compared with standard staging investigations. Such a qualitative approach does not assume a priori knowledge of patient responses and allows full capture of the range of patient experience without the inhibition of a group setting.

Participants

Participants were recruited from 10 sites participating in Streamline C and/or Streamline L (see *Chapter 2*). When consenting to participate in the main Streamline trials, participants could also opt into this participant interview study. A health psychology researcher contacted consenting participants after they had completed their staging investigations (including WB-MRI) and where possible before starting their treatment. The a priori sample size was 25 participants with non-small-cell lung cancer and 25 participants with colorectal cancer (as with 50 participants it was expected that the point of saturation would be reached, after which little new information would be forthcoming with additional interviews). However, the team was permitted to expand the recruitment number if they felt the point of saturation had not been reached.[33]

Interviews

Interviews were conducted either face to face or by telephone by two trained and experienced psychology researchers. A topic guide was created (see *Appendix 69*, *Box 1*) in collaboration with the Streamline investigator team to gauge participants' views on the experience of undergoing WB-MRI compared with standard staging. However, participants were encouraged to discuss any issues that they felt were important to them and all posed questions were deliberately open ended, with flexibility in the order of topics.

Interviews were digitally recorded and participants were paid £20 plus travel expenses for participation.

Data coding and analysis

NVivo version 10 (QSR International, Warrington, UK) was used to manage data. Interviews were transcribed verbatim and analysed using thematic analysis.[46] A coding structure was developed through a combination of transcript reviews and a reflective log created by the interviewing researchers. This was an iterative process and updated as the interviews were conducted. The code was reviewed and agreed by both psychology researchers and enabled relevant themes to be identified, such as participants' beliefs and coping strategies during staging scans, and these were recorded using matrix tables.

Data were entered into SPSS version 20 (SPSS Inc., Chicago, IL, USA). Analysis of variance assessed differences in mean ages between lung and colorectal participant groups, chi-squared tests assessed group differences in sex and interview method and the Mann–Whitney *U*-test was used to assess differences in the median time interval from WB-MRI scan to interview date, between participant groups.

Results

The first 123 trial participants were approached, of whom 91 (74%) initially agreed to participate. Of these 91, 51 participants (56%) underwent an interview (the additional interview was completed because a participant expressed a strong wish to share their views even though their participation was no longer required as the point of saturation was considered to have been reached after 50 participant interviews).

Reasons participants were not interviewed included retracted consent ($n = 8$), participation in study 2 (questionnaire study) ($n = 14$), withdrawal from the main trial ($n = 12$) and interview quota reached before completion of all staging imaging tests ($n = 6$).[33] Interviews lasted between 12 and 86 minutes (mean 48 minutes). Overall, 39 participants (76%) were interviewed over the phone and the rest face to face.

In total, 47 of the 51 participants completed the WB-MRI, and the median interval between WB-MRI and interview was 15 days (range 1–63 days). All Streamline C participants reported a cancer diagnosis whereas around only half of Streamline L participants reported a cancer diagnosis at the time of interview; others were still waiting for the outcome of test results. The full participant demographics of the interviewed cohort are shown in *Appendix 70*, *Table 107*.

There were no significant differences in age or sex between those who initially agreed to take part in the interviews ($n = 91$) and those who declined ($n = 32$) [mean age: 65 years vs. 64 years, $F_{1,121} < 1$, $p = 0.527$; percentage male: 64% vs. 56%, $\chi^2 = 0.562$, degrees of freedom (df) = 1, $p = 0.453$]. There were also no significant differences in demographics between the 51 (of the 91) participants who actually undertook the interviews and the 40 who did not (mean age: 65 years vs. 64 years, $F_{1,89} < 1$, $p = 0.579$; percentage male: 53.4% vs. 46.6%, $\chi^2 = 0.437$, df = 1, $p = 0.508$).

Participant experience of WB-MRI in comparison to standard staging investigations

Participants reported some challenges during standard staging scans, for example the use of intravenous contrast agents: 'you feel a bit, you know, woozy, funny [. . .] and it did make me feel slightly off' (< 50 years of age, male, Streamline C). Concerns were also expressed about exposure to ionising radiation: 'all that radiation dye in my system and everything [. . .] And it just disturbed me. Because I couldn't go near my niece and she's only 6 [years of age]' (< 50 years of age, female, Streamline L).

Among the participants who completed the WB-MRI, there was variation in whether or not the experience of WB-MRI was challenging and the intensity of difficulties encountered. WB-MRI was enjoyable for some: 'I was very relaxed and quite enjoyed the sensations really of the hammer blows and stuff like that [. . .] I was in a long time and it was fun' (≥ 80 years of age, male, Streamline C). For others it elicited no emotional response apart from happiness that it was being performed: 'No nerves, no worries, no anxieties. I feel happy' (50–59 years of age, male, Streamline L). The 'coping challenge' created by WB-MRI formed a continuum: some enjoyed it or experienced no difficulties, perceiving it to be 'just another bit of kit that they're going to put me through. No worry, no apprehension' (≥ 80 years of age, male, Streamline L); however, many participants viewed WB-MRI as more challenging than standard scans, for example stating 'I would say the most challenging of the MRI scans was the body scan [. . .] It's quite different from the other two [i.e. CT and PET-CT scans] [. . .] It's different than what I would call the normal scans' (60–69 years of age, male, Streamline C). Indeed, some found WB-MRI to be very challenging, describing it as 'torture – medieval torture' (50–59 years of age, male Streamline C), whereas CT was described as 'a walk in the park' (60–69 years of age, male, Streamline C) and 'very simple' (50–59 years of age, female, Streamline L). PET-CT was described by some as 'so much easier' (70–79 years of age, female, Streamline L) and 'wasn't so unpleasant' (50–59 years of age, male, Streamline C).

Challenges of WB-MRI

Among the cohort, the factors that were identified as difficult during the interviews were relatively consistent, regardless of how challenging people found the scan. In particular, claustrophobia was often cited as problematic, along with the physical discomfort, noise, scan duration and the challenge of coping with emotions and negative thoughts elicited during the scan, including fear/panic and isolation. The intensity of these problems was variable, but increased in those finding the scan more challenging (*Table 32*).

Participants who found one aspect of the scan very problematic also appeared to find other aspects difficult: 'You're kind of helpless and trapped in a tunnel [. . .] that was even before the noise started and the noise is very kinda distressing really' (50–59 years of age, male, Streamline C).

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 32 Problems and the diversity of challenge experienced related to WB-MRI

	Challenge level			
Problem	**None/enjoyable**	**Low**	**Medium/high**	**Did not complete WB-MRI**
Enclosed space	*I'm not claustrophobic [. . .] that [enclosed space] doesn't frighten me* *50–59 years of age, male, Streamline L*	*Maybe in the back of my mind you're worried about it [enclosed space]* *60–69 years of age, female, Streamline C*	*Coffin came to mind [. . .] initially, you feel like 'Oh god, it's so close in here'* *60–69 years of age, female, Streamline L*	*I had this sense of being in a sarcophagus* *60–69 years of age, male, Streamline L*
Noise	*I thought it like music really [. . .] and it relaxed me* *≥ 80 years of age, male, Streamline C*	*The noise was a bit [. . .] disconcerting because it was quite loud and quite sort of clanking and banging [. . .] and it took some time getting used to it* *< 50, years of age, male, Streamline C*	*That's one of the worst ones that I've had to go through with the noise [. . .] I felt like something was going to fall off and hit me* *< 50 years of age, female, Streamline L*	
Scan duration/discomfort	*It's actually quite comfortable in there* *60–69 years of age, male, Streamline L*	*It was not excruciatingly painful. It just ached a bit, that is all* *60–69 years of age, male, Streamline C*	*My back was uncomfortable. My, the back of my neck was uncomfortable [. . .] My brain was saying 'You can't go much longer'* *60–69 years of age, male, Streamline C*	
Negative emotions (e.g. anxiety, panic, shock, distress)	*I mean, there wasn't, you know, anything in particular to be frightening* *< 50 years of age, male, Streamline C*	*When I first went in there, I was quite shocked [. . .] I wondered what the hell was going on [. . .] After a while I quite enjoyed it* *60–69 years of age, male, Streamline C*	*When I first laid down in the scan I was frightened* *60–69 years of age, female, C* *I'm lying there and all of a sudden they put all these things on top of me [. . .] And I'm thinking, 'Oh, my god! [. . .] This is threatening'* *60–69 years of age, male, Streamline C*	*I don't know whether the people doing the scan knew I had lung cancer. I didn't feel any empathy or any sort of friendly attitude. I walked out of there feeling quite upset actually* *50–59 years of age, male, Streamline L*

Reproduced from Evans *et al.*[33] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

Claustrophobia

One reason offered by participants as to why the CT and PET-CT scans were 'easier' was because the scanner construction was deemed 'a lot more open' (50–59 years of age, female, Streamline L) and 'it was almost like going under a bridge [. . .] as opposed to going down a long dark tunnel underground' (60–69 years of age, male, Streamline L), whereas participants often felt trapped or buried in the WB-MRI: 'it was the feeling of being sort of trapped because you can feel the machine all around your body' (50–59 years of age, male, Streamline C); 'I had this sense of being in this sarcophagus' (60–69 years of age, male, Streamline L). Whereas some experienced intense reactions to being enclosed, only one participant asked to terminate the scan early because of this.

Noise

The noise emitted by MRI machinery during image acquisition was thought to be unpleasant by some participants: 'Horrendous. It's horrible! I think it's one of the worst things there is' (60–69 years of age, female, Streamline L); 'like having my head in a bucket with someone hitting a hammer' (70–79 years of age, female, Streamline L). However, for others it was not disconcerting, and some even compared it to music: 'There was some hammer noise. And, there was bell noises. And there were also some different noises. And I thought it's like music really, you know, disjointed a bit, but like music, isn't it? And, it's relaxed me' (80–89 years of age, Streamline C male). The unpredictability of noise was concerning for some – 'the sound was really irritating [. . .] it was very unpredictable and it changed quite a lot' (50–59 years of age, male, Streamline C) – but engaging to others: 'It was quite good [. . .] I would say, "What's coming next? What's the next noise?" So, I enjoyed that. The not knowing what was coming next' (60–69 years of age, male, Streamline C). The inability to control noise exposure was also viewed as difficult by some: 'You wouldn't have that much noise for an hour in any situation other than that because you would move away from it' (70–79 years of age, female, Streamline C). Certain sounds were sometimes interpreted to indicate the presence of pathology – 'I thought, "If the machine is clicking on, maybe something is wrong there with me," you know?' (60–69 years of age, female, Streamline C) – or created concern that the machine had a fault. Over time, some participants habituated to the noise – 'You get used to it in the end [. . .] I think the first, the first time you hear it, "oh goodness me" ' (60–69 years of age, female, Streamline L) – but others found their tolerance reduced as the scan progressed: 'Right at the very end, you get tired and then, you know, you get tired of the noise at the very end' (60–69 years of age, female, Streamline C).

Scan duration and physical discomfort

The longer time required to complete WB-MRI in comparison to CT and PET-CT was frequently noted: 'The CT scan is very quick [. . .] you're in and out. Whereas, the MRI scan was very long' (60–69 years of age, female, Streamline C). Scan duration was also linked to comfort: 'CT scan is not as uncomfortable as the WB-MRI scan because of the length of the scan' (50–59 years of age, female, Streamline C); 'My back was uncomfortable [. . .] the back of my neck was uncomfortable, maybe 15 minutes too long to be comfortable' (60–69 years of age, male, Streamline C). In contrast, other participants suffered no discomfort and viewed the MRI scan duration as acceptable: 'I think it's over 1 and a half hours or something like that, so it was alright for me [. . .] is comfortable' (60–69 years of age, male, Streamline C). One participant whose PET appointment immediately followed the WB-MRI drank large quantities of water, as instructed, which resulted in numerous toilet trips prior to WB-MRI and led to panic in the last 10 minutes of the scan when she realised she needed to go again: 'And that's when a sense of panic started to set in, because I thought [. . .] I'm thinking to myself, "I think I want to go to the toilet" ' (70–79 years of age, Streamline L).

Emotions elicited

Some participants referred to negative emotions related to WB-MRI including panic and fear: 'I could feel a couple of times [. . .] I was going to have a sort of panic attack and have to be taken out of the machine' (50–59 years of age, male, Streamline C). When asked what caused this, the participant said: 'This claustrophobia, the feeling of being [. . .] the feeling of being trapped, really' (50–59 years of age, male, Streamline C). The novelty of the WB-MRI experience prompted worry in some: 'The oppression comes

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

from the unknown' (70–79 years of age, male, Streamline C). Some feared the scan result, specifically a fear of finding that the cancer was advanced: 'Whether they'd find [. . .] it has gone everywhere and you know nothing you can do, I suppose that was the main worry' (60–69 years of age, female, Streamline L); 'when I first laid down in the scan I was frightened [. . .] I was thinking what if they find something the other scan didn't' (60–69 years of age, female, Streamline C). However, while WB-MRI had the potential to reveal bad news, there was acceptance that it was preferable to know if their cancer had spread: 'but that is a good thing, it will be dealt with' (70–79 years of age, male, Streamline C). Dealing with feelings of panic and fear was a challenge in itself and needed effort to overcome: 'I was trying to put it [fear] out of my head and think positively [. . .] I was sort of telling myself "this isn't the end of the world, you just have to think ahead, don't think the worst" ' (60–69 years of age, male, Streamline L).

Cancer context

As noted in some of the quotations above, the WB-MRI was perceived to have the potential reveal additional findings associated with a poorer prognosis and treatment implications: 'I was lying in there and I was thinking, what if they say it's cancer? Then, what?' (< 50 years of age, female, Streamline L). One participant reported relief when the scan confirmed that initial tumours were localised: 'My biggest fear, I know I've got the cancer in the bowels. And I was worried that they might, it might just disguise I've got other cancers. So, when he said straight away, there was nothing else and I feel relieved' (70–79 years of age, male, Streamline C). One participant had a dilemma of whether or not to go ahead with the scan because of the possibility that its findings might mean he was not offered treatment: 'I was of two minds about the trial scan because I didn't want the trial scan to say the cancer's spread and they wouldn't operate' (50–59 years of age, male, Streamline L); but there was also appreciation of having the WB-MRI even when it did reveal spread: 'The MRI actually showed up even more of what's going on with me and I think if I hadn't had it, they would [. . .] we still wouldn't know. Because they found traces in five other places I feel a lot better now having the MRI scan, knowing that they've picked up on all of it [. . .] Without that, they would [. . .] I would just be having the chemo and not the full treatment that I need' (< 50 years of age, female, Streamline L).

The underlying known or potential cancer diagnosis may have influenced how physical characteristics of the scan where interpreted. One participant described the enclosed experience as follows: 'It's like being in a coffin. You know, I did think that. It's like being in a coffin. But I wasn't. You just have to think that [. . .] Because if you've got something really, really bad, it would be like being in a coffin. Just like thinking, "My god. This is going to be me" ' (60–69 years of age, male, Streamline C).

Coping

'One has just got to face it and cope' (60–69 years of age, male, Streamline C).

The thematic analysis identified two main coping strategies. (1) Strategies focused on coping with distressing thoughts, emotional responses or physical sensations evoked by the scan, for example mental distraction and relaxation: 'Because I love Cyprus I was thinking of Cyprus the whole time. And it sort of takes your mind off of what's going on and how long you're in there' (60–69 years of age, female, Streamline L); 'Breathing, breathing, concentrating [. . .] And you know trying to sort of just to calm [. . .] Because you can't help but get anxious, you know, in a situation like that' (60–69 years of age, female, Streamline L). (2) Strategies related to motivation to complete the scan, such as focusing on benefit beliefs and acceptance that the experience was part of the process to get better: 'I was so fed up with pain that I would have done [. . .] you know, any investigation was better [. . .] It [WB-MRI] was something to get me better' (60–69 years of age, female, Streamline C); 'I was laying there and I was thinking, "Well, it's for my own benefit, so I can live with that" ' (60–69 years of age, male, Streamline C).

Staff contact

Interactions with staff were noted as important by many participants. Staff supported participants with verbal reassurance and information provision as well as making them physically comfortable. For example: 'They told me all about it so I didn't get anxious. I knew what to expect' (60–69 years of age, female,

Streamline L); 'They managed to find something to put under my neck [. . .] They did put it in the right place under my neck and I was OK' (60–69 years of age, male, Streamline C). Some participants spoke of a sense of isolation when in the scanner and hearing a voice confirmed they were not alone: 'You feel a bit isolated because you're in the scanner, on your own, and everybody else is behind a shielded glass in a room' (< 50 years of age, male, Streamline C); 'Just made me feel a bit confident that you wasn't on your own, you know?' (60–69 years of age, female, L). Staff communication also acted as distraction: 'The person speaking keeps my mind occupied' (60–69 years of age, female, Streamline L).

Explaining differences in people's experiences

Site of cancer

Participants from both study cohorts described varied experiences. However, chest symptoms reported by Streamline L participants were associated with discomfort – 'I found lying flat on my back, yes, very uncomfortable [. . .] I was coping OK for about 20, 25 minutes, and then that's when this pressure or aching or uncomfortable feeling on my chest began to kick in' (50–59 years of age, male, Streamline L) – particularly when there was a need to breath hold during sequence acquisition: 'I was trying not to cough because she was saying, "Please, no. If you can, please don't" and when you've got to hold your breath [. . .] was very difficult' (60–69 years of age, female, Streamline L); 'They had to make me do it a few times because I couldn't hold it for the length of time they wanted me to hold it [. . .] I've got COPD [chronic obstructive pulmonary disease]; whether that made it worse, I don't know' (60–69 years of age, male, Streamline L).

All four participants who did not complete the scan were recruited to Streamline L but their decisions to stop the scan were not necessarily related to their symptoms. For example, for one participant it was the discomfort arising from the head coil and for another it was claustrophobia.

Existing musculoskeletal problems

Sometimes participants reported that the WB-MRI exacerbated musculoskeletal problems: 'I do suffer a bit from back pain [. . .] My back started to sort of be pretty uncomfortable' (60–69 years of age, male, Streamline C); 'I was worried about my neck being uncomfortable because I have got a very stiff neck because of arthritis but they managed to sort me out and then I managed to last the hour, or whatever it was, without too much pain' (60–69 years of age, male, Streamline C).

Mental health comorbidities

Mental health issues increased the challenge of the scan for two participants. One participant with prior anxiety found the scan difficult, describing their anxiety as causing feelings of claustrophobia: 'I'm not claustrophobic. I just felt that way because of the anxiety. When I get anxious, it feels like everything's just closing in on me' (< 50 years of age, female, Streamline L). Several aspects of the scan evoked this anxiety, such as wearing the helmet coil, the noise of the scan and the scan duration. Although the scan was difficult, this participant had several methods of coping that enabled her to deal with the increased anxiety, for example taking extra psychiatric medication, relaxing and mental distraction. Another participant described himself as an alcoholic and found the scan experience difficult because he had to abstain from food and alcohol: 'But you see the only problem, I'm an alcoholic as well. And to be without a drink all that time, I was getting really bad at the end [of the scan] [. . .] Because I normally have a few pints about 8 or 9 [in the morning], you know' (60–69 years of age, male, Streamline L).

Prior experiences

Prior experiences of claustrophobia, MRI and experiences outside medical contexts were linked with the ability to cope. One participant who terminated the scan after 10 minutes did so because he did not like being constrained and his aversion came as a surprise: 'A sense of being constrained and being strapped down which is a new one for me [. . .] I haven't encountered that situation before' (60–69 years of age, male, Streamline L). In contrast, another participant (60–69 years of age, male, Streamline C) who knew he suffered from severe claustrophobia requested a sedative and completed the scan with relative ease.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Similarly, one participant who had several MRIs previously – which she reported as difficult – transformed her experience by wearing an eye mask: 'I put an eye mask on. So, you wouldn't be tempted to sort of peep; it would help you to sort of relax a bit more [. . .] I was fine with it, with the mask on' (60–69 years of age, female, Streamline C). Work experiences meant that other participants were used to confined spaces and or loud noises and therefore the scan did not represent a new challenge for them: 'It doesn't bother me. I've worked in pipes and tunnels and all sorts of places' (60–69 years of age, male, Streamline L); 'If I wasn't in the trade that I was in [car mechanic], I suppose that that could really freak you out because it's quite loud' (< 50 years of age, male, Streamline C). Vicarious experiences of friends and family were sources of knowledge for participants with no personal prior experience: 'Of course, when you haven't had it before, you've heard lots of stories from other people' (60–69 years of age, female, Streamline L). This led to expectations that were sometimes reassuring – 'I have [a] few friends and relatives, my cousin, they have already done MRI, so I had something from them. This is not that, you know, this [is not] painful or anything like that. You will enjoy' (60–69 years of age, male, Streamline C) – and sometimes contrary to their own eventual experience: 'It was not as intimidating as I thought because I had heard people talking about going in to an MRI scanner, and they were telling me how nervous they were, and how they hated it [. . .] And when I saw it, I thought, "That's okay" ' (60–69 years of age, male, Streamline C).

Staff contact

Participants varied in whether or not they felt that they received enough information and contact from staff before and during the scan: 'They told me everything I needed to know' (< 50 years of age, male, Streamline C) versus 'There wasn't any [information] to be honest, I think that was the lacking part' (50–59 years of age, female, L) and 'They talked all the way through it' (60–69 years of age, female, L) versus 'They once asked me if I was OK but I didn't really have any communication' (60–69 years of age, female, Streamline C). Participant preferences for information also varied: '[. . .] think that the less you know the better' (70–79 years of age, male, Streamline C) versus 'Personally they could give as much information as they could' (60–69 years of age, male, Streamline C). Whether or not staff were perceived to be reassuring also depended on the relationship established between staff and participants; a change in staff mid-way through the scan lead to panic in two participants: 'She was constantly talking to me, I was fine. But then [. . .] another lady took over and she didn't talk as much [. . .] So that last 10 minutes or so was [. . .] when I panicked' (60-69 years of age, female, Streamline L); 'People I've been speaking with beforehand, these are people I've actually built up the initial, you know, the trust [. . .] Then, suddenly, a complete stranger is now telling me to breathe in and out. [. . .] That complete stranger then said, "You know, I should be coming down now to inject you with dye" [. . .] In my mind, I thought the absolutely bizarre. You know, he killed everyone in the ward. Now, he's out to kill the patients' (60–69 years of age, male, Streamline C). Failure to heed participant preference influenced one participant's decision to terminate the scan: 'They told me that they're going to inject dye. And once again I said to the person, "Don't put that cannula in my left arm", but she did and the vein collapsed. And I said to them, "I can't do anymore now. I told you not to use that vein but you had, the vein's collapsed. And I think I've had enough" ' (50–59 years of age, male, Streamline L).

Generally, though, participants spoke positively about the caring attitude of staff in WB-MRI centres: 'The staff were all very good. They're very caring' (60–69 years of age, female, Streamline C).

Willingness to have another WB-MRI scan

Participants were asked whether or not they would be prepared to have another WB-MRI scan if the doctor recommended it and all said yes, even those who had requested termination of the trial scan. However, this agreement was offered with varying enthusiasm; for example: 'Only if absolutely necessary [. . .] I want a good reason' (50–59 years of age, male, Streamline C) and 'Well, if I have to [. . .] I don't like it one bit but if it has to be done' (60–69 years of age, male, Streamline L) versus 'Yes, without hesitation' (< 50 years of age, male, Streamline C) and 'I'll have the MRI any day' (60–69 years of age, female, Streamline L).

Participant experience study 2: quantitative assessment of perceived WB-MRI scan burden and acceptability – comparison with CT/PET-CT

Methods

When consenting to participate in the main Streamline trials, participants could also opt into this participant questionnaire study, which was commenced as the interview study (study 1) was completing.

Questionnaires

The study utilised two questionnaires and participants were asked to complete both. The first questionnaire was posted to participants within 2 days of being registered for the Streamline trials, to be completed around the time participants were undergoing their staging investigations. This acted as the baseline questionnaire. A second, follow-up questionnaire was completed after participants had undergone staging investigations (post staging) and was posted 1 month after the baseline questionnaire. Both questionnaires were returned using a stamped addressed reply envelope and participants were paid £20 for participation. Recruitment continued until a minimum of 100 participants had returned both questionnaires (50 for Streamline L and 50 for Streamline C) (see *Power calculation*).

Questionnaire content

The following data were collected in the baseline questionnaire.

- Emotional distress: psychological distress was measured using the General Health Questionnaire-12 items (GHQ-12).[47]
- Comorbidity: participants were asked to report ('yes' or 'no') whether or not they had any of the following – heart or vascular disease, diabetes mellitus, epilepsy, a history of stroke, arthritis, asthma and mental or emotional disorder. There was also an option to provide details of other illness.
- Demographics: participants were asked their age, sex and ethnicity. Missing demographic data on age and sex as well as postcode data were supplied via the central trial database.
- Comparative experience of scans: part of the follow-up questionnaire asked participants about their comparative experience of WB-MRI and chest, abdomen and pelvis CT (standard scan) if recruited to Streamline C or PET-CT (standard scan) if recruited to Streamline L.
- Scan recovery, satisfaction and acceptability: participants rated their post-scan recovery on a 9-point scale ranging from 'immediate' to 'a week'. Participants also rated how satisfied they were with the information received before scanning, communication and departmental facilities, as well as the overall acceptability of scans, on a scale of 1 ('very dissatisfied'/'not at all acceptable') to 4 ('very satisfied'/'very acceptable').
- Scan burden: we adapted a questionnaire previously used to assess acceptability of colonoscopy[48] and CT colonography[49] to measure scan burden. The questionnaire included 26 items and participants completed it for both WB-MRI and standard scans. Participants scored their experience of the scans for each item by ticking agreement on a 7-point Likert scale where 1 and 7 were anchored to bipolar statements related to scan discomfort (13 items), worry (6 items) and satisfaction (7 items). An example discomfort item was 1 ('the noise of the scanner was unbearable') to 7 ('the noise of the scanner was fine').

Power calculation

Power (G*Power 3; Heinrich Heine University Düsseldorf, Düsseldorf) was based on rejecting the null hypothesis that there was no significant difference in perceived burden of WB-MRI when compared with standard staging (related *t*-test). Assuming a medium effect size (Cohen's $d = 0.5$), a Cronbach's alpha of 0.05 and 95% power, a minimum number of 90 participants were required across the two study cohorts (45 in Streamline C and 45 in Streamline L). An effect size of 0.5 is considered the minimally important difference (MID) in quality-of-life measures,[50] where MID is defined as the smallest difference that participants view as important (beneficial or harmful) and would result in a doctor considering a change in the participant's management.[51]

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Statistical analysis

Data handling

- Emotional distress: using the GHQ-12 binary coding method (0,0,1,1), a mean sum score (if ≥ 50% of items were answered) was created ranging from 0 to 12. A score of ≥ 4 is considered indicative of significant distress levels.[52]
- Comorbidity: a response of 'yes' to any illness was coded and a dichotomous 'comorbidity' variable was created whereby the presence of one or more comorbid illness was reported (either 'yes' or 'no').
- Demographics: post code data were used to calculate an area-based deprivation score for each individual using the 2010 Index of Multiple Deprivation (IMD) scale,[53] categorised into quintiles from 1 (highest level of deprivation) to 5 (lowest).
- Scan recovery, satisfaction and acceptability: recovery data were collapsed into three categories – 'immediate', 'up to 30 minutes' and 'over 30 minutes' – for analysis.
- Scan burden: subscores for discomfort, worry and satisfaction scales were computed from the mean of completed items (if < 50% of items were completed, the response was coded as missing). A total score 'scan burden' was computed by taking the mean of discomfort, worry and reverse scored satisfaction subscales, with higher scores equating to greater scan burden.

Analysis was performed using SPSS version 22. Differences in demographic and psychological characteristics between Streamline L and Streamline C cohorts were assessed using the Mann–Whitney *U*-test, and chi-squared or Fisher's exact test (if ≥ 20% of the cells in the contingency table had expected counts of < 5) as appropriate. Related samples Wilcoxon signed-rank test were used to assess differences between WB-MRI and CT/PET-CT in terms of scan recovery time, scan acceptability, satisfaction with scan-related information, facilities, communication and scan burden. Linear regression tested the predictive value for WB-MRI scan burden of data collected in the baseline questionnaire. Individual predictors were entered in unadjusted analyses and those items achieving statistical significance were then entered into a multivariate analysis. Statistical significance was assigned at the 5% level, two-tailed.[34]

Results

A total of 350 participants (89.3%) of the 392 recruited to the trials during the period that study 1 and study 2 were being performed consented to undergo the interview study (study 1) or the current questionnaire study (study 2). None of the participants who took part in the present study took part in the prior interview study.

The full recruitment pathway and reasons for exclusion are presented in *Figure 7*. In total, 214 participants were sent both questionnaires. Ninety-nine were excluded. Reasons for exclusion were non-response ($n = 71$), returned baseline questionnaire only ($n = 27$) and withdrawal from the trial ($n = 1$). Overall, 103 participants (median age, 66 years; male, $n = 58$) completed both questionnaires and 12 (median age, 60 years; male, $n = 9$) completed the post staging questionnaire only. A total of 115 participants completing the follow-up questionnaires were included in the analyses. A total of 61 participants were recruited to Streamline C and 54 participants to Streamline L. Participants were more likely to return the post-staging questionnaire if they had lower levels of deprivation (linear $\chi^2 = 7.113$, df = 1; $p = 0.008$). There were no differences in sex ($p = 0.059$), age ($p = 0.676$) or cancer type (linear $\chi^2 = 0.442$; df = 1; $p = 0.506$) between those who did and did not return the post-staging questionnaire (see *Appendix 71*, *Table 108*).

Full demographics of the 115 participants are shown in *Appendix 72*, *Table 109*.

Female participants recruited to Streamline C ($n = 24$) were significantly younger than those recruited to Streamline L ($n = 24$) (median age 60 vs. 73 years, respectively; $p = 0.003$), with no significant age difference between males recruited to Streamline C ($n = 37$) and Streamline L ($n = 30$) (median age 66 vs. 66 years, respectively; $p = 0.480$).

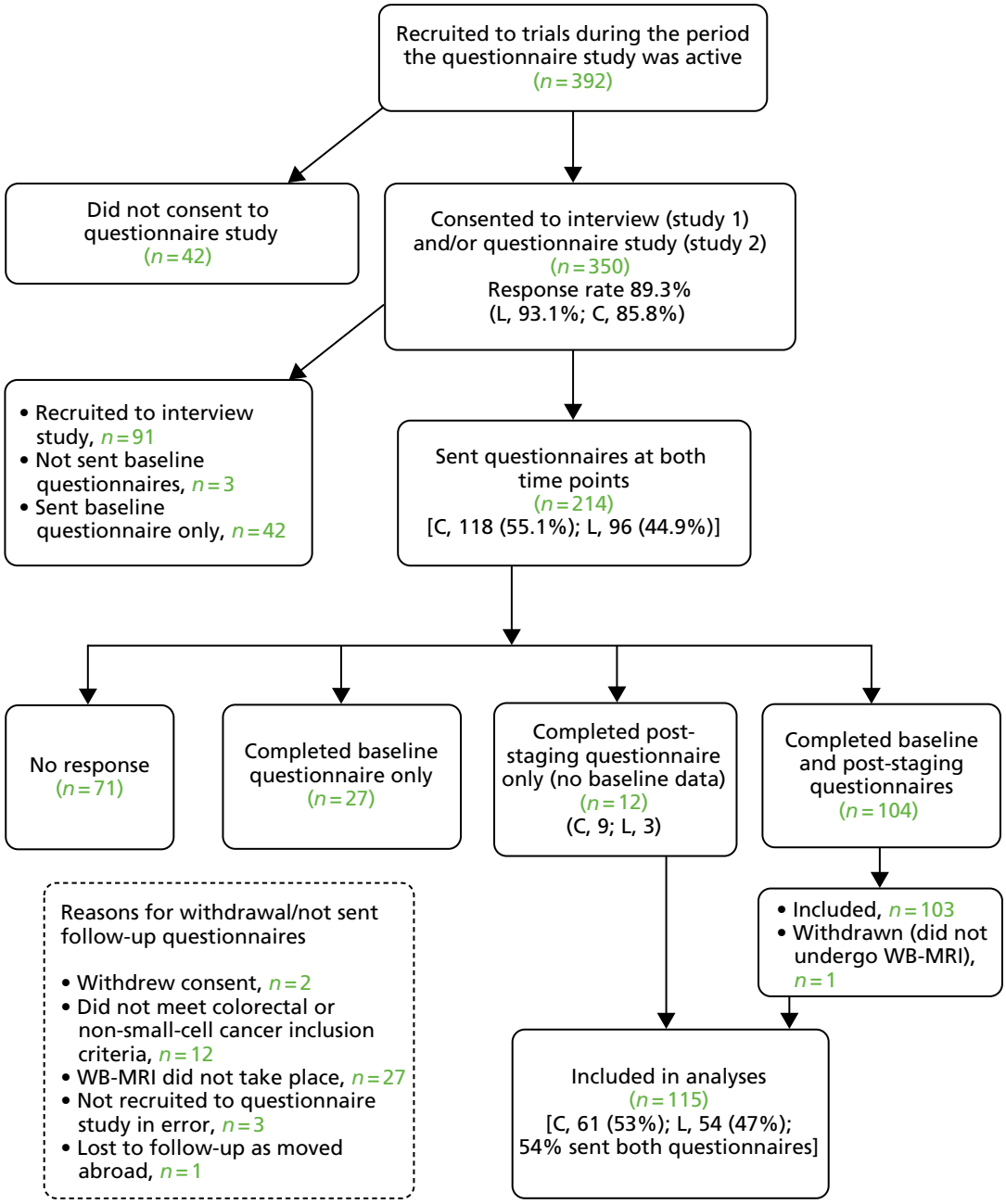


FIGURE 7 Flow diagram of participants through the studies 1 and 2. C, Streamline C; L, Streamline L. Reproduced from Evans *et al.*[34] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

Participants recruited to Streamline L were significantly more likely to report additional comorbidity than those recruited to Streamline C (66.7% vs. 40.4%; $p = 0.008$) (see *Appendix 72*, *Table 109*).

Post-scan participant recovery and scan acceptability

Table 33 shows participant scan recovery time and overall scan acceptability and *Appendix 73*, *Table 110*, shows comparative scan information, communication and facilities.

There were no significant differences in recovery time after WB-MRI compared with CT/PET-CT, with 64% ($n = 69$) of participants who completed this item reporting 'immediate' recovery following WB-MRI, compared with 65% following CT/PET-CT. However, in general, participants rated WB-MRI as less

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 33 Comparative experience of WB-MRI vs. CT/PET-CT: recovery time and scan acceptability

Variable	Overall, % (*n*)	Streamline L,[a] % (*n*)	Streamline C,[b] % (*n*)	Group differences (Wilcoxon signed-rank test) (*p*-value)
Recovery time				
WB-MRI[c]				0.465,[d] 0.735[e]
Immediate	63.9 (69)	61.5 (32)[d]	66.1 (37)[e]	
≤ 30 minutes	25.9 (28)	23.1 (12)	28.6 (16)	
> 30 minutes	10.2 (11)	15.4 (8)	5.4 (3)	
CT/PET-CT[c]				
Immediate	65.1 (69)	58.8 (30)[d]	70.9 (39)[e]	
≤ 30 minutes	21.7 (23)	23.5 (12)	20.0 (11)	
> 30 minutes	13.2 (14)	17.6 (9)	9.1 (5)	
Scan acceptability				
WB-MRI[c]				
Very	65.2 (73)	64.8 (35)[d]	65.5 (38)[e]	0.035,[d] 0.005[e]
Fairly	30.4 (34)	29.6 (16)	31.0 (18)	
Slightly	3.6 (4)	3.7 (2)	3.4 (2)	
Not at all	0.9 (1)	1.9 (1)	0.0 (0)	
CT/PET-CT[c]				
Very	77.8 (84)	75.0 (39)[d]	80.4 (45)[e]	
Fairly	21.3 (23)	23.1 (12)	19.6 (11)	
Slightly	0.0 (0)	0.0 (0)	0.0 (0)	
Not at all	0.9 (1)	1.9 (1)	0.0 (0)	

a WB-MRI vs. PET-CT
b WB-MRI vs. CT.
c Missing data > 5%.% is valid per cent.
d Comparison between WB-MRI and PET-CT (Streamline L).
e Comparison between WB-MRI and CT (Streamline C).
Reproduced from Evans *et al.*[34] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

acceptable than either CT/PET-CT. Participants' satisfaction with information before the scan and facilities, together with communication during the scan, were not significantly different between WB-MRI and either CT or PET-CT (see *Appendix 72*, *Table 109*).

Scan burden

The comparative burden induced by WB-MRI and CT/PET-CT is shown in *Table 34*.

Mean ratings for scan discomfort and worry ranged from 1.63 to 2.65 (7 represents maximum discomfort or worry). Mean satisfaction scores ranged from 6.25 to 6.53 (7 represents maximum satisfaction). In general, therefore, the burden generated by all the scans was relatively low.

However, the overall burden induced by WB-MRI was significantly greater than both PET-CT and CT (see *Table 34*). Specifically, both discomfort scores and satisfaction scores were worse for WB-MRI

TABLE 34 Comparative scan burden: WB-MRI vs. CT/PET-CT

Burden	Overall, mean (SD)	Streamline L,[a] mean (SD)	Streamline C,[b] mean (SD)	Group differences (Wilcoxon signed-rank test) (*p*-value)
Total participant burden				
WB-MRI[c]	2.21 (1.1)	2.33 (0.94)[d]	2.09 (1.18)[e]	< 0.001[d]
CT/PET-CT[c]	1.87 (0.98)	2.05 (0.82)[d]	1.70 (1.1)[e]	< 0.001[e]
Discomfort subscale				
WB-MRI[c]	2.51 (1.26)	2.65 (1.14)[d]	2.30 (1.22)[e]	< 0.001[d]
CT/PET-CT[c]	1.83 (1.05)	2.04 (.90)[d]	1.63 (1.15)[e]	< 0.001[e]
Worry subscale				
WB-MRI[c]	2.47 (1.32)	2.62 (1.15)[d]	2.23 (1.31)[e]	0.208[d]
CT/PET-CT[c]	2.24 (1.23)	2.52 (1.15)[d]	2.00 (1.28)[e]	0.041[e]
Satisfaction subscale				
WB-MRI[c]	6.25 (1.06)	6.27 (0.85)[d]	6.26 (1.23)[e]	0.036[d]
CT/PET-CT[c]	6.49 (0.89)	6.43 (0.76)[d]	6.53 (1.01)[e]	0.001[e]

SD, standard deviation.
a WB-MRI vs. PET-CT.
b WB-MRI vs CT.
c Missing data > 5%.
d Comparison between WB-MRI and PET-CT (Streamline L).
e Comparison between WB-MRI and CT (Streamline C).
Reproduced from Evans *et al.*[34] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

compared with those of both CT and PET-CT. By way of example, WB-MRI conferred significantly greater feelings of claustrophobia than both CT (mean score 2.81 vs. 1.51; $p < 0.001$) and PET-CT (mean score 3.04 vs. 1.98; $p < 0.001$) and greater burden from scan-related noise compared with both CT (mean 2.84 vs. 1.73; $p < 0.001$) and PET-CT (mean 2.85 vs. 1.63; $p < 0.001$). In general, the intravenous injections required for each of the three scan types resulted in low levels of discomfort that did not differ between scan type (WB-MRI vs. CT: 1.59 vs. 1.56, $p = 0.637$; WB-MRI vs. PET-CT: 1.86 vs 1.73, $p = 0.225$).

Questionnaire items related to 'worry' did not differ for WB-MRI in comparison to PET-CT but were inferior for WB-MRI in comparison to CT.

The WB-MRI burden was not rated differently between those recruited to Streamline C or Streamline L. However, participants recruited to Streamline L reported significantly more worry and discomfort during PET-CT compared with the equivalent ratings for CT by those recruited to Streamline C (worry 2.52 vs. 2.00, respectively, $p < 0.001$; discomfort 2.04 vs. 1.63, respectively, $p < 0.001$).

Predictors of WB-MRI scan burden

The presence of comorbidity, psychological distress and deprivation were significant predictors of WBI-MRI burden in unadjusted analysis ($\beta = 0.242$, $p = 0.015$; $\beta = 0.305$, $p = 0.002$; $\beta = –0.265$, $p = 0.005$, respectively), with age, sex, cancer type and ethnicity non-significant predictors ($\beta = 0.059$, $p = 0.535$; $\beta = 0.083$, $p = 0.389$; $\beta = –0.122$, $p = 0.201$; $\beta = –0.179$, $p = 0.081$, respectively). In the adjusted analyses only psychological distress and presence of comorbidities remained significantly predictive ($\beta = 0.223$, $p = 0.025$; $\beta = 0.191$, $p = 0.048$, respectively).

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Participant experience study 3: participant preferences for whole body MRI or conventional staging pathways in lung and colorectal cancer – a discrete choice experiment

Study 3 built on the findings of study 1 and 2 and was designed to elicit overall participant preferences for cancer staging pathways using a DCE. We followed international DCE guidelines for study design and analysis.[54,55]

A DCE elicits preferences by asking individuals to indicate their choice between two or more options, where each option contains characteristics or attributes that are varied and is differentiated by values or levels of each attribute.[35] By analysing the choices people make, the relative importance of different attributes can be determined.[35] In the case of cancer staging pathways, example attributes could include exposure to ionising radiation or number of scans required to complete staging.

Study 3 was conducted after completion of studies 1 and 2. When consenting to participate in the main Streamline trials, participants could also opt into this study. In total, 264 participants consented to take part, of whom 126 did not respond to the DCE questionnaire. The final study cohort was therefore 148 participants (66 recruited to Streamline C and 72 to Streamline L).

The DCE questionnaires were posted to participants within 2 days of them consenting to the Streamline trials and participants were asked to complete them during their staging process, returning them in a stamped addressed envelope. Participating participants were paid £20 for completing the questionnaire.

Discrete choice experiment questionnaire attributes and levels

Attributes were selected by the trial team to capture known or potential important differences between WB-MRI and standard staging pathways;[35] these were informed by findings from the participant interview study (study 1) and questionnaire study (study 2). Based in particular on the findings of these two studies, the following attributes were included: scan duration, need for the whole body and head to be enclosed by the scanner and increased cancer risk due to exposure to ionising radiation. In addition, and to align with the main outcomes of the Streamlines trials (see *Chapter 2*), the following pathway attributes were also included: pathway accuracy, scan number and time to diagnosis.

Credible levels for each attribute were chosen based on either known levels, such as scan duration, or after appropriate literature review, for example radiation exposure and scan/pathway accuracy (see *Chapter 1*).[35,56] At the time of study design, actual pathway accuracy and scan number required to reach a final diagnosis was not yet known (see *Chapters 3* and *4*).

Table 35 summarises the selected attributes and levels.

TABLE 35 Attributes and attribute levels

Attribute	Attribute levels		
Time in scanner	10 minutes	30 minutes	60 minutes
Time to reach a final diagnosis	1 week	3 weeks	5 weeks
Associated increase in cancer risk due to radiation exposure	None	1 in 1000 patients	2 in 1000 patients
Number of additional staging scans before final diagnosis	0	1	2
Accuracy for metastatic disease (%)	85	90	95
Need for whole body and head to be in scanner	No	Yes	–

Reproduced from Miles *et al.*[35] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

Discrete choice experiment questionnaire design

Of the six attributes, five had three levels and one had two levels. The total number of attribute combinations was therefore 486 ($3^5 \times 2^1$).[35] By presenting each question to participants as a binary choice set (pathway A vs. pathway B), there were a possible 235,710 choices (486 × 485). We reduced this choice set to 18 using an orthogonal fractional main effects design for pathway A.[57] Pathway B was generated by shifting the attribute level up by one category for each attribute. We did not include an opt-out or 'neither' option as participants are unlikely to choose not to undergo staging. The 18 choice sets were split into two blocks of nine, and participants were randomly assigned to complete either choice sets 1 to 9 (questionnaire A) or 10 to 18 (questionnaire B).[35] The choice sets were presented in a random order in each questionnaire. Prior to administering the DCE questionnaire, its content was reviewed and modified for clarity by the Trial Management Group (TMG), which included two patient representatives.

An example of a choice set is shown in *Figure 8*.

A range of demographic and health-related variables were also collected from participants including age, sex, highest educational qualification, ethnicity, car and home ownership, employment and marital status along with self-rated health, presence of comorbidities, and positive and negative mood (using the Positive and Negative Affect Schedule,[58] phrased to ask about current mood).[35]

Participants were also asked whether they preferred WB-MRI or standard tests ('If you had to have JUST ONE of the tests which one would you prefer?').[35]

Choice 1	Scan A	Scan B
You will be in the scanner for:	30 minutes	60 minutes
You will know your final full diagnosis after:	1 week	3 weeks
The lifetime chance of getting cancer because you had a scan that exposed you to radiation is:	1 in 1000 patients	2 in 1000 patients
The number of additional scans after this one that you will need before you know your final full diagnosis is:	2	0
The accuracy of the scans to tell you if the cancer has or has not spread is:	Correct in 95 out of 100 patients	Correct in 85 out of 100 patients
Your whole body including your head will need to be inside the scanner:	Yes	No
Based on the descriptions above, which scan would you prefer? (Tick one box only)	Prefer Scan A ☐	Prefer Scan B ☐

FIGURE 8 Example of a choice set. Reproduced from Miles *et al.*[35] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Sample size calculation

Sample size calculations are difficult for DCEs and depend on the question format, the complexity of the choice tasks, the desired precision of the results, the degree of heterogeneity in the target population, the availability of respondents and the need to conduct subgroup analyses. Researchers commonly apply rules of thumb, based on the number of attribute levels, to estimate sample size. Our calculations were based on the formula:

$$(500 \times c)\ \delta\iota\omega\ (t \times a), \quad (1)$$

where t is the number of sets of choices ($n = 9$), a is the number of scenarios to choose between in each choice ($n = 2$) and c is the largest number of levels for any one attribute ($n = 3$).[59] Thus, our target sample size was 84 in total, or 42 participants for each cancer type, which we comfortably exceeded.

Analysis

Descriptive statistics for participant characteristics were compared between colorectal and lung cancer cohorts using chi-squared tests for categorical variables and independent *t*-tests for continuous variables.

Discrete choice experiment data were analysed using an alternative-specific regression model where the outcome was test preference (scan A or B) and the variables in the equation were the individual attributes.[35] Alternative-specific constant terms were not included. We ran the model on the whole sample, as well as stratifying by recruitment to either Streamline L or Streamline C, and by participant characteristics. The relative importance of each attribute was calculated as the difference in preference weights between the best or most preferred level of each attribute and the worst or least preferred level of the same attribute.[60] Preference weights for continuous measures were calculated as the product of the coefficient for each attribute and the best and worst levels of the attribute. The resulting range for each attribute provides an estimate of the relative importance of that attribute over the range of levels included in the experiment.

We used the regression coefficients to compute marginal rates of substitution (MRS). The MRS allow direct assessment of the relative importance of each attribute by showing how much of one attribute participants are willing to trade for one unit of another attribute. It therefore enables a comparison of how important different attributes are to participants on a common scale.[35]

The choice of which attribute is used as the denominator allows different presentations of participants' trade-offs to be made. Any of the attributes can be chosen as the denominator. We calculated the MRS values using each attribute in turn as the denominator so that participants' preferences and the trade-offs could be compared in terms of (1) willingness to spend longer in the scanner (minutes), (2) willingness to wait for diagnosis (weeks), (3) willingness for an extra 1/1000 risk of cancer due to radiation exposure, (4) willingness to have an extra scan and (5) willingness for a 1% increase in accuracy for metastatic disease.[35] We also used the regression analysis results to calculate the predicted probabilities of choosing alternative pathways (e.g. based on WB-MRI) compared with a standard staging pathway. The selected default standard pathway was PET-CT plus one additional scan (Streamline L) or CT plus one additional scan (Streamline C).

We compared default staging pathways to alternative pathways with varying attribute levels based around PET-CT, CT and WB-MRI. We considered several scenarios for WB-MRI based pathways, although fixed the following attributes: (1) 60 minutes in the scanner, (2) no risk of cancer from radiation exposure and (3) requirement for the whole body and head to be enclosed. We then varied combinations of time to diagnosis, number of additional scans and accuracy of WB-MRI individually and jointly. Non-traders were included in the analysis.[35]

All data were analysed using SPSS version 24 and Stata 13.0.

Study 3 results

Of the 138 participants recruited, 128 completed all nine choice sets. The full demographics of the cohort is shown in *Appendix 74*, *Table 111*. Participants recruited to Streamline C were more likely to have educational qualifications, own a home and be in employment than participants recruited to Streamline L. They were also less likely to report comorbidities, more likely to rate their current health as good or very good and more likely to report higher levels of positive mood than Streamline L participants.

The majority of participants had undergone WB-MRI by the time they completed the DCE [113/131 participants (86%) answering the question], with no significant difference between the cohorts [Streamline C: 55/64 participants (86%); Streamline L: 58/67 (87%)].[35]

Importance of attributes

Overall, participants preferred (1) to wait less time for a diagnosis, (2) a lower dose of radiation exposure, (3) fewer additional scans and (4) greater test accuracy (*Table 36*).

Preferences across the combined cohort (Streamline C and Streamline L) were not influenced significantly by time in the scanner or the need for the whole body and head to be enclosed. However, preferences differed significantly between participants with lung cancer and those with colorectal cancer. Notably, time in the scanner significantly influenced the preferences of lung cancer participants. Both cohorts preferred

TABLE 36 Results of conditional logit regression analysis by cohort

		Participants						
		All		Streamline L		Streamline C		
Attribute	Levels	Coefficient (95% CI)	RI	Coefficient (95% CI)	RI	Coefficient (95% CI)	RI	*p*-value[a]
Time in scanner	Minutes	–0.002 (–0.007 to 0.002)[a]	–	–0.008 (–0.014 to –0.002)	0.04	0.005 (–0.002 to 0.012)[b]	–	0.01
Time to diagnosis	Weeks	–0.355 (–0.411 to –0.300)	1.42	–0.372 (–0.449 to –0.295)	1.49	–0.349 (–0.432 to –0.265)	1.40	0.70
Radiation dose	Risk of cancer (/1000)	–0.421 (–0.521 to –0.320)	0.84	–0.413 (–0.551 to –0.274)	0.83	–0.436 (–0.587 to –0.286)	0.87	0.83
Number of additional scans	Number	–0.192 (–0.299 to –0.084)	0.38	–0.179 (–0.330 to –0.028)	0.36	–0.224 (–0.382 to –0.067)	0.45	0.69
Accuracy	Percentage	0.128 (0.107 to 0.150)	1.28	0.109 (0.079 to 0.138)	1.09	0.156 (0.122 to 0.190)	1.56	0.03
Need for whole body and head to be in scanner	No	–	–	–	–	–	–	
	Yes	0.020 (–0.129 to 0.170)[b]	–	0.017 (–0.190 to 0.224)[b]	–	–0.007 (–0.233 to 0.220)[b]	–	0.88
Observations/ respondents	–	2362/138	–	1230/72	–	1132/66	–	0.02

RI, relative importance.

a *p*-values are from chi-squared tests where coefficients are equal for lung cancer and colorectal cancer participants; $p < 0.05$ indicates coefficients are significantly different between groups. The *p*-value in the bottom row is for joint test across all coefficients.

b Coefficient not significantly different from zero; all other coefficients significant at $p < 0.05$.

Different attributes do not have the same unit of change so cannot be directly compared with one another. RI is calculated for attributes with coefficients that were significantly different from zero.

Reproduced from Miles *et al.*[35] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

tests with higher accuracy, but the preference was significantly greater for participants with colorectal cancer ($p = 0.03$). For the other attributes, preferences were not significantly different between the two cohorts.

Relative importance of the attributes

Based on the relative importance weightings, for Streamline L participants time to diagnosis was the attribute valued most highly, followed by accuracy, radiation dose, number of additional scans and time in scanner (see *Table 36*). For Streamline C participants, accuracy was valued most highly, followed by time to diagnosis, radiation dose and number of additional scans.

Within each cohort, there were no significant differences in preferences according to subgroups stratified by sex, age, comorbidities, employment status, marital status and positive mood. For participants with lung cancer (but not colorectal cancer) there were significant variations when participants were stratified by home ownership, education and self-rated health (see *Appendix 75*, *Table 112*, *Appendix 76*, *Table 113* and *Appendix 77*, *Table 114*). For example, the influence of diagnostic accuracy on preferences was greater for lung cancer participants who were home owning or had higher self-rated health.

Overall, 32 out of 59 (54.2%) lung cancer participants and 45 out of 61 (73.8%) colon cancer participants who answered the question selected WB-MRI over standard scans. For Streamline L participants, those stating an overall preference for standard staging scans preferred less time in scanner and to not have their whole body and head enclosed (see *Appendix 78*, *Table 115*). There were no significant differences in attribute preferences between Streamline C participants who preferred WB-MRI and those who stated a preference for standard staging scans.

Traders versus non-traders

A total of 37% (51/138) of participants were 'non-traders', with preferences determined by a single attribute that they did not trade-off against any of the other attributes presented. The most common attributes participants would not trade were higher accuracy, faster time to diagnosis and reduced cancer risk due to scan-related radiation exposure (see *Appendix 79*, *Table 116*).

Marginal rates of substitution

Table 37 shows results of the MRS analysis. Lung cancer participants were willing to wait just over 1 extra week (MRS –1.11) in return for a 1 in 1000 reduction in the risk of cancer from radiation exposure. They were willing to wait around half an extra week (MRS –0.48) to avoid an additional scan and around one-third of an extra week (MRS 0.29) for every 1% increase in accuracy (i.e. 1.45 weeks for a 5% increase in accuracy). The willingness to wait longer for a diagnosis for a reduction in the time in scanner was negligible (MRS –0.02). These figures were broadly similar for colorectal cancer participants. For example, they were willing to wait just under half a week (MRS 0.45) for every 1% increase in accuracy (i.e. 2.25 weeks for a 5% increase in accuracy).[35]

Predicted probabilities

Figures 9 and *10* detail the predicted probabilities of choosing alternative pathways, compared with a default standard staging pathway, for Streamline L (PET-CT plus one additional scan) and Streamline C (CT plus one additional scan), respectively. Lung cancer participants were more likely to prefer a WB-MRI-based pathway (probability 0.64) if it was as accurate, required the same total number of scans and had the same time to diagnosis as the default staging pathway. If the WB-MRI pathway was more accurate, reduced time to diagnosis and/or required fewer scans than the default staging pathway, the preference for WB-MRI was even stronger. For example, the probability of choosing WB-MRI if it was more accurate than the default pathway was 0.76, rising to 0.89 if WB-MRI was more accurate, reduced time to diagnosis and meant fewer scans. The same patterns were also found for colorectal cancer participants compared with their default staging pathway.

TABLE 37 Marginal rates of substitution across all attributes

	Participants								
	Streamline L					Streamline C			
Numerator of MRS	Time in scanner: willingness to wait in scan (minutes)	Time to diagnosis: willingness to wait for diagnosis (weeks)	Radiation dose: willingness to have an additional 1/1000 cancer risk due to radiation exposure	Number of additional scans: willingness to have an extra scan	Accuracy: willingness for a 1% increase in accuracy	Time to diagnosis: willingness to wait for diagnosis (weeks)	Radiation dose: willingness to have an additional 1/1000 cancer risk due to radiation exposure	Number of additional scans: willingness to have an extra scan	Accuracy: willingness for a 1% increase in accuracy
Time in the scanner	–	–0.02	–0.02	–0.05	0.08	NS	NS	NS	NS
Time to diagnosis	–45.23	–	–0.90	–2.08	3.42	–	–0.80	–1.55	2.23
Radiation exposure	–50.20	–1.11	–	–2.31	3.80	–1.25	–	–1.94	2.79
Number of additional scans	–21.77	–0.48	–0.43	–	1.65	–0.64	–0.51	–	1.44
Accuracy	13.21	0.29	0.26	0.61	–	0.45	0.36	0.70	–

NS, not significant.

NS indicates that the coefficient on time in scanner is non-significant, so the MRS is not computed. The MRS for willingness for time in the scanner for Streamline C participants is not reported because the coefficient on time in scanner in this group is non-significant. The MRS for need for whole body and head to be in scanner is not reported for either Streamline C or Streamline L participants because in both cases the coefficient is non-significant.

Reproduced from Miles *et al.*[35] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

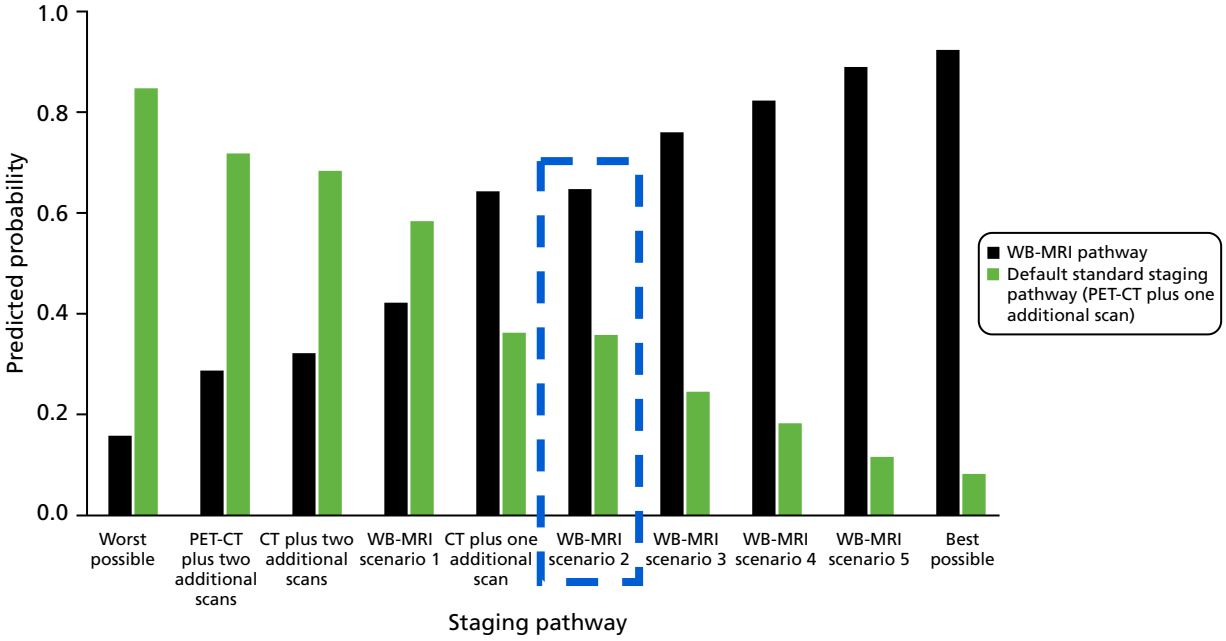


FIGURE 9 Predicted probabilities of choosing an alternative staging pathway in comparison to a default staging pathway (PET-CT plus one additional scan): Streamline L. The comparison indicated by the dashed box (WB-MRI scenario 2) is one in which WB-MRI differs from the default staging pathway according to established differences (time in scanner, exposure to ionising radiation, need for the whole body and head to be inside the scanner) but for which other attributes (time to diagnosis, number of additional scans, accuracy) are assumed to be the same between the two pathways. Reproduced from Miles *et al.*[35] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

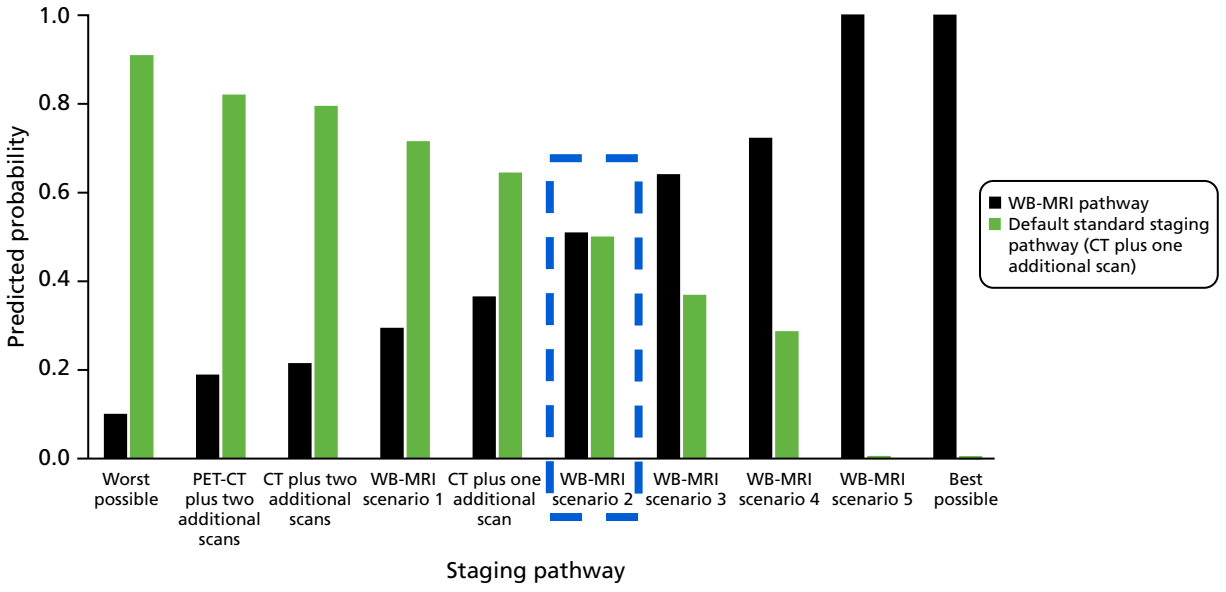


FIGURE 10 Predicted probabilities of choosing an alternative staging pathways in comparison to a default staging pathway (CT plus one additional scan): Streamline C. The comparison indicated by the dashed box (WB-MRI scenario 2) is one in which WB-MRI differs from the default staging pathway according to established differences (time in scanner, exposure to ionising radiation, need for the whole body and head to be inside the scanner) but for which other attributes (time to diagnosis, number of additional scans, accuracy) are assumed to be the same between the two pathways. Reproduced from Miles *et al.*[35] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

Description of tests in Streamline L

- Default staging pathway (PET-CT plus one additional scan) in every case: 30 minutes in scanner, 3 weeks to diagnosis, 2/1000 increase in cancer risk due to radiation exposure, one additional scan, 90% accuracy and no need for whole body and head to be in scanner.
- Worst possible test: 60 minutes in scanner, 5 weeks to diagnosis, 2/1000 increase in cancer risk due to radiation exposure, two additional scans, 85% accuracy and need for whole body and head to be in scanner.
- PET-CT plus two additional scans: 30 minutes in scanner, 5 weeks to diagnosis, 2/1000 increase in cancer risk due to radiation exposure, two additional scans, 90% accuracy and no need for whole body and head to be in scanner.
- CT plus two additional scans: 10 minutes in scanner, 5 weeks to diagnosis, 2/1000 increase in cancer risk due to radiation exposure, two additional scans, 90% accuracy and no need for whole body and head to be in scanner.
- WB-MRI scenario 1: longer scan time, no radiation, whole body enclosed, longer time to diagnosis, more scans – 60 minutes in scanner, 5 weeks to diagnosis, 0/1000 increase in cancer risk due to radiation exposure, two additional scans, 90% accuracy and need for whole body and head to be in scanner.
- CT plus one additional scan: 10 minutes in scanner, 3 weeks to diagnosis, 1/1000 increase in cancer risk due to radiation exposure, one additional scan, 90% accuracy and no need for whole body and head to be in scanner.
- WB-MRI scenario 2: longer scan time, no radiation, whole body enclosed – 60 minutes in scanner, 3 weeks to diagnosis, 0/1000 increase in cancer risk due to radiation exposure, one additional scan, 90% accuracy and need for whole body and head to be in scanner.
- WB-MRI scenario 3: longer scan time, no radiation, whole body enclosed, more accurate – 60 minutes in scanner, 3 weeks to diagnosis, 0/1000 increase in cancer risk due to radiation exposure, one additional scan, 95% accuracy and need for whole body and head to be in scanner.
- WB-MRI scenario 4: longer scan time, no radiation, whole body enclosed, quicker time to diagnosis, fewer scans – 60 minutes in scanner, 1 week to diagnosis, 0/1000 increase in cancer risk due to radiation exposure, no additional scans, 90% accuracy and need for whole body and head to be in scanner.
- WB-MRI scenario 5: longer scan time, no radiation, whole body enclosed, more accurate, quicker time to diagnosis, fewer scans – 60 minutes in scanner, 1 week to diagnosis, 0/1000 increase in cancer risk due to radiation exposure, no additional scans, 95% accuracy and need for whole body and head to be in scanner.
- Best possible pathway: 10 minutes in scanner, 1 week to diagnosis, 0/1000 increase in cancer risk due to radiation exposure, no additional scans, 95% accuracy and no need for whole body and head to be in scanner.

Description of tests in Streamline C

- Default staging pathway (CT plus one additional scan) in every case: 10 minutes in scanner, 3 weeks to diagnosis, 1/1000 increase in cancer risk due to radiation exposure, one additional scan, 90% accuracy and no need for whole body and head to be in scanner.
- Worst possible pathway: 60 minutes in scanner, 5 weeks to diagnosis, 2/1000 increase in cancer risk due to radiation exposure, two additional scans, 85% accuracy and need for whole body and head to be in scanner.
- PET-CT plus two additional scans: 30 minutes in scanner, 5 weeks to diagnosis, 2/1000 increase in cancer risk due to radiation exposure, two additional scans, 90% accuracy and no need for whole body and head to be in scanner.
- CT plus two additional scans: 10 minutes in scanner, 5 weeks to diagnosis, 2/1000 increase in cancer risk due to radiation exposure, two additional scans, 90% accuracy and no need for whole body and head to be in scanner.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

- WB-MRI scenario 1: longer scan time, no radiation, whole body enclosed, longer time to diagnosis, more scans – 60 minutes in scanner, 5 weeks to diagnosis, 0/1000 increase in cancer risk due to radiation exposure, two additional scans, 90% accuracy and need for whole body and head to be in scanner.
- PET-CT plus one additional scan: 30 minutes in scanner, 3 weeks to diagnosis, 2/1000 increase in cancer risk due to radiation exposure, one additional scan, 90% accuracy and no need for whole body and head to be in scanner.
- WB-MRI scenario 2: longer scan time, no radiation, whole body enclosed – 60 minutes' time in scanner, 3 weeks to diagnosis, 0/1000 increase in cancer risk due to radiation exposure, one additional scan, 90% accuracy and need for whole body and head to be in scanner.
- WB-MRI scenario 3: longer scan time, no radiation, whole body enclosed, more accurate – 60 minutes' time in scanner, 3 weeks to diagnosis, 0/1000 increase in cancer risk due to radiation exposure, one additional scan, 95% accuracy and need for whole body and head to be in scanner.
- WB-MRI scenario 4: longer scan time, no radiation, whole body enclosed, quicker time to diagnosis, fewer scans – 60 minutes' time in scanner, 1 week to diagnosis, 0/1000 increase in cancer risk due to radiation exposure, no additional scans, 90% accuracy and need for whole body and head to be in scanner.
- WB-MRI scenario 5: longer scan time, no radiation, whole body enclosed, more accurate, quicker time to diagnosis, fewer scans – 60 minutes' time in scanner, 1 week to diagnosis, 0/1000 increase in cancer risk due to radiation exposure, no additional scans, 95% accuracy and need for whole body and head to be in scanner.
- Best possible pathway: 10 minutes in scanner, 1 week to diagnosis, 0/1000 increase in cancer risk due to radiation exposure, no additional scans, 95% accuracy and no need for whole body and head to be in scanner.

Chapter 7 Influence of WB-MRI sequences on diagnostic accuracy

Introduction

Whole-body magnetic resonance imaging is time intensive in terms of both acquisition and reporting. This stems largely from the multiple sequences acquired as part of standard protocols. This multitude of sequences allow the radiologist to not only detect abnormalities but also crucially characterise them to maintain adequate diagnostic sensitivity and specificity. For example, sites of metastasis may be more conspicuous on one sequence than another, although using a sequence with high sensitivity but lower specificity alone risks false-positive diagnosis without other sequences to help characterise areas of potential abnormality.

Although WB-MRI protocols may differ a little according to the primary cancer and the nature and usual sites of metastatic disease, they usually consist of T1- and T2-weighted imaging supplemented with DWI.[61] Intravenous contrast-enhanced sequences are also frequently employed, particularly to improve detection and characterisation of metastasis in the liver, lung and bone.[61] The addition of intravenous contrast, however, adds to the length, cost and invasiveness of WB-MRI and for some cancer types such as prostate may be omitted.[62]

There is interest in decreasing the number of sequences acquired as part of standard WB-MRI protocols to reduce the overall scan acquisition time and, potentially, radiologist reporting time. For example, small studies have suggested that protocols using DWI instead of contrast-enhanced sequences may achieve similar sensitivity for extracranial metastatic disease in melanoma.[63] However, in a single-centre explanatory study of 203 participants with non-small-cell lung cancer, Ohno *et al.*[22] reported that sensitivity for metastasis detection of WB-MRI using only DWI sequences was 58% compared with 70% when DWI was added to other sequences including post-contrast T1-weighted image acquisition. There are no such data on WB-MRI protocol optimisation in staging colon cancer.

Given this existing evidence base, the Streamline WB-MRI protocol included whole-body axial DWI, whole-body axial T2-weighted imaging, whole-body axial or coronal T1-weighted imaging pre contrast and axial T1-weighted imaging post contrast of at least the liver, lungs and head (see *Chapter 1* and *Appendix 1*). This allowed us to investigate the impact of reduced sequence WB-MRI protocols on diagnostic accuracy.

Methods

As described in *Chapter 2*, all radiologists interpreting WB-MRI as part of the Streamline trials did so in three blocks in the following order:

1. diffusion-weighted images and non-contrast-enhanced T1-weighted images
2. diffusion-weighted images and non-contrast-enhanced T1- and T2-weighted images
3. diffusion-weighted images, non-contrast-enhanced T1- and T2-weighted images and contrast-enhanced T1-weighted images.

After viewing each sequence block (and before reviewing the next block) the radiologist completed a WB-MRI imaging booklet CRF – available on the project web page [URL: www.journalslibrary.nihr.ac.uk/programmes/hta/106801 (accessed 12 September 2019)] – documenting their findings. Items recorded on the CRFs included the location and T and N stage of the local tumour (TNM 7th edition) and the presence,

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

location and size of metastatic disease. The presence of metastatic disease was recorded for a range of soft tissue and bony anatomical sites using six confidence levels grouped into normal (confidence levels 1 and 2), equivocal (levels 3 and 4) and abnormal (levels 5 and 6). If disease presence was recorded as equivocal or abnormal (i.e. confidence level 3 or above), the size of the largest two metastatic deposits per organ was recorded, along with the additional number of deposits sized < 6 mm, 6–9 mm and ≥ 10 mm. Reporting radiologists were instructed to interpret the WB-MRI as they would in their routine clinical practice, taking into account the known morphology and characteristics of potential metastatic disease across the various sequences.[22]

Analysis

The primary outcome was the difference in per-patient sensitivity for metastatic disease based on the three combinations of sequences, against the consensus reference standard. In participants for whom sequence data were missing from sequence block 3, prior sequences were substituted. Analysis included participants with complete data only. Sensitivity, specificity and differences between sequences were analysed using the methods for the primary outcome. Influences of WB-MRI sequences on radiologists' confidence and diagnosis were summarised descriptively.

Results

A total of 187 (Streamline L) and 289 (streamline C) participants had full data for all sequences. In 10 Streamline C participants, data were missing from sequence block 3, so available MRI sequences were substituted (sequence block 2 substitution in two participants and sequence block 2 in eight participants). In Streamline L there were no missing sequence data.

Streamline C and Streamline L

Tables 38 and *39* show the sensitivity and specificity for participants with metastatic disease across the three sequence blocks according to the participant cohort.

For both Streamline C and Streamline L cohorts, the combination of T1-weighted (pre and post gadolinium contrast), DWI and T2-weighted imaging achieved significantly higher sensitivity for metastatic disease than with a simple protocol of just pre-contrast T1-weighted imaging combined with DWI. Addition of contrast-enhanced sequences to DWI and T2-weighted imaging sequences produced a small non-significant increase in sensitivity for both cancers. There was no difference in specificity across the tested sequence combinations.

TABLE 38 Streamline C: sensitivity and specificity for metastatic disease according to WB-MRI protocol sequences

	Diagnostic performance, % (95% CI)			Difference in sensitivity/specificity,[a] % (95% CI)	
Diagnostic accuracy	S1: T1- and diffusion-weighted imaging	S2: T1-, T2- and diffusion-weighted imaging	S3: T1-, T2-, and diffusion-weighted and contrast-enhanced imaging	S2 – S1: T1-, T2- and diffusion-weighted imaging – T1- and diffusion-weighted imaging	S3 – S1: T1-, T2- and diffusion-weighted and contrast-enhanced imaging – T1- and diffusion-weighted imaging
Sensitivity	63 (51 to 73)	70 (58 to 80)	72 (60 to 81)	7 (–1 to 15)	9 (1 to 17)
Specificity	84 (78 to 88)	86 (81 to 90)	85 (80 to 89)	2 (–2 to 6)	1 (–2 to 5)

S1, sequence set 1; S2, sequence set 2; S3, sequence set 3.
a Participants with all sequence data available.

TABLE 39 Streamline L: sensitivity and specificity for metastatic disease according to WB-MRI protocol sequences

Diagnostic accuracy	Diagnostic performance, % (95% CI)			Difference in sensitivity/specificity,[a] % (95% CI)	
	S1: T1- and diffusion-weighted imaging	S2: T1-, T2- and diffusion-weighted imaging	S3: T1-, T2-, and diffusion-weighted and contrast-enhanced imaging	S2 – S1: T1-, T2- and diffusion-weighted imaging – T1- and diffusion-weighted imaging	S3 – S1: T1-, T2- and diffusion-weighted and contrast-enhanced imaging – T1- and diffusion-weighted imaging
Sensitivity	42 (28 to 56)	48 (35 to 61)	52 (39 to 65)	6 (–3 to 15)	10 (1 to 19)
Specificity	82 (74 to 87)	84 (76 to 89)	86 (79 to 91)	2 (–3 to 6)	4 (–2 to 9)

S1, sequence set 1; S2, sequence set 2; S3, sequence set 3.
a Participants with all sequence data available.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Chapter 8 Interobserver agreement in WB-MRI interpretation

Introduction

Interobserver agreement is an important facet in the evaluation of any imaging technology. As noted in *Chapter 1*, many WB-MRI studies use a consensus of two readers for interpretation, which does not reflect how WB-MRI would be reported in NHS clinical practice. Where small single-site studies have tested interobserver agreement, it is generally reported as good. For example, in a small study of 34 participants with a mix of primary tumours, Grueneisen *et al.*[64] reported very high interobserver agreement ($\kappa = 0.96$) between two radiologists for lesion characterisation. In a similar study of 51 participants in which two radiologists interpreted WB-MRI based on DWI only, Paruthikunnan *et al.*[65] reported a kappa of 0.77, 0.89 and 0.93 for detection of lung, liver and bone metastasis, respectively. Blackledge *et al.*[66] tested interobserver agreement for manual segmentation of disease on diffusion weighted images, and again reported good observer agreement for DWI-defined volume (intraclass correlation coefficient 1.0). In a 140-participant comparison of WB-MRI, PET-CT and PET-MRI in non-small-cell lung cancer, Ohno *et al.*[67] reported that interobserver agreement across all three modalities for final TNM stage was high, with kappa ranging from 0.63 to 0.97. However, in the calculation of kappa in these studies, if both readers are incorrect in their interpretation, this counts as 'agreement'.

To date, almost all the literature has tested interobserver agreement between pairs of readers, many of whom are highly experienced in interpreting WB-MRI. To date, interobserver agreement has not been formally tested across a large number of readers who are typical of those who would report WB-MRI in NHS clinical practice. The Streamline trials afforded an opportunity to rectify this.

Methods

All participants recruited to the main Streamline trials gave permission for their imaging data sets to be used for the interobserver variation substudy as part of the study consent process.

Seventeen radiologists took part in the study, 15 of whom had interpreted WB-MRI as part of the main Streamline trials. Two invited Streamline trial radiologists were unable to participate in the current study and were replaced by two radiologists fulfilling the original experience criteria for the main trials (see *Chapter 2*). One radiologist who had interpreted WB-MRI as part of the Streamline trials did not undertake the required reads for the current study and could not be replaced owing to study timelines.

The WB-MRI examinations of participants recruited to the Streamline trials were shared via upload to an online viewing platform (3Dnet) (see *Chapter 2*). The platform has all the image viewing functionality of a standard PACS system but can be accessed from any PC via the internet with appropriate password credentials. The system allowed participating radiologists to access the WB-MRI examinations from any geographical location without the need to receive copies on compact disc (CD) or hard drive, thus facilitating study efficiency.

Data set selection

The interobserver study was planned during the recruitment phase of the main trials, although reads took part after recruitment was complete in order to lock the 3Dnet platform so readers could not access uploaded WB-MRI reports. Overall, 192 Streamline L and 193 Streamline C participants had complete data for trial standard and WB-MRI reads by 21 July 2016. Of these, 70 data sets were identified for each

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

interobserver study from which to select the final study data sets. Selection was undertaken to (1) enrich for participants likely to have metastases to increase statistical power, against the expected mid-trial prevalence of 21% and 22% for lung and colon cancer, respectively. Enrichment used available trial data including consensus reference data (available for 21 Streamline L and 69 Streamline C participants at this time) and also WB-MRI and standard imaging reads and participant follow-up data at 3, 6, 9 and 12 months where available. Selection was also undertaken to (2) enrich for 'stress' data sets where trial readings of WB-MRI and standard pathways disagreed on the presence of metastases, or conflicted with actual or expected consensus reference.

Case reader randomisation used random sequence generation (random uniform distribution Stata 14.2) to assign observers to data set reads. Readers were selected from two different centres or imaging hubs, and were different from the radiologists who had interpreted the WB-MRI data set as part of the main trials (i.e. no radiologist was given a data set they had already interpreted as part of the main trials). For the current analysis, participant data sets were randomly selected to ensure an equivalent prevalence of 'stress' participants in the interobserver study to that in the main trials (13% and 18% for Streamline L and Streamline C, respectively). 'Stress' participants were therefore not over- or under-represented in the interobserver substudy. The prevalence of participants with metastases, however, remained enriched (Streamline L prevalence 63%, compared with main trial prevalence 28%; Streamline C prevalence 63%, compared with main trial prevalence 23%).

Once radiologists were allocated their cases, in their own time (but within 8 weeks) they interpreted the examinations, noting their findings on a CRF based on that used for the main Streamline trials – available on the project web page [URL: www.journalslibrary.nihr.ac.uk/programmes/hta/106801 (accessed 12 September 2019)]. Data sets were either interpreted using the 3Dnet platform or could be downloaded onto the reader's preferred workstation. Each reader was allocated between 10 and 15 data sets. For the current interobserver study, reads were performed using all available WB-MRI protocol sequences.

Analysis

The primary outcome of the interobserver study was to determine agreement between two reads for identification of participants either with or without metastases, based on the consensus panel at 12 months (see *Chapter 2*). Data were analysed for the Streamline L and C cohorts separately.

Analysis related to comparisons of the two reads within a participant (WB-MRI data set) and the outcomes related to the agreement between two reads for the per-participant correct identification of metastasis (to mirror the primary outcome of the Streamline trials; see *Chapter 2*). Equivocal results (confidence scores 3 and 4) were handled as for the primary outcome (see *Chapter 2*), with sensitivity analyses treating them as either all positive or all negative (Streamline C) or all negative (Streamline L).

Interobserver variability was measured by agreement between two radiologists, grouping results by the consensus reference as positive, negative and across all participants, with 95% CI based on paired proportions. Agreement between the reads regardless of 'correctness' against the reference standard is also reported to allow more direct comparison with the existing literature. Prevalence-adjusted bias-adjusted kappa (PABAK) was also reported[68] and interpreted as follows: 0.01–0.20 (slight agreement), 0.21–0.40 (fair agreement), 0.41–0.60 (moderate agreement), 0.61–0.80 (substantial agreement) and 0.81–0.99 (almost perfect agreement).[69]

Statistical analysis was performed using Stata version 14.2.

Results

In total, 40 data sets from Streamline C and 43 data sets from Streamline L were included.

The level of agreement between both reads and the consensus reference standard according to Streamline C and L is shown in *Table 40*, and between both reads (irrespective of agreement with the consensus reference) is shown in *Appendix 80*, *Table 117*.

Overall, agreement was better for WB-MRI staging of colorectal cancer (Streamline C) than non-small-cell lung cancer (Streamline L). For Streamline C, agreement between two reads against the reference standard was 76% (95% CI 57% to 89%) and 87% (95% CI 62% to 96%) for participants with and without metastatic disease, respectively (PABAK 0.60). For Streamline L, agreement between two reads against the reference standard was 37% (95% CI 22% to 56%) and 56% (95% CI 33% to 77%) for participants with and without metastatic disease, respectively (PABAK –0.12). This in part reflects the overall low sensitivity of the two radiologist reads of 67% and 44% for the Streamline L datasets.

Agreement between reads for the presence or absence of metastatic disease regardless of agreement with the reference standard was, as expected, higher than that when the reference standard was considered (see *Appendix 80*, *Table 117)*. For the primary outcome, overall reads agreed in 35 out of 40 (88%) data sets for Streamline C and 28 out of 43 (65%) data sets for Streamline L. When equivocal lesions were considered negative, agreement was unchanged for Streamline C but improved to 32 out of 43 (74%) for Streamline L.

TABLE 40 WB-MRI interobserver variability against the consensus refence standard: Streamline C and Streamline L

	Streamline C participants ($n = 40$)					
	Participants with metastatic disease[a] ($n = 25$)			Participants without metastatic disease[a] ($n = 15$)		
Outcome	R1 (n)	R2 (n)	% positive agree (95% CI)[b]	% negative agree (95% CI)[b]	% overall agree	κ
Primary outcome[c]	19	22	76 (57 to 89)	87 (62 to 96)	80	0.60
Equivocal lesions considered positive	19	23	76 (57 to 89)	73 (48 to 89)	75	0.50
Equivocal lesions considered negative	18	21	72 (52 to 86)	87 (62 to 96)	78	0.55
	Streamline L participants ($n = 43$)					
	Participants with metastatic disease[a] ($n = 27$)			Participants without metastatic disease[a] ($n = 16$)		
Outcome	R1 (n)	R2 (n)	% positive agree (95% CI)[b]	% negative agree (95% CI)[b]	% overall agree	κ
Primary outcome[c]	18	12	37 (22 to 56)	56 (33 to 77)	44	–0.12
Equivocal lesions considered negative	10	10	26 (13 to 45)	69 (44 to 86)	42	–0.16

R1, first read; R2, second read.
a By consensus reference standard.
b Both reads agree with consensus reference standard.
c Per-participant sensitivity for metastatic disease. Equivocal results considered positive for colonic tumours and negative for rectal tumours (Streamline C). Equivocal results considered positive (Streamline L).

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Chapter 9 Discussion

Parts of this chapter are reproduced from Taylor *et al.*[25] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Parts of this chapter are reproduced from Taylor *et al.*[26] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Diagnostic accuracy

Whole-body magnetic resonance imaging is advocated as a replacement for current complex multimodality cancer staging pathways. We prospectively tested its diagnostic accuracy, influence on staging efficiency (in terms of time, number of tests and cost) and impact on the first major treatment decisions in the two most common causes of cancer deaths in the UK: colorectal and non-small-cell lung cancer. Such comprehensive evaluation of diagnostic tests, although recommended,[24] is rarely performed. We also investigated participant experience, influence of MRI sequences on diagnostic accuracy and interobserver variation in the interpretation of WB-MRI.

At the time of writing and to the best of our knowledge, Streamline C and Streamline L are the largest prospective multicentre trials of WB-MRI in cancer staging.

Multidisciplinary consensus panel review is standard methodology for diagnostic test accuracy studies for which an independent reference standard does not exist or is impossible because of incorporation bias.[23,28] We convened a number of consensus panels at recruitment sites that considered a full 12 months of participant data including all imaging, histopathology and clinical outcomes. The panels were attended by at least one radiologist external to the recruitment site to provide an independent review of all imaging, as well as clinician representation. We a priori provided the panels with criteria for diagnosing metastatic disease based on whether or not the primary tumour was left in situ, given its potential to seed metastatic disease if not removed soon after diagnosis (thereby incorrectly penalising staging investigations). To the best of our knowledge, seeding of metastatic disease from in situ primary tumours has not been considered by imaging diagnostic accuracy studies to date. In the absence of hard evidence on which to base the criteria, the trial investigators derived pragmatic definitions; any metastatic disease assumed as present at diagnosis if the primary tumour was removed within 3 months, and for primary tumours left in situ, metastatic disease appearing for the first time after 6-month follow-up was assumed to be seeded. Although such criteria are arguably relatively arbitrary, they were applied equally to both staging pathways and, given that our primary outcome was the difference between pathways rather than their absolute sensitivity, will not affect our findings.

Overall, we found that the sensitivity and specificity of the WB-MRI staging pathway for identifying participants with metastatic disease was not significantly different from standard staging pathways for either colorectal or non-small-cell lung cancer. Based on a thorough literature review at the time of study design (see *Report Supplementary Material 1*), we hypothesised that the WB-MRI pathway would have higher sensitivity than standard pathways. For colorectal cancer this was mainly based on a higher sensitivity for liver metastasis, and for non-small-cell lung cancer for brain and bone metastasis. Although in Streamline L WB-MRI did detect more brain and pleural metastasis for example, overall it held no

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

significant diagnostic accuracy advantage in either cancer. The Streamline trials are essentially the first to compare complete staging pathways and our power calculations were informed mainly by the single modality comparisons available in the literature. Indeed, it is owing to the known limitations of individual cross-sectional imaging techniques that staging pathways are multimodality to maximise sensitivity for metastatic disease. As part of our secondary analysis, we did replicate the existing literature and compared WB-MRI with PET-CT alone (Streamline L) and with chest, abdomen and pelvis CT alone (Streamline C). Interestingly, we did find WB-MRI alone had significantly higher sensitivity and lower specificity that chest, abdomen and pelvis CT in staging colon cancer (76% vs. 50% and 86% vs. 99%, respectively), although this was not significantly different from PET-CT in staging non-small-cell lung cancer. The Streamline trials clearly demonstrate that it is overly simplistic to compare single imaging modalities and that true clinical practice can be properly tested only by comparing complete staging pathways. For Streamline L, we recruited sufficient numbers of participants with metastatic disease to meet the requirements of the power calculation. However, for Streamline C, the prevalence of participants with metastatic disease according to the 12-month consensus reference (23%) was lower than anticipated (40%) and the number of participants with metastasis was less than that suggested as required by the pre-study power calculation. The UK bowel screening programme has reduced the prevalence of advanced disease below that of historical cohorts,[70] which in part explains the lower prevalence of metastasis, and participants were not preselected (e.g. based on age or primary tumour stage) to artificially boost the prevalence (other than exclusion of polyp cancers). Instead, Streamline C recruited consecutive eligible participants who required staging investigations, specifically to mimic clinical practice. The IDMC did perform an interim analysis to look at both prevalence and other assumptions in the sample size calculation. The decision was made to recruit to the number of evaluable participants in the original trial design, accounting for the increased withdrawal rate. The number of participants recruited was not adjusted to account for the lower prevalence, as this interim analysis showed no evidence that a difference in diagnostic accuracy between pathways would be identified. The IDMC confirmed that the trial should be continued and would include a sufficient number of participants for results to guide current practice. The study was designed to have sufficient power to detect a clinically relevant sensitivity difference of 10% between WB-MRI and standard staging pathways (predominantly chest, abdomen and pelvis CT). Interestingly, we did find that WB-MRI alone has greater sensitivity than chest, abdomen and pelvis CT alone. Overall and to the best of our knowledge, Streamline C is by far the largest trial of WB-MRI in staging pathways colon cancer to date.

In recent years, several meta-analyses have been published concerning the accuracy of WB-MRI in cancer staging. Many combine disparate tumour types and, although several consider lung cancer in isolation, they typically focus on one site of metastatic disease, usually bone or lymph node (see *Chapter 1*). Those combining different primary tumours generally concur with our finding that the sensitivity of WB-MRI is similar to conventional tests such as PET-CT. For example, Xu *et al.*[9] analysed nine studies, including 1070 participants with a range of primary cancers such as breast, lung, melanoma, nasopharyngeal and colorectal. Overall sensitivity for participants with metastasis was 86% (95% CI 70% to 94%) for WB-MRI and 85% (95% CI 68% to 94%) for PET-CT. Li *et al.*[11] also reported identical sensitivity (90%) for WB-MRI and PET-CT in their 2014 meta-analysis across a range of primary tumours and metastatic sites. Those meta-analyses concerning lung cancer alone have mainly considered bone and nodal metastasis. Accuracy of bone metastases detection was evaluated across a range of studies and participant cohorts by Qu *et al.*,[15] who reported that WB-MRI had lower sensitivity than PET-CT [77% (95% CI 65% to 87%) (in three studies of 252 participants) vs. 92% [95% CI 88% to 95%] (in 17 studies, of which only two included WB-MRI), respectively]. Of note, only Li *et al.*[11] have directly compared the sensitivity and specificity of WB-MRI and PET-CT and reported that there was no significant difference; the other meta-analyses and systematic reviews report only stand-alone data from each test. Crucially, data were not reported for studies comparing the imaging methods in the same participants (i.e. direct within-participant comparison). Furthermore, no study presents results stratified by the largest lesion size, a key determinate of sensitivity in Streamlines L and C.

Overall, for both cancers we found the sensitivity of WB-MRI (and standard) pathways to be lower than that reported in the literature. There are several probable reasons for this observation. Most studies to date have been small single-site explanatory studies using a consensus of two experienced radiologists for

WB-MRI interpretation. Such study design inflates the apparent diagnostic accuracy of any new technology. Conversely, we tested WB-MRI as it would be used if disseminated in the NHS; images were acquired across a range of MRI platforms and interpreted by radiologists representative of those who would report the examination in clinical practice. It should be noted that detailed quality assurance review of WB-MRI data sets was reassuring: just under half were judged to be of optimal quality, and although 43% were suboptimal they remained fully diagnostic by our quality assurance criteria. Only 64 data sets were scored as degraded and many (23) of these were because gadolinium could not be administered, meaning they still had T1-, T2- and diffusion-weighted imaging sequences available (which achieve similar sensitivity as per our substudy in *Chapter 7*). There were also few perceptual errors noted by the consensus panels on retrospective review of the imaging that influenced the primary outcome (per-patient analysis). Interestingly, however, there were more WB-MRI perceptual errors in Streamline L than in Streamline C. This probably reflects the conspicuity of metastases according to their location. Most metastases in Streamline C participants were in the liver or lung, with relatively few perceptual errors for either organ. This was mirrored in Streamline L, with few perceptual errors in the lung or liver. Conversely, in Streamline L many metastases were located in the bone, where they are more prone to being small, subtle and difficult to characterise. Indeed, there were perceptual errors in seven of 15 bone metastases. It is important to stress that although the consensus panels suggested that some missed lesions were perceptual errors, this does not necessarily imply that they should have been prospectively reported. Many lesions although visible in retrospect and with the advantage of follow-up imaging are very difficult to detect prospectively, particularly if reasonable specificity is to be maintained. Overall, in terms of the primary outcome, there was no material difference in the number of perceptual errors in WB-MRI and standard staging pathways. This suggests that WB-MRI was interpreted with at least the same competency as standard imaging, justifying the stipulated experience criteria for radiologists. We of course acknowledge that diagnostic sensitivity of WB-MRI may be higher if interpreted by highly experienced radiologists with large reporting volumes. However, to restrict interpretation to these individuals would not be representative of a national workforce. We also blinded radiologists to the patient history and, in the case of WB-MRI, to contemporaneous imaging, which does not mirror usual clinical practice but was necessary to isolate diagnostic test accuracy as far as possible. Indeed, we used novel cloud-based imaging interpretation software for WB-MRI to ensure full blinding of WB-MRI reporters to standard imaging and to WB-MRI for those reporting standard imaging. Again, to the best of our knowledge this approach is novel and added to the integrity of the trials. For both trials our consensus reference standard considered a full 12 months of participant follow-up, and several participants 'developed' metastasis during this follow-up period according to our criteria that was not visible even in retrospect on any imaging modality (e.g. six participants in Streamline C). Such disease, almost certainly present at diagnosis, is currently beyond the resolution of cross-sectional imaging and explains the large number of colon cancer participants who relapse within 2 years of attempted curative resection and the relatively high rate of 'futile thoracotomy' for apparently early-stage non-small-cell lung cancer.[71] Most diagnostic accuracy studies do not include such a prolonged period of participant follow-up, and, as such, the Streamline trials provide a realistic reflection of cancer staging accuracy.

The trials were not powered to detect differences between pathways for metastasis detection in specific organs, but although the ability of WB-MRI to detect lung metastasis has been questioned,[21] we found no strong evidence that it was inferior to standard pathways for either cancer type. Although WB-MRI pathways diagnosed two additional participants with brain metastasis in Streamline L, we can conclude that WB-MRI pathways are unlikely to significantly reduce the rates of unexpected relapse. In the case of Streamline L, we excluded participants with locally advanced or metastatic disease on their diagnostic chest CT (which usually includes the lower neck, liver and adrenal glands); such participants generally undergo treatment without curative intent and detecting additional disease sites has limited therapeutic impact. Such exclusion is unusual in the literature. In our cohort, metastases were therefore either occult on diagnostic chest CT or at remote sites outside the chest and upper abdomen. This pre-selection of participants to only those in true clinical need of detailed staging probably reduced overall diagnostic accuracy of both standard imaging and WB-MRI in comparison to most trials in the literature in which such participants with metastatic disease at initial diagnosis are included. In both trials, as would be expected, sensitivity for metastasis $\geq$ 1 cm in size was much higher than for metastasis <1 cm in size. In the case of Streamline C, most (48/68) participants with metastatic disease had liver involvement, and the dependence of imaging accuracy on lesion size is well established; in

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

their 2010 meta-analysis, Niekel *et* al.[72] reported that sensitivity of even dedicated liver MRI is frequently < 50% for liver metastasis < 1 cm in size. It is important to differentiate WB-MRI cancer staging protocols from those specially designed to stage the liver. Our WB-MRI protocol complied with accepted international standards,[62] including DWI and post-gadolinium-enhanced sequences, but by necessity had to compromise, for example on slice thickness, to ensure reasonable scan times. A recent meta-analyses suggests 93% sensitivity for dedicated liver MRI protocols with liver specific contrast agent, albeit with much heterogeneity between studies.[73] As MRI technology evolves, addition of sequences following liver-specific contrast agents to standard WB-MRI protocols will be increasingly feasible, although would add to cost.

The Streamline trials were prospective real-world evaluations of staging pathways and the findings certainly raise the possibility that diagnostic accuracy studies in the radiological literature tend to inflate the sensitivity of cross-sectional imaging owing to participant selection bias, reading paradigms and applied standards of reference. Indeed, recent data suggests the assumed accuracy of standard staging in non-small-cell lung cancer is overly optimistic,[74] reflected in the poor overall 5-year survival rate.

Overall, on a per-participant basis there was a moderate number of equivocal lesions (Streamline C, 11 and 21 participants for WB-MRI and standard pathways, respectively; Streamline L, two and six participants for WB-MRI and standard pathways, respectively). The number of equivocal lesions was higher for WB-MRI as a stand-alone test (Streamline C, $n = 34$; Streamline L, $n = 19$). WB-MRI generated several additional staging investigations, probably as a result of equivocal findings, and these additional tests clarified their nature, resulting in high specificity for the complete WB-MRI pathway. Again, this underlines the principle that it is insufficient to compare stand-alone staging investigations, and that only full staging pathways are relevant. In general, as part of the primary analysis we treated equivocal imaging findings as positive given their impact on patient management, for example generating additional tests. On the advice of the IDMC, however, equivocal findings were treated differently in the Streamline C cohort depending on whether the primary tumour was rectal or non-rectal (in the case of rectal cancer they were considered negative). The IDMC rationale was that many participants with rectal cancer undergo long-course chemotherapy radiation and then restaging, giving time for equivocal lesions to 'declare themselves'. However, we performed a sensitivity analysis for both Streamline C and Streamline L, and there remained no significant difference between WB-MRI and standard staging pathways when equivocal findings were either considered all positive or all negative.

As part of the secondary analysis we also investigated the agreement of pathways for local T and N staging against the consensus reference standard. Were WB-MRI pathways to replace current standard pathways, their ability to stage the tumour locally as well as determining metastatic status is of importance. We found that for Streamline C, the WB-MRI staging pathway had potentially lower agreement for T stage than standard staging (54% vs. 60%, respectively, although this difference was not statistically significant) and that there was no significant difference in nodal staging agreement or sensitivity to detect nodal disease (76% with both imaging pathways). MRI is promising in its ability to assign the T stage of extrarectal colonic cancer,[75] but as yet there is no firm evidence that it will exceed the sensitivity of CT. WB-MRI protocols must compromise on slice thickness to ensure that protocols can be completed in a reasonable time. Local staging of colonic tumours would probably be improved by high-resolution T2-weighted images through the primary tumour (akin to rectal cancer staging protocols), although this would require planning by radiographers and increased scan time. In the case of Streamline L, agreement for T stage was no different between WB-MRI and standard pathways, but agreement for N stage was higher in the standard pathway (75% in the standard pathway vs. 65% in the WB-MRI pathway). Sensitivity for nodal disease was also significantly higher [84% vs 71%, respectively, a 13% (95% CI 3% to 22%) difference]. For MRI alone, the sensitivity to detect nodal disease was 67% (95% CI 56% to 77%). In their 2016 meta-analysis, which included 1122 participants, Peerlings *et al.*[18] reported that thoracic MRI had 87% (95% CI 78% to 92%) sensitivity for per-patient nodal staging in lung cancer, although Shen *et al.*[19] reported a sensitivity of 68% (95% CI 63% to 73%), which is more consistent with our findings. Indeed, in a study of 113 participants with a histological reference standard, Kim *et al.*[76] reported that morphological MRI observations had a per-patient sensitivity of 68% and a per-nodal station sensitivity of 53%, again very similar to our findings. To date, there is little evidence that WB-MRI alone

will increase the sensitivity for nodal disease compared with PET-CT. For example, in a recent meta-analysis, Wu *et al.*[20] reported MRI had a sensitivity of 72% (95% CI 63% to 80%) compared with 75% (95% CI 68% to 81%) for PET-CT, although only three studies included MRI, of which just one had a direct comparison with PET-CT. Although MRI holds promise in nodal staging of non-small-cell lung cancer, our data suggest that it is inferior to standard pathways, particularly PET-CT. The impact of this will probably be tempered by the increasing use of EBUS to both diagnose and stage lung cancer[77] (for the most part EBUS was considered a diagnostic rather than staging test in the Streamline trials). As discussed in *First major treatment decision*, this reduction in N stage agreement and sensitivity for the WB-MRI pathway did not appear to adversely influence patient management.

First major treatment decision

The Streamline trials include a novel trial design where the first patient management decision was recorded during the actual participant pathway. An important finding of the Streamline trials was that the level of agreement between treatment decisions based on the WB-MRI pathway and both the final MDT treatment decision and retrospective consensus panel optimal treatments was essentially identical to that between standard staging pathways and both the final MDT treatment decision and retrospective consensus panel optimal treatments.[24] This suggests that the lower sensitivity of WB-MRI pathways for local staging (particularly N stage in non-small-cell lung cancer) did not materially affect patient management, even in retrospect. To date, there has been little study on the impact of WB-MRI on treatment decisions. For the current study, we grouped the range of possible decisions into four major categories for Streamline C and two major categories (curative and non-curative intent) for Streamline L. Such grouping was necessary for analysis. Although the contemporaneous MDT first major treatment decision is arguably clinically the most relevant when assessing staging pathways, it may ultimately be incorrect if WB-MRI and/or standard tests are misleading and those findings were acted on by the MDT. We countered this by also collating the optimum retrospective treatment decision based on the outcome of participants at 12 months. Perhaps not surprisingly, agreement of both pathways was higher with the MDT treatment decision than with the retrospective consensus decision, as the latter considers follow-up data, which would have influenced initial treatment had it been known at the time, for example undiagnosed metastatic disease and histopathological staging of resected tumours. For example, in 22 participants with (non-rectal) colon cancer, where both WB-MRI and standard pathways agreed with the MDT treatment decision that primary surgery without chemotherapy was recommended the consensus panel suggested that additional chemotherapy would have been optimal, based on consideration of all follow-up data at 12 months. Of course, such retrospective optimal treatment decisions are theoretical as much information would often not be available at the time of the first major treatment decision. Overall, however, that the WB-MRI staging pathway does not lead to a greater number of incorrect contemporaneous or retrospective optimal treatment decisions than standard pathways is an important finding of the Streamline trials and suggests that implementation would not have a detrimental impact on patient outcomes compared with current practice. As noted in *Main trial strengths and limitations*, a limitation of our methods is that the MDT was unblinded to the standard pathway tests and treatment decision prior to documenting the decision based on the WBI-MRI pathway. However, it was important that the same individuals made the decision based on WB-MRI to avoid contamination by the variability in treatment decisions due to differing opinions between different clinicians.

In summary, the Streamline trials provide, to the best of our knowledge for the first time, a realistic assessment of the accuracy of WB-MRI and standard staging pathways in NHS clinical practice. We can conclude that WB-MRI pathways achieve similar diagnostic accuracy for identifying participants with metastatic disease as current multimodality standard staging pathways, and result in very similar primary treatment decisions.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Staging pathway efficiency

Another important aspect of cancer staging pathways rarely addressed in the literature is pathway efficiency, that is, the time to complete staging, the number of tests required and overall cost-effectiveness. The Streamline trials identified that the WB-MRI pathway was more efficient than standard pathways. For both cancer types, the WB-MRI pathway reduced the time taken to complete staging (by median 6 days for non-small-cell lung cancer and median 5 days for colorectal cancer). The time saving was greater for participants with metastatic disease. Although we were able to measure the actual time taken to complete standard staging pathways, by necessity we modelled the timing of the WB-MRI staging pathway. However, this was based on real test waiting times collated from recruitment sites as part of the trial, and so is probably an accurate reflection of staging times, assuming local access to WB-MRI (which had been agreed upfront by imaging hubs). We defined the end of staging as the date that the last staging investigation was performed, as this was a robust data point easily collected. However, we did also record the length of time between this last staging investigation and the final first major MDT treatment decision. This added a median 4 (95% CI 3 to 5) days for streamline C and 6 (95% CI 6 to 7) days for streamline L. Prolonged staging pathways increase participant anxiety[8] and time to first treatment is an important yardstick in assessing the quality of hospital cancer services. For example, accelerated diagnostic pathways in participants with lung cancer are associated with a more rapid reduction in stress levels over time.[78] Furthermore, timeliness of lung cancer treatment is recognised as a care quality indicator, and reducing time to treatment decisions by just 2 weeks is associated with improved survival.[77] It is as yet unknown if reduced staging times will have any meaningful effect on colorectal cancer outcomes, but reduction in participant anxiety would perhaps be expected given the data from other cancer types.

The number of tests was also lower in the WB-MRI pathway for both cancers (significantly so for colorectal cancer). We specifically excluded diagnostic tests, assuming that these would be common to both pathways, and compared the tests undertaken as part of staging after a proven or highly suspected cancer diagnosis. One of the potential advantages of WB-MRI is its ability to reduce the number of tests required to fully stage the patient, thereby simplifying the staging pathway. The Streamline trials show this advantage to hold true in the case of colorectal and non-small-cell lung cancer. Indeed, WB-MRI generated relatively few additional tests (21 for Streamline C and 45 for Streamline L). We asked the MDT to state whether or not, after reviewing the WB-MRI findings, they would request additional tests to complete staging. These decisions were therefore made contemporaneously by the same clinical team that defined the content of the standard staging pathway, and, outside a RCT design, this is the most robust way to define the content of staging pathways utilising initial WB-MRI. Unsurprisingly, the vast majority (97%) of participants recruited to Streamline L underwent PET-CT as part of standard staging, although just 14% would have required PET-CT as part of the WB-MRI pathway. As discussed in *Cost-effectiveness*, this was a major component of the cost savings of the WB-MRI pathway. For participants recruited to Streamline C, WB-MRI generated relatively few additional tests and just six dedicated pelvic MRI scans. Overall, 130 recruited participants had rectal cancer, for whom it would be expected that pelvic MRI would usually be performed for local staging. Indeed, 120 participants underwent a dedicated pelvic MRI as part of standard staging. It is of interest that the MDT were for the most part content with the local staging provided by the WB-MRI protocol, although this included relatively thick standard axial T2-weighted images, rather than high-resolution T2-weighted imaging angled to the plane of the tumour (which forms the bedrock of dedicated rectal cancer staging protocols). In Streamline C, the agreement for T and N staging of WB-MRI as a stand-alone investigation was lower than that for standard pathways, which (as would be expected) suggests that dedicated rectal staging MRI has superior accuracy than WB-MRI protocols. However, as noted in *First major treatment decision*, treatment decisions were not adversely influenced by the inferior local staging provided by WB-MRI pathways. This suggests that for most participants the staging information provided by WB-MRI (and additional tests generated) was sufficient to guide the first major treatment decision. A similar observation can be made for the need for dedicated liver MRI, utilised in 12% of participants as part of standard pathways.

Overall, the WB-MRI pathway was more efficient that standard pathways, reducing both the time needed to complete staging and the number of tests performed. As discussed in *Participant experience*, with these attributes, participants generally prefer the WB-MRI pathway to standard staging pathways.

Cost-effectiveness

There was in excess of 90% concordance between the standard staging pathways and the WB-MRI staging pathway for staging of both non-small-cell lung cancer and colorectal cancer. As a consequence, and based on our study protocol,[23] the economic analysis focused solely on the cost of the treatment decision pathways for both cancers. We observed that the mean cost per patient in each trial was lower for WB-MRI pathways than for conventional staging pathways, and that the 95% bootstrapped CIs did not overlap, leading us to suggest that the differences in costs are significant. The WB-MRI staging pathway costs in Streamline C were around three-quarters of the standard staging costs, and the WB-MRI pathway staging costs in Streamline L were around half the standard staging pathways costs. The staging costs for non-small-cell lung cancer are more expensive than for colorectal cancer. Assuming that there are no other differences in costs or effects on patient management of using the different staging pathways, our conclusion is that WB-MRI is more cost-effective than standard staging in both cancers. We note that an added consideration could be that the use of WB-MRI could have a different impact on participants in the long term in that, theoretically, it could be less harmful for them than the standard investigations (e.g. owing to no radiation exposure). However, no data could be collected on this as all participants also received the standard battery of tests, so it would be difficult to isolate the impact of reduced radiation exposure. Our costs were also based on standard NHS reference costs, as this was considered the most robust and fair way of comparing pathways without undue influence from local tariffs that may differ markedly between hospitals. For example, in the case of Streamline L, the increased cost of standard pathways was mainly driven by the use of PET-CT. Our costings maintained a 2 : 1 ratio of the cost of PET-CT and WB-MRI, which is consistent with local tariffs at most participating hospitals, suggesting our comparisons were fair. Finally, we note that the WB-MRI pathway required fewer tests than the standard pathway. As well as affecting the cost difference, this may also release capacity in imaging departments, potentially allowing more participants to be scanned, and theoretically reducing waiting times. The economic impact would depend on the availability and capacity of imaging equipment and staff, together with the clinical demand for the imaging platforms (overall, not just for lung and colorectal cancer participants). Consideration of this question would require a wider analysis, considering these variables, but would be of interest.

Main trial strengths and limitations

A major strength of our study is its pragmatic design. We recruited from a representative range of district general and teaching hospitals, with all imaging performed and interpreted according to usual local protocols, to increase the generalisability of our results. The radiologists interpreting WB-MRI were representative of those who would do so in daily NHS practice. We specifically avoided using a small number of highly experienced radiologists as they do not represent the national workforce. We utilised MDT meetings to mirror exactly how patient care is delivered in the NHS. Doing so captured the entirety of standard pathways, including contemporaneous treatment decisions for both staging pathways. To facilitate this, we used a novel cloud-based image repository to maintain blinding and control MDT access to WB-MRI until the appropriate time in the decision-making process. We were then able to model the content and timing of the WB-MRI staging pathway, and the potential impact on decision-making, which is a major strength of the Streamline trials. Conversely, previous research usually reports head-to-head comparisons between single imaging platforms, failing to capture pathway complexity. Previous studies have not measured staging pathway times or effects on treatment decisions. We tested the alternative pathways in the same participants, who were representative of those with the condition, which is the best methodology when testing diagnostic accuracy. By including the standard staging pathway in its entirety, we were able to fully evaluate the impact of replacement with WB-MRI. Such data are crucial to policy makers such as the NICE, and under-reported in the literature.[79] To the best of our knowledge, our trial design is unique.

The trials do have limitations in addition to those discussed in previous sections. We blinded radiologists to patient history and, for WB-MRI, to contemporaneous imaging. This does not mirror daily practice but was necessary to isolate diagnostic test accuracy as far as possible within a pragmatic setting. Participants were

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

representative of those undergoing staging in daily practice, although we did exclude pregnant women and participants with contraindications to MRI. We modelled timing of the WB-MRI staging pathway on real waiting times collated from recruitment sites during the trial and imaging hubs that had agreed they had capacity to perform WB-MRI. Treatment decisions based on WB-MRI pathways were made after the MDT was unblinded to all standard imaging tests, which may introduce bias. However, this was unavoidable if the full complexity of standard staging pathways was to be captured without interference from WB-MRI findings, and treatment decisions (and changes) could be recorded contemporaneously. Furthermore, alternative pathway agreement with a retrospective optimal treatment based on 12 months of follow-up data remained essentially identical. Our costings reflect an English NHS perspective and may differ in other settings, but full listing of the tests ordered for each pathway allows costing for alternative economic settings. It is acknowledged that access to MRI platforms to perform WB-MRI examinations may be limited in many health-care settings, which would influence the time of staging pathways. However, our data suggest that investment in increased provision would ultimately reduce the cost and complexity of colorectal and non-small-cell lung cancer staging.

Participant experience

We evaluated the comparative experience of the WB-MRI staging pathway and standard staging pathways via a series of three connected studies.

Study 1[33] was a qualitative interview study. In general, participants reported that the challenge of WB-MRI was different from standard investigations, although the extent of this challenge varied considerably between individuals. Of note, among the cohort interviewed just four participants requested premature termination and all were prepared to attempt a repeat WB-MRI in the future. To deal with the challenge of undergoing WB-MRI, participants adopted a variety of coping strategies. These mainly focused on the physical and emotional experiences while undergoing the examination, as well as focusing on beliefs that bolstered motivation to complete the scan.[33] Some strategies dealt with the immediate demands imposed by the scan, whereas others addressed the contextual threat of what the scan might reveal about their underlying condition. We found that hospital staff were often important in supporting participants; positive interactions bolstered coping whereas negative interactions had the opposite effect. For example, verbal contact from staff during the scan helped distract some participants from the physical challenges of being in the scanner and provided emotional reassurance. However, this only appeared effective if a rapport was established that was initiated at the start of the appointment before scanning started. If that trust was absent, staff interactions could increase participants' vulnerability.

Our data concur with previous qualitative work investigating patient experience during MRI scanning,[43,44] although ours is the first to specifically investigate WB-MRI. Comparable with our findings, these two previous studies report that many patients find MRI a unique experience, with references to isolation within the scanner and of 'being in another world'.[44] The challenges described by participants in our study also feature in these previous reports, particularly the enclosed space, scanner noise, physical discomfort, lying in the scanner and long scan duration. The benefit of staff support is similarly conveyed. The themes identified by our participant interviews also concur with previous quantitative work, particularly the negative impact of scanner noise and confinement.[37,41] We also found that comorbidity influenced tolerance, although unlike previous research[41] we also found that prior experience of MRI increased coping ability, even if this experience was relatively negative.

An important finding of our study was the greater impact of physical comorbidity on patient tolerance of WB-MRI in participants with lung cancer than in those with colorectal cancer. For example, musculoskeletal problems and difficulties with breath holding due to pre-existing chest disease were frequently cited. Participants recruited to the Streamline trials were either suspected to have or newly diagnosed with underlying cancer and this also affected their experience of WB-MRI. Some, for example, feared WB-MRI would reveal cancer dissemination and several noted that the MRI machine magnet resembled a coffin or sarcophagus. Carlsson *et al.*[43] also reported that fear was greater when MRI was used to confirm suspicions of serious illness.

Although the impact of comorbidities and the emotional vulnerability of cancer patients should be acknowledged, it is important to note that many participants found WB-MRI to be acceptable, and most tolerated scanning to completion.

Overall, study 1 suggested that good staff communication, both prior to and during the scan, is undoubtedly beneficial, reducing anxiety and boosting coping efforts by acting as a source of distraction, motivation and emotional reassurance.[33] Staff training to build rapport and facilitate relaxation can reduce the MRI non-completion rate and increase patient satisfaction.[80] Participant experience during WB-MRI is variable, but most employ some form of coping strategy and complete the scan.

Study 2[34] built on the findings of study 1, which suggested that for some participants WB-MRI is a greater challenge than standard staging scans. We quantified participant experiences during staging pathways using questionnaires and in particular compared WB-MRI with CT (Streamline C participants) and with PET-CT (Streamline L participants).

Overall, study 2 suggested that in general participants tolerate WB-MRI well (96% rate it as very or fairly acceptable) and, when quantified from one to seven, absolute discomfort and worry levels were low and satisfaction was high. However, when compared with standard staging scans, the burden of WB-MRI was significantly greater. This was particularly apparent when Streamline C participants compared WB-MRI with CT, but also held true when Streamline L participants compared WB-MRI with PET-CT. Of the standard staging scans, the burden of CT was less than that of PET-CT. However, as noted in study 1 and confirmed by study 2, comorbidities in the lung cancer participant cohort may have influenced their tolerance of PET-CT, so a direct comparison with Streamline C participants undergoing CT has limitations. Interestingly, Adams *et al.*[81] reported that lymphoma participants in general describe a better experience during WB-MRI than during CT, although the latter required ingestion of oral contrast and an intravenous contrast injection. An intravenous contrast injection was part of the Streamline WB-MRI protocol (see *Chapter 2*).

A high level of baseline distress was associated with subsequent higher WB-MRI burden, as was additional comorbidity and high social deprivation (in the unadjusted analysis). It is known that the presence of comorbidity is associated with suboptimal treatment and decreased cancer survival.[82] Furthermore, higher social deprivation is linked with worse outcomes, particularly in lung cancer.[83] Our observation that both of these factors increased the burden of WB-MRI is certainly interesting, and further research should explore their association with WB-MRI image quality, which may affect diagnostic accuracy.

Study 2 found that although WB-MRI is generally reasonably well tolerated, it induces more participant burden than CT and PET-CT. Study 3[35] used a DCE to investigate which attributes of staging pathways are most important to participants, whether or not they are willing to trade between these attributes and what criteria would need to be met for a WB-MRI staging pathway to be preferred to standard staging. We found that both Streamline L and Streamline C participants rate time to diagnosis, diagnostic accuracy, cancer risk from radiation exposure and the number of additional scans as important pathway attributes. In addition, for participants with lung cancer, time in scanner was also rated as important. Streamline C participants rated diagnostic accuracy as the most important pathway attribute, whereas Streamline L participants rated time to diagnosis as most important (although diagnostic accuracy had a greater influence on the preferences of lung cancer participants who were home owning or had higher self-rated health). That time in the scanner is important for participants with lung cancer (but not colon cancer) concurs with the findings in study 1 and 2 in which comorbidities (more prevalent in lung cancer participants) influence the burden of WB-MRI. It is also interesting that both participant cohorts rated exposure to ionising radiation as important. The doses received for staging scans such as CT and PET-CT increase the theoretical cancer risk by a relatively small percentage in a cohort of participants who already have a probable cancer diagnosis and the prognostic implications this entails. For example, radiotherapy imparts much higher doses to participants than staging investigations. It is possible that improved participant education about the potential influence of diagnostic radiation exposure on their long-term

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

prognosis would reduce the weighting of this attribute, but nevertheless it is important to acknowledge that it is of importance to participants regardless of their underlying diagnosis.

Both study 1 and study 2 showed that claustrophobia was problematic for some participants. However, in study 3[35] the need to enclose the whole body and head did not influence scan preferences when balanced against other test attributes. By definition, participants recruited to the Streamline trials were willing to undergo WB-MRI, so the prevalence of pre-existing claustrophobia may be lower than the general population. Furthermore, the DCE methodology looks at relative weighting of attributes to each other rather than absolute levels, and our study 3 sample may not have been fully representative of the whole cohort. It is of note that Streamline L participants who chose standard staging over WB-MRI (when given a binary choice) preferred less time in the scanner and not to have their whole body and head enclosed compared with those preferring WB-MRI.

We found that around two-thirds of participants were willing to trade attributes against each other. For example, both Streamline C and Streamline L participants were willing to wait additional time and/or undergo additional scans to achieve greater pathway diagnostic accuracy. Of the 'non-traders' (i.e. those not willing to compromise on a particular attribute), diagnostic accuracy was the most common 'fixed' attribute.

The final part of study 3 established the criteria that would need to be met for the WB-MRI staging pathway to be preferred to standard staging. Participants from both cohorts were more likely to prefer the WB-MRI pathway compared with default staging as long as it was as accurate and resulted in the same scan number and time to diagnosis. As described in *Chapters 3* and *4*, the results of Streamline trials suggest that the WB-MRI pathway is as accurate as standard staging pathways, but more efficient, actually reducing time and number of tests required to complete staging. Overall, the DCE in study 3 suggests that, given these findings, participants in general would prefer the WB-MRI staging pathway than standard pathways.

Our set of three participant experience studies has limitations. As noted above, by definition participants recruited to the Streamline trials were willing to undergo WB-MRI and so may not be representative of the population in general. Furthermore, the majority (86%) of participants had already undergone WB-MRI when completing the DCE questionnaire, meaning that we were unable to investigate the influence of prior experience of WB-MRI owing to the small number who had not yet experienced it. Future work should investigate what attributes WB-MRI would need to possess to appeal to people who are more reluctant to undergo it as part of cancer staging. The sample size for all three studies was based on the established criteria (study 1) and on power calculations (studies 2 and 3). However, power was limited to detect differences within each cancer type and our null findings for some predictors (e.g. age, sex and cancer type) may be due to lack of statistical power to detect small effects. The proportion of participants who completed the study 2 questionnaires was arguably quite low, at around 50%. However, this is in line with postal survey completion rates observed in other similar studies.

In conclusion, the majority of participants are able to tolerate WB-MRI and view it as an acceptable investigation. However, overall burden is greater than standard staging investigations, with the enclosed scanner space, scanner noise, discomfort and scan duration particularly problematic. Most participants develop coping strategies, which are aided by positive staff interactions both before and during the scan. Participants with high levels of existing stress and with existing comorbidities have greater burden during WB-MRI, and participants with lung cancer (but not participants with colon cancer) view time in the scanner as a relatively important attribute. Participants rate time to diagnosis, diagnostic accuracy, cancer risk from radiation exposure and the number of additional scans as important pathway attributes and around two-thirds are willing to trade attributes. Based on the findings of the Streamline trial (see *Chapters 3* and *4*), a DCE suggests that the majority of participants would prefer the WB-MRI staging pathway to standard staging pathways for both cancer types.

Influence of WB-MRI sequences on diagnostic accuracy

Although there is reasonable consensus as to the basic WB-MRI sequence protocol,[61] optimised approaches differ between cancer types, depending on the nature of metastatic disease and the specific organ predilection, for example bone or soft tissue. The routine use of intravenous contrast for example is currently not advocated in prostate cancer staging[62] but probably has utility in staging myeloma.[84] Ultimately, the optimal scan protocol is a compromise between acquisition time, cost, radiologist reporting time and diagnostic accuracy. Recent data on potential retention in the brain have also highlighted the need for judicious use of gadolinium.[85]

Colon cancer and lung cancer differ in their predilection for metastatic disease. Bone and brain metastasis are frequent in lung cancer, whereas they are unusual in colorectal cancer, with liver and lung being the most common sites for metastatic disease. DWI is increasingly part of WB-MRI protocols given its reported high sensitivity for metastasis, particularly in bone and liver.[13,61,63,86] However, in a large single-site study of WB-MRI staging in non-small-cell lung cancer, DWI sequences alone had sensitivity of 58% for metastatic disease compared with 70% when DWI was added to other sequences including post-contrast T1-weighted image acquisition.[22] This was predominantly caused by missed brain metastasis.

We prospectively tested the diagnostic accuracy of three WB-MRI sequence combinations across all participating radiologists as part of their real-time interpretation of the Streamline WB-MRI data sets. Overall, we found that the combination of T1-, T2-, diffusion-weighted sequences and post-contrast T1-weighted sequences was significantly more sensitive for metastatic colorectal cancer than a combination of non-contrast T1-weighted and diffusion-weighted sequences alone [72% (95% CI 60% to 81%) vs. 63% (95% CI 51% to 73%), respectively], whereas specificity was no different. However, there was no significant difference in sensitivity or specificity when using all available sequences compared with a combination of T1-, T2- and diffusion-weighted imaging without post-contrast-enhanced images [70% sensitivity (95% CI 58% to 80%)].

Findings were similar for staging non-small-cell lung cancer. Again, the combination of T1-, T2- and diffusion-weighted sequences and post-contrast T1-weighted sequences was significantly more sensitive for metastasis than a combination of non-contrast T1-weighted and diffusion-weighted sequences alone [52% (95% CI 39% to 65%) vs. 42% (95% CI 29% to 56%), respectively], whereas specificity was no different. The combination of T1-, T2- and diffusion-weighted imaging without post-contrast-enhanced images had 48% sensitivity (95% CI 35% to 61%).

The data suggest that, for both colorectal and non-small-cell lung cancer, time-efficient limited protocols based on non-contrast T1-weighted and diffusion-weighted sequences alone are insufficient and cannot be recommended. Addition of other anatomical (T2-weighted) sequences is required and thereafter there does appear to be a small incremental (although non-statistically significant) advantage in the use of intravenous gadolinium. Given the reported high sensitivity of DWI imaging in the bone and liver, we perhaps may have expected that the additional sequences would have improved specificity rather than sensitivity. This was not the case. However, there is evidence that non-contrast WB-MRI protocols have lower sensitivity for lung metastasis than standard pathways[87] and that detection of brain metastasis is also improved by addition of gadolinium-enhanced sequences.[22] Although the diagnostic advantage of adding gadolinium to T1-, T2- and diffusion-weighted sequences was marginal in the Streamline cohorts, and did not reach statistical significance, it should be remembered that the analysis was on a per-participant basis. Given the major clinical importance of detecting metastatic disease (e.g. a solitary brain metastasis in a participant with otherwise operable non-small-cell lung cancer), it seems prudent that WB-MRI protocols should include T1-weighted imaging sequences pre and post gadolinium enhancement and T2- and diffusion-weighted imaging sequences in both cancer types.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Interobserver agreement in WB-MRI interpretation

As noted in *Chapter 8*, the existing literature suggests that interobserver agreement for WB-MRI interpretation is high. However, most studies have been single-site studies and used highly experienced radiologists. Conversely, we tested interobserver agreement across 17 radiologists, representative of those who would report WB-MRI in daily clinical practice in the NHS. Our study is one of the largest in the literature. We tested agreement in two ways: (1) against the final reference standard diagnosis and (2) between reads regardless of the consensus reference standard adjudication of metastatic disease presence or absence. Such analyses are important when evaluating true interobserver agreement in clinical practice and its potential impact on patient management. Most studies 'reward' agreement between radiologists by their use of simple percentage agreement and kappa, even when both readers are wrong. Although this type of 'technical' agreement is of interest, it does not fully reflect the potential impact on patient care. For our primary analyses, such 'incorrect' agreement was not considered as true agreement. However, we also present simple agreement, which allows direct comparison with the existing literature. Another strength of our study design was the deliberate selection of data sets to mirror the prevalence of 'stress' cases as in the main trials. Thus, our data sets were representative of what was encountered (although we increased the prevalence of metastasis to increase statistical power). Not surprisingly, agreement was lower in the first analysis as it incorporated the technical failures of WB-MRI technology as well as interpretive differences. In both analyses, however, agreement was superior for colorectal cancer than for lung cancer. Simple agreement between the two reads for the primary outcome (i.e. per-patient detection of metastases) was 88% for colorectal cancer and 65% for non-small-cell lung cancer. This split between cancer types mirrors the data on the number of perceptual errors in the trial and reinforces the concept that metastases in non-small-cell lung cancer seem more subtle and less conspicuous than those in colorectal cancer (predominantly affecting the liver and lung). There is no specific evidence that our radiologists were more experienced in staging lung than colorectal cancer, but our data nonetheless suggest that dedicated training in WB-MRI over and above that achieved in usual clinical practice may be fruitful. This should perhaps focus on difficult disease organ sites such as bone. Further research into the impact of organ site on perceptual error and observer agreement is also indicated. It is notable that the way in which equivocal lesions were handled did affect agreement between radiologists, particularly for non-small-cell lung cancer. When equivocal lesions were classified as negative, simple agreement between the reads increased to 76%, suggesting that some of the disagreement was due to overdiagnosis of potential lesions. The clinical implications of the data are unclear, but a suggested topic for future research is the impact of dedicated radiologist training modules on the diagnostic accuracy of WB-MRI. The use of computer-assisted detection software is also of interest and currently under investigation.

Overall conclusions

- When tested in a prospective multicentre trial setting, the WB-MRI staging pathway had similar diagnostic accuracy to standard pathways for identifying patients with metastatic disease in newly diagnosed colorectal and non-small-cell lung cancer, and precipitated similar treatment decisions.
- WB-MRI pathways are inferior to standard pathways for defining N stage in non-small-cell lung cancer, although this does not adversely influence the first major treatment decision.
- The WB-MRI staging pathway was more efficient than standard pathways, reducing the time to staging completion and, in the case of colorectal cancer, overall number of tests.
- When using English NHS reference costs, the WB-MRI staging pathway cost less than standard staging pathways.
- Patients often find WB-MRI more challenging than standard staging investigations and devise coping strategies to complete the examination. Positive staff interaction is important. The quantified burden of WB-MRI, although relatively low, is greater than that of either CT or PET-CT. However, the majority of patients prefer the WB-MRI staging pathway if it is as accurate as standard pathways and reduces number of tests and/or staging time.

- Time-efficient WB-MRI protocols based on DWI and non-contrast enhanced T1-weighted sequences are insufficient and protocols including additional T2-weighted and post-gadolinium-enhanced T1-weighted sequences significantly improve sensitivity for patients with metastatic disease. Addition of post-gadolinium-enhanced T1 sequences has a small non-significant incremental advantage over protocols based on non-enhanced T1-, T2- and diffusion-weighted imaging sequences.
- There is generally moderate agreement between radiologists interpreting WB-MRI. However, interobserver agreement is inferior for non-small-cell lung cancer compared with colorectal cancer.

Implications for practice

In a NHS setting, the WB-MRI staging pathway is as accurate as current standard staging pathways for detecting patients with metastasis in both newly diagnosed colorectal and non-small-cell lung cancer. Sensitivity and overall agreement for N staging is lower in non-small-cell lung cancer but, for both cancers, the WB-MRI pathway results in the same treatment decisions as standard staging pathways. The WB-MRI pathway is more efficient, reducing time to complete staging and, in the case of colorectal cancer, the number of tests in the staging pathway. For both colorectal and non-small-cell lung cancer, the WB-MRI staging pathway is less costly than standard staging in an English NHS setting. Although patients find WB-MRI more challenging than standard staging investigations, in general they prefer the WB-MRI staging pathway over standard staging if it matches the diagnostic accuracy of standard pathways but reduces the time to staging and/or number of tests. Time-efficient WB-MRI protocols using just T1- and diffusion-weighted sequences have lower sensitivity than protocols that also include T2- and post-gadolinium-enhanced T1-weighted sequences. Radiologists must receive adequate training in interpreting WB-MRI, and interobserver variation is inferior for non-small-cell lung cancer compared with colorectal cancer when tested across NHS radiologists.

Recommendations for future research

Future research should investigate the following:

- The diagnostic accuracy, patient acceptability and cost-effectiveness of WB-MRI for staging and treatment follow-up of other primary cancer sites, notably breast, prostate and myeloma, compared with standard investigations. The NICE currently recommends WB-MRI in diagnosis, staging and treatment follow-up of myeloma, but there are few data concerning its comparative accuracy and cost-effectiveness with standard investigations such as CT. WB-MRI is used ad hoc in the NHS for staging and treatment follow-up in breast and prostate cancer and the evidence base needs developing.
- The diagnostic accuracy, patient acceptability and cost-effectiveness of WB-MRI for investigating patients with clinically suspected recurrence of colorectal and non-small-cell lung cancer compared with standard investigations. Early detection of cancer recurrence may increase the chance of successful treatment. Patients with suspected recurrence often undergo multiple testing and WB-MRI may be a more efficient and cost-effective method.
- The diagnostic accuracy, patient acceptability and cost-effectiveness of WB-MRI in routine post-cancer therapy surveillance in comparison with standard investigations. Surveillance of patients after cancer therapy using cross-sectional imaging is routine. WB-MRI may be a more accurate and cost-effective method.
- The impact on diagnostic accuracy and cost-effectiveness of adding liver-specific contrast agents to WB-MRI protocols. Liver metastases are common, particularly in colorectal cancer, and their presence usually drives treatment paradigms. There are good data suggesting that liver-specific contrast agents increase the sensitivity of liver MRI for metastasis detection over conventional gadolinium contrast, but the utility or otherwise and cost-effectiveness of adding such agents to WB-MRI protocols is unknown.
- The impact of formal WB-MRI training programmes on radiologist performance. Interobserver agreement in WB-MRI interpretation is at best moderate, and there were a number of perceptual errors. The impact of formal training programmes on radiologist performance is unknown.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Acknowledgements

The authors thank the NIHR University College London Hospital Biomedical Research Centre and NIHR Birmingham Biomedical Research Centre at the University Hospitals Birmingham NHS Foundation Trust and the University of Birmingham.

The authors are grateful to Biotronics3D, which hosted the WB-MRI images on its online server.

Streamline investigators

Barking, Havering and Redbridge University Hospitals NHS Trust

Simon Ball, Revanth Jannapureddy, Tina Mills-Baldock, Kishor Barhate, Zoltan Nagy, Sherif Raouf, Akosa Aboagye, Girija Anand, Rommel Butawan, Elizabeth Hadley, Adesewa Onajobi and Kathryn Tarver.

Barts Health NHS Trust

Steve Ellis, Ayshea Hameeduddin, Tanjil Nawaz, Catherine Norman, Nathalie Rich, Khawaja Shahabuddin, Sidra Tulmuntaha, Shafi Ahmed, Louise Lim, Fiona McKirdy, Jenna Couture, Shahanara Ferdous, Payal Julka, Ali Mohammed, Terry O'Shauhnessy, William Ricketts, Mohamed Thaha and Alistair Rienhardt.

Birkbeck, University of London

Ruth Evans.

Bradford Teaching Hospitals NHS Foundation Trust

Marie Jackson, Clive Kay, Andy Lowe, Janet McGowan, Amjad Mohammed and Jon Robinson.

Chelsea and Westminster Hospital NHS Foundation Trust

Lara Curry, Sasithar Maheswaran, Subramanian Ramesh and Pippa Riddle.

Croydon Health Services NHS Trust

Shaki Balogun, Yvonne Campbell, Nelesh Jeyadevan, Aji Kavidasan, Imogen Locke, Tuck-Kay Loke and Ibiyemi Olaleye.

East and North Hertfordshire NHS Trust

Clare Collins, Elizabeth Green, Colm Prendergast and Thida Win.

Epsom and St Helier University Hospitals NHS Trust

Amy Davis.

Guy's and St Thomas' NHS Foundation Trust

Lyn Blakeway, Sofia Gourtsoyianni, Adrian Green, Christian Kelly-Morland, Sahar Naaseri, Davide Prezzi, David Snell and Dorothee Boisfer.

Homerton University Hospital NHS Foundation Trust

Peter Boavida, Keyury Desai, Balinder Hans, Sophia Hans, Eleni Ntala, Adnam Alam, Stephen Burke, Helen Pardoe, Sanjaya Wijeyekoon and Angshu Bhowmik.

Imperial College Healthcare NHS Trust

Nishat Bharwani, Gule Hanid, Lesley Honeyfield, Tina Stoycheva, Katherine van Ree, Dominic Blunt and Nicola Strickland.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Kingston Hospital NHS Foundation Trust
Farid Bazari, Helen Beedham, Jane De Los, Reyes Lauigan, Priya Limbu, Nicola Lucas, Sally O'Connor and Anita Rhodes.

Lewisham and Greenwich NHS Trust
Laletha Agoramoorthy, Martha Handousa, Abel Jalloh, Stefania Stegner, Shanna Wilson, David Birch, Suzanne Chukundah, John O'Donohue, Priscilla Phiri, Raj Srirajaskanthan, Eleni Karapanagiotou and Daniel Smith.

Norfolk and Norwich University Hospital NHS Foundation Trust
Hameed Rafiee.

North Middlesex University Hospital NHS Trust
Ferrial Syeed, Chloe van Someren, Rudi Borgstein, Jamila Roehrig, David Chao and Lorraine Hurl.

Paul Strickland Scanner Centre, Mount Vernon Hospital
Andrew Gogbashian, Andre Nunes, Ian Simcock and James Stirling.

Portsmouth Hospitals NHS Trust
Richard Beable, Maureen Furneaux, Nicola Gibbons and Antony Higginson.

Princess Royal University Hospital
Howard Curtis and Kitrick Perry.

Royal Free London NHS Foundation Trust
Anita Amadi, Robert Glyne-Jones, Heather Hughes, Prital Patel, Matt Train, Gary Atkin, Colin Elton, Stephen Karp, Lisa Woodrow, Dominic Yu and Sajid Khan.

St Mark's Hospital and Academic Institute
Pooja Datt, Rajapandian Ilangovan, Ian Jenkins, Saba Mahmud and Uday Patel.

Princess Alexandra Hospital NHS Trust
Teresa Light, Joanne Kellaway, Ann O'Callaghan, William Partridge, Amelia Daniel, Ugo Ekeowa, Michael Long and Peter Russell.

Royal Marsden NHS Foundation Trust
Erica Scurr, Veronica Morgan and Nina Tunariu.

University College London Cancer Trials Centre
Elizabeth Chang, Marian Duggan, Laura Hughes, Ellice Marwood, Katie Prior, Krystyna Reczko, Meena Reddi, Kara Sargus, Abby Sharp and Jonathan Teague.

University College London Hospitals NHS Foundation Trust
Teresita Beeston, Gauraang Bhatnagar, Elizabeth Isaac, Adoracion Jayme, Edward Johnston, Jagadish Kalasthry, Wivijin Piga, Farzana Rahman, Harbir Sidhu, Shraddha Weir, Ashley Groves, Aileen Austria, James Crosbie, Alec Engledow, Jonathan McCullogh, Austen Obichere, Kai-Keen Shiu, Christopher Wanstall, Celia Simeon and Amy Smith.

University Hospital Southampton NHS Foundation Trust
Andrew Bateman, David Breen, Liane Davis, Chris Everitt, Alice Johnson, Paul Nichols and Beth Shepherd.

Whittington Health NHS Trust

Kayleigh Gilbert, Azmina Verjee, Michelle Saull, Jonathan Wilson, Rashidat Adeniba, Veronica Conteh, Sarah Howling and Sara Lock.

Oversight committees

The TSC and IDMC met at least annually and included the following members.

- TSC: Nick Reed (chairperson), Andrew Clamp (subject expert), Fergus Macbeth (subject expert), Richard Stephens (subject expert), Damian Tolan (subject expert) and Moira Heath (public representative).
- IDMC: Stuart Williams (chairperson), Richard Adams (subject expert), Caroline Kelly (statistician) and Peter Schmid (subject expert).

Contribution of authors

Stuart A Taylor (https://orcid.org/0000-0002-6765-8806) (Professor of Medical Imaging) was the chief investigator, conceived the study design, contributed to the protocol writing and study management as member of the TMG, interpreted trial imaging and performed the initial drafting and final editing of the report.

Susan Mallett (https://orcid.org/0000-0002-0596-8200) (Senior Statistician) helped to conceive the study design, contributed to the protocol writing and study management as member of the TMG, wrote the statistical analysis plan, led the statistical analysis and contributed to the writing and editing of the final report.

Anne Miles (https://orcid.org/0000-0003-3122-9900) (Senior Health Psychologist) helped to conceive the study design, contributed to the protocol writing and study management as member of the TMG, led the participant experience analysis plan and contributed to the writing and editing of the final report.

Stephen Morris (https://orcid.org/0000-0002-5828-3563) (Senior Health Economist) contributed to the study design and protocol writing, contributed to the health economic analysis plan, provided health economic analysis support and contributed to the writing and editing of the final report.

Laura Quinn (https://orcid.org/0000-0001-9660-4631) (Statistician) contributed to the statistical analysis plan, statistical analysis and the writing and editing of the final report.

Caroline S Clarke (https://orcid.org/0000-0002-4676-1257) (Health Economist) contributed to the health economic analysis plan, provided health economic analysis support and contributed to the writing and editing of the final report.

Sandy Beare (https://orcid.org/0000-0002-9055-1804) (Clinical Trial Centre Lead) contributed to the study design, study management as member of the TMG, data acquisition and the writing and editing of the final report.

John Bridgewater (https://orcid.org/0000-0001-9186-1604) (Consultant Oncologist) contributed to the study design and study management as member of the TMG, provided clinical advice and contributed to data acquisition and the writing and editing of the final report.

Vicky Goh (https://orcid.org/0000-0002-2321-8091) (Consultant Radiologist) contributed to the study design and study management as member of the TMG, interpreted trial imaging, collected trial data and contributed to the writing and editing of the final report.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Sam Janes (https://orcid.org/0000-0002-6634-5939) (Consultant Chest Physician) contributed to the study design and study management as member of the TMG, provided clinical advice and contributed to data acquisition and the writing and editing of the final report.

Dow-Mu Koh (https://orcid.org/0000-0001-7654-8011) (Consultant Radiologist) contributed to the study design and study management as member of the TMG, interpreted trial imaging, collected trial data and contributed to the writing and editing of the final report.

Alison Morton (https://orcid.org/0000-0002-2089-5257) (Patient Representative) contributed to the study design, study management as member of the TMG and the writing and editing of the final report.

Neal Navani (https://orcid.org/0000-0002-6412-7516) (Consultant Chest Physician) contributed to the study design and study management as member of the TMG, provided clinical advice and contributed to data acquisition and the writing and editing of the final report.

Alfred Oliver (https://orcid.org/0000-0002-7380-1074) (Patient Representative) contributed to the study design, study management as member of the TMG and the final report.

Anwar Padhani (https://orcid.org/0000-0002-5830-5777) (Consultant Radiologist) contributed to the study design and study management as member of the TMG, interpreted trial imaging, collected trial data and contributed to the writing and editing of the final report.

Shonit Punwani (https://orcid.org/0000-0002-1014-0870) (Consultant Radiologist) contributed to the study design and study management as member of the TMG, interpreted trial imaging, collected trial data and contributed to the writing and editing of the final report.

Andrea Rockall (https://orcid.org/0000-0001-8270-5597) (Consultant Radiologist) contributed to the study design and study management as member of the TMG, interpreted trial imaging, collected trial data and contributed to the writing and editing of the final report.

Steve Halligan (https://orcid.org/0000-0003-0632-5108) (Consultant Radiologist) contributed to the study design, protocol writing and study management as member of the TMG, interpreted trial imaging and contributed to the writing and editing of the final report.

Publications

Evans R, Taylor S, Janes S, Halligan S, Morton A, Navani N, *et al.* Patient experience and perceived acceptability of whole-body magnetic resonance imaging for staging colorectal and lung cancer compared with current staging scans: a qualitative study. *BMJ Open* 2017;**7**:e016391.

Taylor SA, Mallett S, Miles A, Beare S, Bhatnagar G, Bridgewater J, *et al.* Streamlining staging of lung and colorectal cancer with whole body MRI; study protocols for two multicentre, non-randomised, single-arm, prospective diagnostic accuracy studies (Streamline C and Streamline L). *BMC Cancer* 2017;**17**:299.

Evans RE, Taylor SA, Beare S, Halligan S, Morton A, Oliver A, *et al.* Perceived patient burden and acceptability of whole body MRI for staging lung and colorectal cancer; comparison with standard staging investigations. *Br J Radiol* 2018;**91**:20170731.

Miles A, Taylor SA, Evans REC, Halligan S, Beare S, Bridgewater J, *et al.* Patient preferences for whole-body MRI or conventional staging pathways in lung and colorectal cancer: a discrete choice experiment. *Eur Radiol* 2019;**29**:3889–900.

Taylor SA, Mallett S, Ball S, Beare S, Bhatnagar G, Bhowmik A, *et al.* Diagnostic accuracy of whole-body MRI versus standard imaging pathways for metastatic disease in newly diagnosed non-small-cell lung cancer: the prospective Streamline L trial. *Lancet Respir Med* 2019;**7**:523–32.

Taylor SA, Mallett S, Beare S, Bhatnagar G, Blunt D, Boavida P, *et al.* Diagnostic accuracy of whole-body MRI versus standard imaging pathways for metastatic disease in newly diagnosed colorectal cancer: the prospective Streamline C trial. *Lancet Gastroenterol Hepatol* 2019;**4**:529–37.

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 are 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.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

References

1. Cancer Research UK. *Statistics by Cancer Type*. URL: www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type (accessed 1 June 2018).
2. Buyse M, Burzykowski T, Carroll K, Michiels S, Sargent DJ, Miller LL, *et al.* Progression-free survival is a surrogate for survival in advanced colorectal cancer. *J Clin Oncol* 2007;**25**:5218–24. https://doi.org/10.1200/JCO.2007.11.8836
3. Mody K, Bekaii-Saab T. Clinical trials and progress in metastatic colon cancer. *Surg Oncol Clin N Am* 2018;**27**:349–65. https://doi.org/10.1016/j.soc.2017.11.008
4. Gilligan D, Nicolson M, Smith I, Groen H, Dalesio O, Goldstraw P, *et al.* Preoperative chemotherapy in patients with resectable non-small cell lung cancer: results of the MRC LU22/NVALT 2/EORTC 08012 multicentre randomised trial and update of systematic review. *Lancet* 2007;**369**:1929–37. https://doi.org/10.1016/S0140-6736(07)60714-4
5. National Institute for Health and Care Excellence (NICE). *Colorectal Cancer: Diagnosis and Management. Clinical Guideline (CG131)*. London: NICE; 2011. URL: www.nice.org.uk/guidance/cg131 (accessed 1 September 2018).
6. National Institute for Health and Care Excellence (NICE). *Lung Cancer: Diagnosis and Management. Clinical Guideline (CG121)*. London: NICE; 2011. URL: www.nice.org.uk/guidance/cg121 (accessed 1 September 2018).
7. Shah DJ, Sachs RK, Wilson DJ. Radiation-induced cancer: a modern view. *Br J Radiol* 2012;**85**:e1166–73. https://doi.org/10.1259/bjr/25026140
8. Brocken P, Prins JB, Dekhuijzen PN, van der Heijden HF. The faster the better? A systematic review on distress in the diagnostic phase of suspected cancer, and the influence of rapid diagnostic pathways. *Psycho-Oncology* 2012;**21**:1–10. https://doi.org/10.1002/pon.1929
9. Xu GZ, Li CY, Zhao L, He ZY. Comparison of FDG whole-body PET/CT and gadolinium-enhanced whole-body MRI for distant malignancies in patients with malignant tumors: a meta-analysis. *Ann Oncol* 2013;**24**:96–101. https://doi.org/10.1093/annonc/mds234
10. Duo J, Han X, Zhang L, Wang G, Ma Y, Yang Y. Comparison of FDG PET/CT and gadolinium-enhanced MRI for the detection of bone metastases in patients with cancer: a meta-analysis. *Clin Nucl Med* 2013;**38**:343–8. https://doi.org/10.1097/RLU.0b013e3182817af3
11. Li B, Li Q, Nie W, Liu S. Diagnostic value of whole-body diffusion-weighted magnetic resonance imaging for detection of primary and metastatic malignancies: a meta-analysis. *Eur J Radiol* 2014;**83**:338–44. https://doi.org/10.1016/j.ejrad.2013.11.017
12. Liu T, Xu JY, Xu W, Bai YR, Yan WL, Yang HL. Fluorine-18 deoxyglucose positron emission tomography, magnetic resonance imaging and bone scintigraphy for the diagnosis of bone metastases in patients with lung cancer: which one is the best? – a meta-analysis. *Clin Oncol* 2011;**23**:350–8. https://doi.org/10.1016/j.clon.2010.10.002
13. Wu LM, Gu HY, Zheng J, Xu X, Lin LH, Deng X, *et al.* Diagnostic value of whole-body magnetic resonance imaging for bone metastases: a systematic review and meta-analysis. *J Magn Reson Imaging* 2011;**34**:128–35. https://doi.org/10.1002/jmri.22608
14. Yang HL, Liu T, Wang XM, Xu Y, Deng SM. Diagnosis of bone metastases: a meta-analysis comparing ^{18}FDG PET, CT, MRI and bone scintigraphy. *Eur Radiol* 2011;**21**:2604–17. https://doi.org/10.1007/s00330-011-2221-4

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

15. Qu X, Huang X, Yan W, Wu L, Dai K. A meta-analysis of ^{18}FDG-PET-CT, ^{18}FDG-PET, MRI and bone scintigraphy for diagnosis of bone metastases in patients with lung cancer. *Eur J Radiol* 2012;**81**:1007–15. https://doi.org/10.1016/j.ejrad.2011.01.126

16. Wu Q, Yang R, Zhou F, Hu Y. Comparison of whole-body MRI and skeletal scintigraphy for detection of bone metastatic tumors: a meta-analysis. *Surg Oncol* 2013;**22**:261–6. https://doi.org/10.1016/j.suronc.2013.10.004

17. Smets AM, Deurloo EE, Slager TJE, Stoker J, Bipat S. Whole-body magnetic resonance imaging for detection of skeletal metastases in children and young people with primary solid tumors – systematic review. *Pediatr Radiol* 2018;**48**:241–52. https://doi.org/10.1007/s00247-017-4013-8

18. Peerlings J, Troost EG, Nelemans PJ, Cobben DC, de Boer JC, Hoffmann AL, Beets-Tan RG. The diagnostic value of MR imaging in determining the lymph node status of patients with non-small cell lung cancer: a meta-analysis. *Radiology* 2016;**281**:86–98. https://doi.org/10.1148/radiol.2016151631

19. Shen G, Hu S, Deng H, Kuang A. Performance of DWI in the nodal characterization and assessment of lung cancer: a meta-analysis. *AJR Am J Roentgenol* 2016;**206**:283–90. https://doi.org/10.2214/AJR.15.15032

20. Wu LM, Xu JR, Gu HY, Hua J, Chen J, Zhang W, *et al.* Preoperative mediastinal and hilar nodal staging with diffusion-weighted magnetic resonance imaging and fluorodeoxyglucose positron emission tomography/computed tomography in patients with non-small cell lung cancer: which is better? *J Surg Res* 2012;**178**:304–14. https://doi.org/10.1016/j.jss.2012.03.074

21. Squillaci E, Manenti G, Mancino S, Cicciò C, Calabria F, Danieli R, *et al.* Staging of colon cancer: whole-body MRI vs. whole-body PET-CT – initial clinical experience. *Abdom Imaging* 2008;**33**:676–88. https://doi.org/10.1007/s00261-007-9347-5

22. Ohno Y, Koyama H, Onishi Y, Takenaka D, Nogami M, Yoshikawa T, *et al.* non-small cell lung cancer: whole-body MR examination for M-stage assessment – utility for whole-body diffusion-weighted imaging compared with integrated FDG PET/CT. *Radiology* 2008;**248**:643–54. https://doi.org/10.1148/radiol.2482072039

23. Taylor SA, Mallett S, Miles A, Beare S, Bhatnagar G, Bridgewater J, *et al.* Streamlining staging of lung and colorectal cancer with whole body MRI; study protocols for two multicentre, non-randomised, single-arm, prospective diagnostic accuracy studies (Streamline C and Streamline L). *BMC Cancer* 2017;**17**:299. https://doi.org/10.1186/s12885-017-3281-x

24. Ferrante di Ruffano L, Hyde CJ, McCaffery KJ, Bossuyt PM, Deeks JJ. Assessing the value of diagnostic tests: a framework for designing and evaluating trials. *BMJ* 2012;**344**:e686. https://doi.org/10.1136/bmj.e686

25. Taylor SA, Mallett S, Beare S, Bhatnagar G, Blunt D, Boavida P, *et al.* Diagnostic accuracy of whole-body MRI versus standard imaging pathways for metastatic disease in newly diagnosed colorectal cancer: the prospective Streamline C trial. *Lancet Gastroenterol Hepatol* 2019;**4**:529–37. https://doi.org/10.1016/S2468-1253(19)30056-1

26. Taylor SA, Mallett S, Ball S, Beare S, Bhatnagar G, Bhowmik A, *et al.* Diagnostic accuracy of whole-body MRI versus standard imaging pathways for metastatic disease in newly diagnosed non-small cell lung cancer: the prospective Streamline L trial. *Lancet Respir Med* 2019;**7**:523–32. https://doi.org/10.1016/S2213-2600(19)30090-6

27. Edge SB, Byrd DR, Compton CC, Fritz AG, Greene FL, Trotti A, editors. *AJCC Cancer Staging Manual*. 7th edn. New York, NY: Springer; 2010.

28. Rutjes AW, Reitsma JB, Coomarasamy A, Khan KS, Bossuyt PM. Evaluation of diagnostic tests when there is no gold standard. A review of methods. *Health Technol Assess* 2007;**11**(50). https://doi.org/10.3310/hta11500

29. Alonzo TA, Pepe MS, Moskowitz CS. Sample size calculations for comparative studies of medical tests for detecting presence of disease. *Stat Med* 2002;**21**:835–52. https://doi.org/10.1002/sim.1058

30. 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

31. Huppertz A, Schmidt M, Wagner M, Puettcher O, Asbach P, Strassburg J, *et al.* Whole-body MR imaging versus sequential multimodal diagnostic algorithm for staging patients with rectal cancer: cost analysis. *Rofo* 2010;**182**:793–802. https://doi.org/10.1055/s-0029-1245463

32. Department of Health and Social Care (DHSC). *NHS Reference Costs 2016/17*. London: DHSC; 2017. URL: https://improvement.nhs.uk/resources/reference-costs/ (accessed 1 September 2018).

33. Evans R, Taylor S, Janes S, Halligan S, Morton A, Navani N, *et al.* Patient experience and perceived acceptability of whole-body magnetic resonance imaging for staging colorectal and lung cancer compared with current staging scans: a qualitative study. *BMJ Open* 2017;**7**:e016391. https://doi.org/10.1136/bmjopen-2017-016391

34. Evans RE, Taylor SA, Beare S, Halligan S, Morton A, Oliver A, *et al.* Perceived patient burden and acceptability of whole body MRI for staging lung and colorectal cancer; comparison with standard staging investigations. *Br J Radiol* 2018;**91**:20170731. https://doi.org/10.1259/bjr.20170731

35. Miles A, Taylor SA, Evans REC, Halligan S, Beare S, Bridgewater J, *et al.* Patient preferences for whole-body MRI or conventional staging pathways in lung and colorectal cancer: a discrete choice experiment. *Eur Radiol* 2019;**29**:3889–900. https://doi.org/10.1007/s00330-019-06153-4

36. van Minde D, Klaming L, Weda H. Pinpointing moments of high anxiety during an MRI examination. *Int J Behav Med* 2014;**21**:487–95. https://doi.org/10.1007/s12529-013-9339-5

37. 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

38. McIsaac HK, Thordarson DS, Shafran R, Rachman S, Poole G. Claustrophobia and the magnetic resonance imaging procedure. *J Behav Med* 1998;**21**:255–68. https://doi.org/10.1023/A:1018717016680

39. Dewey M, Schink T, Dewey CF. Claustrophobia during magnetic resonance imaging: cohort study in over 55,000 patients. *J Magn Reson Imaging* 2007;**26**:1322–7. https://doi.org/10.1002/jmri.21147

40. Dantendorfer K, Amering M, Bankier A, Helbich T, Prayer D, Youssefzadeh S, *et al.* A study of the effects of patient anxiety, perceptions and equipment on motion artifacts in magnetic resonance imaging. *Magn Reson Imaging* 1997;**15**:301–6. https://doi.org/10.1016/S0730-725X(96)00385-2

41. Harris LM, Cumming SR, Menzies RG. Predicting anxiety in magnetic resonance imaging scans. *Int J Behav Med* 2004;**11**:1–7. https://doi.org/10.1207/s15327558ijbm1101_1

42. Shortman RI, Neriman D, Hoath J, Millner L, Endozo R, Azzopardi G, *et al.* A comparison of the psychological burden of PET/MRI and PET/CT scans and association to initial state anxiety and previous imaging experiences. *Br J Radiol* 2015;**88**:20150121. https://doi.org/10.1259/bjr.20150121

43. Carlsson S, Carlsson E. 'The situation and the uncertainty about the coming result scared me but interaction with the radiographers helped me through': a qualitative study on patients' experiences of magnetic resonance imaging examinations. *J Clin Nurs* 2013;**22**:3225–34. https://doi.org/10.1111/jocn.12416

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

44. Törnqvist E, Månsson A, Larsson EM, Hallström I. It's like being in another world – patients' lived experience of magnetic resonance imaging. *J Clin Nurs* 2006;**15**:954–61. https://doi.org/10.1111/j.1365-2702.2006.01499.x

45. Wiljer D, Walton T, Gilbert J, Boucher A, Ellis PM, Schiff S, *et al.* Understanding the needs of colorectal cancer patients during the pre-diagnosis phase. *J Cancer Educ* 2013;**28**:402–7. https://doi.org/10.1007/s13187-013-0465-1

46. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nurs Health Sci* 2013;**15**:398–405. https://doi.org/10.1111/nhs.12048

47. Goldberg DP, Gater R, Sartorius N, Ustun TB, Piccinelli M, Gureje O, Rutter C. The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychol Med* 1997;**27**:191–7. https://doi.org/10.1017/S0033291796004242

48. Salmon P, Shah R, Berg S, Williams C. Evaluating customer satisfaction with colonoscopy. *Endoscopy* 1994;**26**:342–6. https://doi.org/10.1055/s-2007-1008988

49. von Wagner C, Smith S, Halligan S, Ghanouni A, Power E, Lilford RJ, *et al.* Patient acceptability of CT colonography compared with double contrast barium enema: results from a multicentre randomised controlled trial of symptomatic patients. *Eur Radiol* 2011;**21**:2046–55. https://doi.org/10.1007/s00330-011-2154-y

50. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care* 2003;**41**:582–92. https://doi.org/10.1097/01.MLR.0000062554.74615.4C

51. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR, Clinical Significance Consensus Meeting Group. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc* 2002;**77**:371–83. https://doi.org/10.4065/77.4.371

52. Knott C. General Mental and Physical Health. In Craig R, Mindell J, editors. *Health Survey for England 2012: Health, Social Care and Lifestyle*. Leeds: NHS Digital; 2013.

53. Olomu AB, Corser WD, Stommel M, Xie Y, Holmes-Rovner M. Do self-report and medical record comorbidity data predict longitudinal functional capacity and quality of life health outcomes similarly? *BMC Health Serv Res* 2012;**12**:398. https://doi.org/10.1186/1472-6963-12-398

54. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. *PharmacoEconomics* 2008;**26**:661–77. https://doi.org/10.2165/00019053-200826080-00004

55. Bridges JF, Hauber AB, Marshall D, Lloyd A, Prosser LA, Regier DA, *et al.* Conjoint analysis applications in health – a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. *Value Health* 2011;**14**:403–13. https://doi.org/10.1016/j.jval.2010.11.013

56. Hall EJ, Brenner DJ. Cancer risks from diagnostic radiology. *Br J Radiol* 2008;**81**:362–78. https://doi.org/10.1259/bjr/01948454

57. Hahn G, Shapiro S. *A Catalogue and Computer Program for the Design and Analysis of Orthogonal Symmetric and Asymmetric Fractional Factorial Experiments*. New York, NY: General Electric Research and Development Centre; 1966.

58. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. *J Pers Soc Psychol* 1988;**54**:1063–70. https://doi.org/10.1037/0022-3514.54.6.1063

59. de Bekker-Grob EW, Donkers B, Jonker MF, Stolk EA. Sample size requirements for discrete-choice experiments in healthcare: a practical guide. *Patient* 2015;**8**:373–84. https://doi.org/10.1007/s40271-015-0118-z

60. Hauber AB, González JM, Groothuis-Oudshoorn CG, Prior T, Marshall DA, Cunningham C, *et al.* Statistical methods for the analysis of discrete choice experiments: a report of the ISPOR Conjoint Analysis Good Research Practices Task Force. *Value Health* 2016;**19**:300–15. https://doi.org/10.1016/j.jval.2016.04.004

61. Morone M, Bali MA, Tunariu N, Messiou C, Blackledge M, Grazioli L, Koh DM. Whole-Body MRI: current applications in oncology. *AJR Am J Roentgenol* 2017;**209**:W336–W349. https://doi.org/10.2214/AJR.17.17984

62. Padhani AR, Lecouvet FE, Tunariu N, Koh DM, De Keyzer F, Collins DJ, *et al.* Metastasis reporting and data system for prostate cancer: practical guidelines for acquisition, interpretation, and reporting of whole-body magnetic resonance imaging-based evaluations of multiorgan involvement in advanced prostate cancer. *Eur Urol* 2017;**71**:81–92. https://doi.org/10.1016/j.eururo.2016.05.033

63. Petralia G, Padhani A, Summers P, Alessi S, Raimondi S, Testori A, *et al.* Whole-body diffusion-weighted imaging: is it all we need for detecting metastases in melanoma patients? *Eur Radiol* 2013;**23**:3466–76. https://doi.org/10.1007/s00330-013-2968-x

64. Grueneisen J, Beiderwellen K, Heusch P, Gratz M, Schulze-Hagen A, Heubner M, *et al.* Simultaneous positron emission tomography/magnetic resonance imaging for whole-body staging in patients with recurrent gynecological malignancies of the pelvis: a comparison to whole-body magnetic resonance imaging alone. *Invest Radiol* 2014;**49**:808–15. https://doi.org/10.1097/RLI.0000000000000086

65. Paruthikunnan SM, Kadavigere R, Karegowda LH. Accuracy of whole-body DWI for metastases screening in a diverse group of malignancies: comparison with conventional cross-sectional imaging and nuclear scintigraphy. *AJR Am J Roentgenol* 2017;**209**:477–90. https://doi.org/10.2214/AJR.17.17829

66. Blackledge MD, Tunariu N, Orton MR, Padhani AR, Collins DJ, Leach MO, Koh DM. Inter- and intra-observer repeatability of quantitative whole-body, diffusion-weighted imaging (WBDWI) in metastatic bone disease. *PLOS ONE* 2016;**11**:e0153840. https://doi.org/10.1371/journal.pone.0153840

67. Ohno Y, Koyama H, Yoshikawa T, Takenaka D, Seki S, Yui M, *et al.* Three-way comparison of whole-body MR, coregistered whole-body FDG PET/MR, and integrated whole-body FDG PET/CT imaging: TNM and stage assessment capability for non-small cell lung cancer patients. *Radiology* 2015;**275**:849–61. https://doi.org/10.1148/radiol.14140936

68. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol* 1993;**46**:423–9. https://doi.org/10.1016/0895-4356(93)90018-V

69. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;**33**:159–74. https://doi.org/10.2307/2529310

70. Borowski DW, Cawkwell S, Zaidi SMA, Toward M, Maguire N, Garg DK, Gill TS. The NHS bowel cancer screening programme achieves the anticipated survival improvement, but participation must be improved. *Int J Health Care Qual Assur* 2018;**31**:106–15. https://doi.org/10.1108/IJHCQA-11-2016-0169

71. Fischer B, Lassen U, Mortensen J, Larsen S, Loft A, Bertelsen A, *et al.* Preoperative staging of lung cancer with combined PET-CT. *N Engl J Med* 2009;**361**:32–9. https://doi.org/10.1056/NEJMoa0900043

72. Niekel MC, Bipat S, Stoker J. Diagnostic imaging of colorectal liver metastases with CT, MR imaging, FDG PET, and/or FDG PET/CT: a meta-analysis of prospective studies including patients who have not previously undergone treatment. *Radiology* 2010;**257**:674–84. https://doi.org/10.1148/radiol.10100729

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

73. Choi SH, Kim SY, Park SH, Kim KW, Lee JY, Lee SS, Lee MG. Diagnostic performance of CT, gadoxetate disodium-enhanced MRI, and PET/CT for the diagnosis of colorectal liver metastasis: systematic review and meta-analysis. *J Magn Reson Imaging* 2018;**47**:1237–50. https://doi.org/10.1002/jmri.25852

74. Navani N, Fisher D, Tierney JF, Stephens RJ, Burdett S. The accuracy of clinical staging of stage I-IIIa non-small cell lung cancer: an analysis based on individual participant data. *Chest* 2018;**3**:502–9. https://doi.org/10.1016/j.chest.2018.10.020

75. Hunter C, Blake H, Jeyadevan N, Abulafi M, Swift I, Toomey P, Brown G. Local staging and assessment of colon cancer with 1.5-T magnetic resonance imaging. *Br J Radiol* 2016;**89**:20160257. https://doi.org/10.1259/bjr.20160257

76. Kim HY, Yi CA, Lee KS, Chung MJ, Kim YK, Choi BK, *et al.* Nodal metastasis in non-small cell lung cancer: accuracy of 3.0-T MR imaging. *Radiology* 2008;**246**:596–604. https://doi.org/10.1148/radiol.2461061907

77. Navani N, Nankivell M, Lawrence DR, Lock S, Makker H, Baldwin DR, *et al.* Lung cancer diagnosis and staging with endobronchial ultrasound-guided transbronchial needle aspiration compared with conventional approaches: an open-label, pragmatic, randomised controlled trial. *Lancet Respir Med* 2015;**3**:282–9. https://doi.org/10.1016/S2213-2600(15)00029-6

78. Brocken P, van der Heijden EH, Oud KT, Bootsma G, Groen HJ, Donders AR, *et al.* Distress in suspected lung cancer patients following rapid and standard diagnostic programs: a prospective observational study. *Psycho-Oncology* 2015;**24**:433–41. https://doi.org/10.1002/pon.3660

79. Yu AM, Balasubramanaiam B, Offringa M, Kelly LE. Reporting of interventions and 'standard of care' control arms in pediatric clinical trials: a quantitative analysis. *Pediatr Res* 2018;**84**:393–8. https://doi.org/10.1038/s41390-018-0019-7

80. Lang EV, Ward C, Laser E. Effect of team training on patients' ability to complete MRI examinations. *Acad Radiol* 2010;**17**:18–23. https://doi.org/10.1016/j.acra.2009.07.002

81. Adams HJ, Kwee TC, Vermoolen MA, Ludwig I, Bierings MB, Nievelstein RA. Whole-body MRI vs. CT for staging lymphoma: patient experience. *Eur J Radiol* 2014;**83**:163–6. https://doi.org/10.1016/j.ejrad.2013.10.008

82. Sogaard M, Thomsen RW, Bossen KS, Sorensen HT, Norgaard M. The impact of comorbidity on cancer survival: a review. *Clin Epidemiol* 2013;**5**:3–29. https://doi.org/10.2147/CLEP.S47150

83. Faggiano F, Partanen T, Kogevinas M, Boffetta P. Socioeconomic differences in cancer incidence and mortality. *IARC Sci Publ* 1997;**138**:65–176.

84. Dutoit JC, Vanderkerken MA, Verstraete KL. Value of whole body MRI and dynamic contrast enhanced MRI in the diagnosis, follow-up and evaluation of disease activity and extent in multiple myeloma. *Eur J Radiol* 2013;**82**:1444–52. https://doi.org/10.1016/j.ejrad.2013.04.012

85. Olchowy C, Cebulski K, Łasecki M, Chaber R, Olchowy A, Kałwak K, Zaleska-Dorobisz U. The presence of the gadolinium-based contrast agent depositions in the brain and symptoms of gadolinium neurotoxicity – a systematic review. *PLOS ONE* 2017;**12**:e0171704. https://doi.org/10.1371/journal.pone.0171704

86. Wei C, Tan J, Xu L, Juan L, Zhang SW, Wang L, Wang Q. Differential diagnosis between hepatic metastases and benign focal lesions using DWI with parallel acquisition technique: a meta-analysis. *Tumour Biol* 2015;**36**:983–90. https://doi.org/10.1007/s13277-014-2663-9

87. Siegel MJ, Acharyya S, Hoffer FA, Wyly JB, Friedmann AM, Snyder BS, *et al.* Whole-body MR imaging for staging of malignant tumors in pediatric patients: results of the American College of Radiology Imaging Network 6660 Trial. *Radiology* 2013;**266**:599–609. https://doi.org/10.1148/radiol.12112531

Appendix 1 WB-MRI acquisition: minimum data set

MRI protocol

The aim is to complete the whole protocol in ≤ 60 minutes. Depending on the MRI technology available, it may be possible to use the 5-mm slices for all axial imaging, although up to 7 mm is acceptable if time constraints are problematic.

These specifications are a set of minimum requirements; where higher field strengths/scanner software permits, the resolution/quality of sequences should be optimised (while keeping the maximum imaging time limited to 1 hour).

Scanning may be performed at either 1.5 T or 3 T.

Whole-body coverage is defined as head to mid-thigh, optimised for detection of metastases.

Imaging procedure

1. A standard safety questionnaire should be completed.
2. For patients undergoing contrast enhancement, set up intravenous line in a vein in the antecubital fossa, connected to an automated injector with two syringes (contrast and saline flush).
3. Unless contraindicated, administer 20 mg of buscopan or 1 mg of glucagon, to be given just before the start of the scan.

Whole-body diffusion-weighted imaging

- Axial: using short T1 inversion recovery-echo planar imaging (or other fat saturation technique) DWI. Fixed slice thickness of 5–7 mm (to match T2- and T1-weighted axials as below), two *b*-values (b50 and b900). A minimum acquisition matrix of 128 × 128 (or an interpolated equivalent) [rectangular field of view (FOV) should be used if available and appropriate for the patient]; as a reference, a minimum signal-to-noise ratio of 6 on b50 images (for liver) should be maintained if possible by increasing the number of averages. All imaging should be performed in gentle respiration (recommended as four stations of 50 slices beginning from the vertex to mid-thighs). Diffusion imaging through the brain is optional.

Whole-body T2-weighted imaging

- Axial: axial T2-weighted (without fat suppression) imaging should be performed from vertex to mid-thigh. A 5- to 7-mm slice thickness should be used for all scanners. Where possible (within the 60-minute imaging time), respiratory and electrocardiogram triggering should be used for the chest and respiratory triggering alone for the upper abdomen. The head, neck, pelvis and legs should be scanned without any triggering. The number of stacks should be adjusted to cover the imaging volume.

Pre-contrast T1-weighted imaging

Dixon technique to be applied if available.

- Axial: whole-body T1-weighted gradient echo (GRE) (e.g. fast low-angle shot two-dimensional) non-contrast enhanced non-fat-saturation. Image resolution and slice thickness should be ideally matched to T2-weighted imaging.
- Coronal: T1-weighted fat-saturated volume interpolated GRE imaging [e.g. three dimensional (3D)] pre contrast.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Post-contrast T1-weighted imaging (if gadolinium not contraindicated or refused)

Minimum data set.
Axial liver (60–70 seconds).
Axial lung (equilibrium phase).
Small field of view (SFOV) axial head.
Optional: coronal (organ specific or whole body).

- Axial: post contrast; for example T1-weighted fat-saturated volume interpolated GRE imaging (3D), breath hold of the liver (60- to 70-second delay) and lungs. Multiple breath holds employed to provide full volume coverage if required. A minimum of a 256 × 256 (rectangular FOV acquisition if possible and appropriate for the patient) acquisition matrix should be employed. Slice thickness 5–7 mm.
- Coronal: post contrast, whole body, for example T1-weighted fat-saturated volume interpolated GRE imaging (3D), and post contrast. Slice thickness 5 mm. Breath hold.
- Axial: fat-saturated T1-weighted imaging of the brain (SFOV). An acquisition matrix of 256 × 256 should be employed.

Appendix 2 Quality assurance grading score definitions

TABLE 41 Quality assurance grading score definitions

Technical quality	Definition
General	
1	More than one sequence with substantial degradation of images severely limiting interpretation of those sequences; not repeated
2	One sequence with substantial degradation of images severely limiting interpretation of that sequence; not repeated
3	More than one sequence has minor artefacts, but all remain fully diagnostic; repeat, although optimal, not necessary, or all sequences initially technically inadequate (score 1 or 2) correctly repeated
4	One sequence has a minor artefact but remains fully diagnostic; repeat, although optimal, not necessary
5	All sequences technically optimal with no artefacts or degradation
Anatomical coverage	
1	Wrong examination performed
2	More than one sequence does not adequately cover the body (skull to mid-thigh) or designated organ(s)
3	One sequence does not optimally cover the body or designated organ(s), but examination remains fully diagnostic
4	All sequences optimally cover the body and designated organ(s)

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 3 Flow chart of multidisciplinary team process recording the first major treatment decision based on standard and WB-MRI pathways

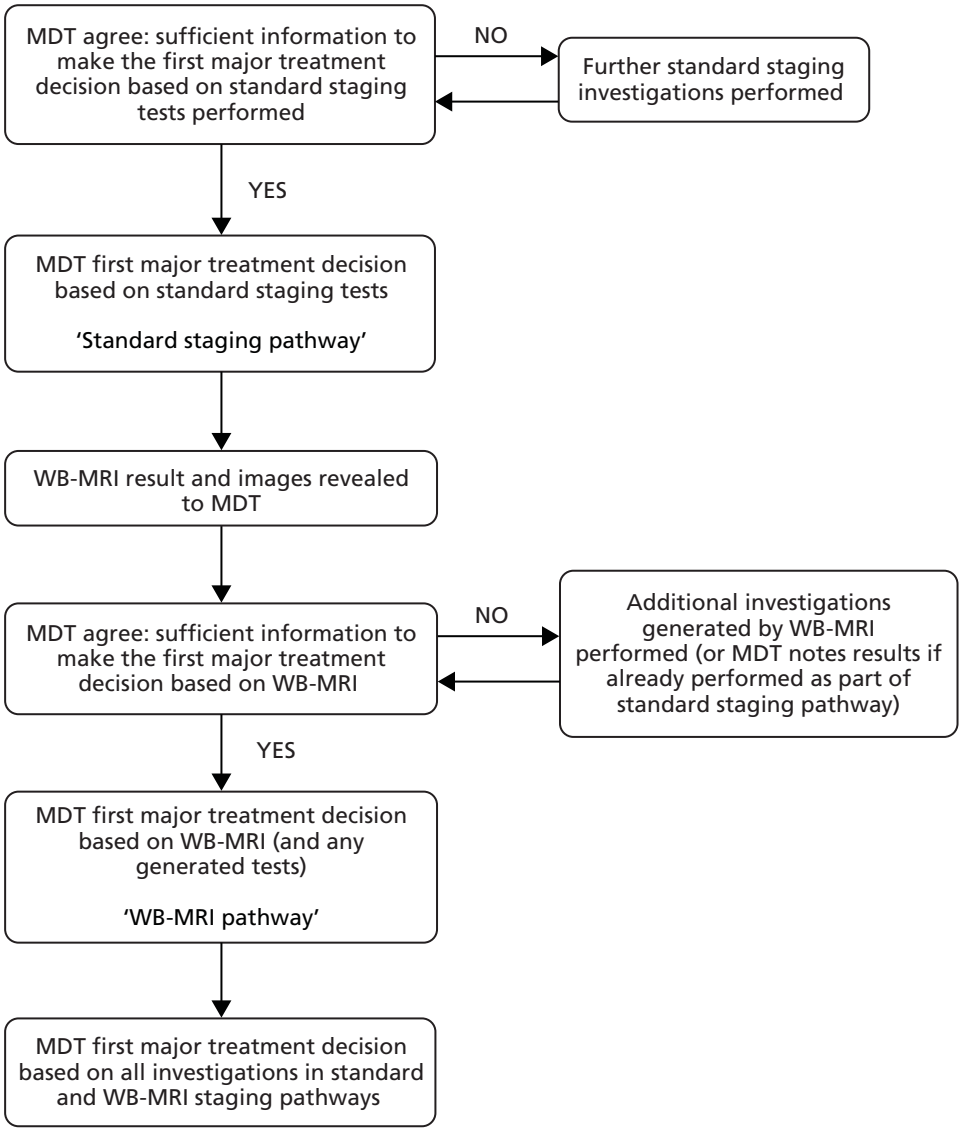


FIGURE 11 Flow chart of MDT process recording the first major treatment decision based on standard and WB-MRI pathways.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 4 Summary of primary and secondary outcomes

TABLE 42 Summary of primary and secondary outcomes

Aim/outcome	Description
Overall aim	To evaluate whether or not early WB-MRI increases detection rate for metastasis in colorectal cancer or non-small-cell lung cancer compared with standard NICE-approved diagnostic pathways
Primary outcome	• To evaluate whether or not early WB-MRI used as a replacement test increases detection rate for metastasis in colorectal cancer or non-small-cell lung cancer compared with standard diagnostic pathways • Comparison to detect metastases between WB-MRI test strategy A[a] and standard diagnostic strategy test strategy B[b] • Difference in per-patient sensitivity • Paired comparison of proportions of each strategy against the reference standard of full clinical diagnosis and 12-month follow-up • Patients without consensus reference standard are not included in analysis • Imputation will be used to account for missing data and imperfect reference data
Secondary outcome	

1	• To determine how WB-MRI influences time to and nature of first major treatment decision following definitive staging compared with standard pathways and to determine whether or not early WB-MRI could reduce or replace standard investigations
	• In this secondary outcome, the following components will be measured up to the time of the first major treatment decision:
	  ○ nature of first major treatment decision and number of patients with changed management decision
	  ○ time taken for diagnostic pathways
	  ○ number of tests – average per patient; number of patients with fewer tests in pathway
	• These will be compared for WB-MRI pathway[a] in comparison to standard staging pathways[b]
	• This information will be reported and discussed as a basis for whether or not early WB-MRI (test strategy A[a]) could replace or reduce standard investigations (test strategy B[b])
	• No absolute pre-specified definition of the combination of these outcomes likely to be considered clinically significant is attempted, although it is noted that an average decrease in diagnosis time of 7 days is likely to be considered clinically significant if accuracy were similar

2	• Diagnostic accuracy of WB-MRI and standard staging pathways for local tumour staging and detection of metastasis in comparison to an expert-derived consensus reference standard
	• This secondary outcome will include per-patient analysis of:
	  ○ difference in specificity for test strategy A[a] minus test strategy B[b] (note primary outcome is difference in sensitivity)
	  ○ difference in sensitivity for test strategy C[c] minus test strategy B[b]
	  ○ difference in specificity for test strategy C[c] minus test strategy B[b]
	• This secondary outcome will include per-organ analysis of:
	  ○ difference in sensitivity for test strategy A[a] minus test strategy B[b]
	  ○ difference in specificity for test strategy A[a] minus test strategy B[b]
	  ○ difference in sensitivity for test strategy C[c] minus test strategy B[b]
	  ○ difference in specificity for test strategy C[c] minus test strategy B[b]

3	Lifetime incremental cost and cost-effectiveness of staging using WB-MRI compared with standard diagnostic pathways
4	Patient experience of staging using WB-MRI in comparison to standard diagnostic pathways and priorities placed by patients on differing attributes related to competing staging pathways

continued

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 42 Summary of primary and secondary outcomes (*continued*)

Aim/outcome	Description
5	• Interobserver variability in WB-MRI analysis and the effect of diagnostic confidence on staging accuracy • A substudy conducted to look at interobserver variability between radiologists interpreting WB-MRI images and the effect of diagnostic confidence on staging accuracy • Each hub radiologist ($n \geq 5$) interpreting a sample of WB-MRI data sets selected at random from the other imaging hubs to define interobserver variation in the reported presence/absence of metastatic disease
6	• Diagnostic accuracy of limited T1- and diffusion-weighted sequences compared with full multisequence WB-MRI protocols • Comparison of diagnostic accuracy of WB-MRI alone: (i) sequence block 1 (ii) sequence block 2 and (iii) sequence block 3 • Reference standard: WB-MRI plus standard tests with 12-month follow-up • Methods and tables: as per secondary outcome 2

Additional analyses	• Adverse events from WB-MRI
	• Descriptive tables of WB-MRI reported, including quality of imaging based on technical quality scored as (i) general properties and (ii) anatomical coverage
	• Streamline C only: subgroup analysis for main trial outcomes – rectal cancer and non-rectal cancer
	• Although this trial includes all colorectal cancers, descriptive analyses will be completed separately for:
	  ○ rectal cancer
	  ○ non-rectal cancer
	• Outcomes to be analysed include primary outcome – the difference in per-patient sensitivity to detect metastases
	• Additional outcomes: secondary outcome 3 (per-patient specificity, per-organ sensitivity and specificity)

a WB-MRI pathway including additional tests requested as a result of WB-MRI findings.
b Standard pathway.
c WB-MRI pathway (plus additional tests generated), plus standard CT pathway tests (add-on test to test strategy B).

Appendix 5 Streamline C: reasons for non-inclusion of screened patients

TABLE 43 Streamline C: reasons for non-inclusion of screened patients

Reason for non-inclusion	*n* (%)
Patients did not want extra visit	91 (14)
Did not want WB-MRI scan	77 (12)
Patients felt would delay potential treatment	70 (11)
Contraindications to WB-MRI	66 (10)
Polyp cancer	65 (10)
No WB-MRI slot available within 3 weeks of final staging	37 (6)
Evidence of severe or uncontrolled systematic disease	31 (5)
Not primary colorectal cancer	31 (5)
Too frail/not fit enough	24 (3)
Psychiatric or other condition likely to impact on informed consent	22 (3)
Language barrier	14 (2)
Other	122 (19)
Total	650

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 6 Streamline C: reasons for withdrawal of recruited participants

TABLE 44 Streamline C: reasons for withdrawal of recruited participants

Reason for withdrawal	*n* (%)
Participant withdrawal owing to failure to undergo WB-MRI	45 (63)
Participant withdrawal owing to non primary colorectal cancer diagnosis	18 (25)
Participant withdrew consent from trial	3 (4)
Participant lost to follow-up	3 (4)
Participant withdrew by consensus; not enough imaging	1 (2)
Participant had no consensus owing to administrative error	1 (2)
Total	71

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 7 Streamline C: primary tumour status at the time of consensus review

TABLE 45 Streamline C: primary tumour status at the time of consensus review

Participant status	*n* (%)
Participant died within 12 months[a]	17 (6)
Primary tumour removed within 3 months of diagnosis	219 (73)
Primary tumour in situ for > 3 months (or incomplete removal) – imaging capable of detecting metastatic disease within 6 months of diagnosis	79 (26)
Primary tumour in situ for > 3 months (or incomplete removal) – no imaging capable of detecting metastatic disease within 6 months of diagnosis	1 (1)
Other	0 (0)

a Participant death within 12-month follow-up recorded separately.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 8 Streamline C: final T and N stage based on consensus reference standard

TABLE 46 Streamline C: final T and N stage based on consensus reference standard

	N stage[a] (*n*)		
T stage[a]	N0	N1	N2
T1	11	0	0
T2	36	12	4
T3	76	62	39
T4	10	17	32

a 232 participants had histological proof for both T and N stage, six participants had histological proof for T stage only and one participant for N stage only.

Reproduced from Taylor *et al.*[25] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 9 Streamline C: local tumour spread based on consensus reference

TABLE 47 Streamline C: local tumour spread based on consensus reference standard

Variable	Yes, *n* (%)	No, *n* (%)	Histological proof, *n* (%)
Resection margin threatened	24 (8)	275 (92)	230 (77)
Extramural vascular invasion[a]	118 (40)	180 (60)	231 (77)

a Data unavailable for extramural vascular invasion in one participant.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 10 Streamline C: detailed organ sites of metastatic disease

TABLE 48 Streamline C: detailed organ sites of metastatic disease

Site	Participants with metastatic disease,[a,b] *n*	Histological proof, *n* (%)	Imaging diagnosis without histological proof, *n* (%)[a]		
			Characteristic imaging appearances	Growth on follow-up	Response to therapy
Liver	48	12 (25)	33 (92)	26 (72)	19 (53)
Lung	20	0 (0)	19 (95)	13 (65)	11 (55)
Pleura	0	–	–	–	–
Brain	0	–	–	–	–
Spleen	0	–	–	–	–
Adrenal	0	–	–	–	–
Kidney	0	–	–	–	–
Pancreas	0	–	–	–	–
Mesentery/peritoneum	7	2 (29)	4 (80)	4 (80)	1 (20)
Bowel	0	–	–	–	–
Neck/chest soft tissue	0	–	–	–	–
Abdomen/pelvis soft tissue	0	–	–	–	–
Limbs soft tissue	0	–	–	–	–
Nodal (metastatic)	11	2 (28)	8 (89)	7 (78)	3 (33)
Other[c]	2	1 (33)	2 (100)	2 (100)	0 (0)
Skull	0	–	–	–	–
Cervical spine	0	–	–	–	–
Thoracic spine	0	–	–	–	–
Lumbar spine	2	0 (0)	2 (100)	1 (50)	0 (0)
Pelvis	1	0 (0)	1 (100)	1 (100)	1 (100)
Sternum	0	–	–	–	–
Clavicle/scapula	0	–	–	–	–
Ribs	0	–	–	–	–
Upper limb	0	–	–	–	–
Lower limb	0	–	–	–	–

a Metastasis may fulfil more than one criterion.
b Participants may have more than one site of metastatic disease.
c One participant has three additional sites of metastatic disease (i.e. skin, ovary and vagina).

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 11 Streamline C: number of metastatic deposits per organ

TABLE 49 Streamline C: number of metastatic deposits per organ

	Number of participants[a]					
Site	0	1	2	3	> 3	Unknown/not measurable
Liver	251	11	8	4	25	0
Lung	279	4	8	2	6	0
Bone	296	1	1	0	0	1
Mesentery/peritoneum	292	0	3	0	0	4
Nodal (metastatic)	288	2	4	1	4	0
Other	297	1	0	0	1	0

a Participants may have more than one site of metastatic disease.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 12 Streamline C: maximum size of metastatic deposits according to organ

TABLE 50 Streamline C: maximum size of metastatic deposits according to organ

Site	Number of participants with metastatic disease[a]	Number of participants[a]			
		< 1cm	≥ 1cm	Not visible on any initial staging investigation	Unknown/not measurable
Liver	48	5	37	6	0
Lung	20	12	5	3	0
Bone	3	1	0	1	1
Mesentery/peritoneum	7	0	1	2	4
Nodal (metastatic)	11	2	8	1	0
Other[b]	2	0	1	2	0

a Each participant is categorised by their largest metastasis in each organ. Participants may have more than one site of metastatic disease.
b One participant has two additional sites of metastatic disease (i.e. skin, ovary and vagina).

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 13 Streamline C: equivocal results for metastatic disease

TABLE 51 Streamline C: equivocal results for metastatic disease

Location	Equivocal results (*n*,%)							
	Participants with metastatic disease[a]				Participants without metastatic disease[a]			
	WB-MRI staging pathway[b]	Standard staging pathway	Standard staging pathway plus WB-MRI[c]	WB-MRI alone[d]	WB-MRI staging pathway[b]	Standard staging pathway	Standard staging pathway plus WB-MRI[c]	WB-MRI alone[d]
Per participant	4	7	6	9	7	14	10	25
Non-skeletal sites[e]	12	26	N/A	23	8	18	N/A	26
Skeletal sites[e]	2	0	N/A	0	2	0	N/A	7

N/A, data were not collected.
a Participants by consensus reference standard.
b WB-MRI plus additional generated tests.
c WB-MRI as an additional test to standard pathway.
d WB-MRI as a single staging investigation.
e Per organ as multiple sites can have equivocal results.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 14 Streamline C: short-interval follow-up for equivocal findings

TABLE 52 Streamline C: short-interval follow-up for equivocal findings

Test	Number of participants		
	WB-MRI staging pathway[a]	Standard staging pathway	Standard staging pathway plus WB-MRI[b]
Neck, chest and abdomen CT	0	0	0
Chest CT	1	2	2
Abdomen and pelvis CT	0	0	0
Chest and abdomen CT	0	0	0
Chest, abdomen and pelvis CT	2	2	2
Liver MRI	2	3	2
PET-CT	1	1	2
Biopsy	2	0	1
Other	8	5	9

a WB-MRI plus additional generated tests.
b WB-MRI as an additional test to standard pathway.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 15 Streamline C: per-organ sensitivity and specificity for metastatic disease – standard staging pathways versus WB-MRI staging pathway against the consensus reference standard

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 53 Streamline C: per-organ sensitivity and specificity for metastatic disease – standard staging pathways vs. WB-MRI staging pathway against the consensus reference standard

Site	Sensitivity[a]				Specificity[a]			
	Number of participants with metastatic disease[b]	WB-MRI staging pathway,[c] % (95% CI)	Standard staging pathway, % (95% CI)	Difference: WB-MRI[c] vs. standard, % (95% CI)	Number of participants without metastatic disease[b]	WB-MRI staging pathway,[c] % (95% CI)	Standard staging pathway, % (95% CI)	Difference: WB-MRI[c] vs. standard, % (95% CI)
Liver	48	74 (60 to 84)	72 (57 to 83)	2 (–10 to 14)	251	99 (97 to 100)	99 (97 to 100)	0 (–2 to 2)
Lung	20	55 (34 to 74)	65 (43 to 82)	–10 (–35 to 15)	279	99 (96 to 99)	99 (96 to 99)	0 (–2 to 2)
Bone	3	33 (6 to 79)	0 (0 to 56)	33 (–53 to 120)	296	99 (98 to 100)	99 (98 to 100)	0 (–2 to 1)
Mesentery/peritoneum	7	14 (3 to 51)	14 (3 to 51)	0 (–34 to 34)	292	99 (97 to 100)	99 (97 to 100)	0 (–2 to 2)
Nodal (metastatic)	11	45 (21 to 72)	27 (10 to 57)	18 (–13 to 49)	288	97 (94 to 98)	98 (95 to 99)	–1 (–3 to 1)
Other	2	0 (0 to 66)	0 (0 to 66)	0 (–50 to 50)	297	98 (95 to 99)	98 (96 to 99)	0 (–3 to 2)

a Equivocal results considered positive for colonic tumours and negative for rectal tumours.
b Participants by consensus reference standard.
c WB-MRI plus additional generated tests.

Reproduced from Taylor *et al.*[25] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Appendix 16 Streamline C: per-participant agreement for tumour T stage – standard staging pathways versus WB-MRI staging pathway against the consensus reference standard in participants with histological proof

TABLE 54 Streamline C: per-participant agreement for tumour T stage – standard staging pathways vs. WB-MRI staging pathway against the consensus reference standard in participants with histological proof

T stage	Number of participants[a]	Staging pathway, *n* (%)		Difference: WB-MRI[b] vs. standard, % (95% CI)
		WB-MRI[b,c]	Standard[c]	
T1	11	3 (27)	3 (27)	0 (–19 to 19)
T2	48	33 (69)	25 (52)	17 (3 to 31)
T3	134	69 (52)	83 (62)	–10 (–20 to –1)
T4	43	7 (16)	10 (23)	–7 (–18 to 4)
Total	236	112 (47)	121 (51)	–4 (–10 to 3)

a 238 participants with histological proof by final consensus reference standard but two participants missing full data.
b WB-MRI plus additional generated tests.
c 11 and 6 participants staged Tx by WB-MRI and standard staging pathways, respectively.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 17 Streamline C: per-participant agreement for tumour N stage – standard staging pathways versus WB-MRI staging pathway against the consensus reference standard in participants with histological proof

TABLE 55 Streamline C: per-participant agreement for tumour N stage – standard staging pathways vs. WB-MRI staging pathway against the consensus reference standard in participants with histological proof

N stage	Number of participants[a]	Staging pathway, *n* (%)		Difference: WB-MRI[b] vs. standard, % (95% CI)
		WB-MRI[b]	Standard	
N0	120	73 (61)	71 (59)	2 (–6 to 10)
N1	64	30 (47)	28 (44)	3 (–8 to 14)
N2	46	23 (50)	20 (43)	7 (–7 to 20)
Total	230	126 (55)	119 (52)	3 (–3 to 9)

a 233 participants with histological proof by final consensus reference standard but three participants were missing full data.
b WB-MRI plus additional generated tests.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 18 Streamline C: breakdown of the primary treatment decision according to the standard staging pathway compared with the retrospective consensus panel optimal treatment decision (non-rectal colon cancer)

TABLE 56 Streamline C: breakdown of the primary treatment decision according to the standard staging pathway compared with the retrospective consensus panel optimal treatment decision (non-rectal colon cancer)

	Retrospective consensus panel optimal treatment decision (*n*)				
Standard pathway	Surgery for the primary tumour but no chemotherapy	Surgery for the primary tumour and chemotherapy (and/or radiotherapy)	Chemotherapy (and/or radiotherapy) without surgery	Surgical metastectomy with or without chemotherapy	Total
Surgery for the primary tumour but no chemotherapy	83	25	3	2	113
Surgery for the primary tumour and chemotherapy (and/or radiotherapy)	8	23	2	1	34
Chemotherapy (and/or radiotherapy) without surgery	1	2	8	0	11
Surgical metastectomy with or without chemotherapy	0	3	2	5	10
Total	92	53	15	8	168

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 19 Streamline C: breakdown of the primary treatment decision according to the WB-MRI staging pathway compared with the retrospective consensus panel optimal treatment decision (non-rectal colon cancer)

TABLE 57 Streamline C: breakdown of the primary treatment decision according to the WB-MRI staging pathway compared with the retrospective consensus panel optimal treatment decision (non-rectal colon cancer)

	Retrospective consensus panel optimal treatment decision (*n*)				
WB-MRI pathway	Surgery for the primary tumour but no chemotherapy	Surgery for the primary tumour and chemotherapy (and/or radiotherapy)	Chemotherapy (and/or radiotherapy) without surgery	Surgical metastectomy with or without chemotherapy	Total
Surgery for the primary tumour but no chemotherapy	82	23	3	2	110
Surgery for the primary tumour and chemotherapy (and/or radiotherapy)	9	25	2	1	37
Chemotherapy (and/or radiotherapy) without surgery	1	2	8	0	11
Surgical metastectomy with or without chemotherapy	0	3	2	5	10
Total	92	53	15	8	168

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 20 Streamline C: breakdown of the primary treatment decision according to the standard staging pathway compared with the retrospective consensus panel optimal treatment decision (rectal cancer)

TABLE 58 Streamline C: breakdown of the primary treatment decision according to the standard staging pathway compared with the retrospective consensus panel optimal treatment decision (rectal cancer)

	Retrospective consensus panel optimal treatment decision (*n*)				
Standard pathway	Surgery for the primary tumour but no chemotherapy	Surgery for the primary tumour and chemotherapy (and/or radiotherapy)	Chemotherapy (and/or radiotherapy) without surgery	Surgical metastectomy with or without chemotherapy	Total
Surgery for the primary tumour but no chemotherapy	37	13	1	1	52
Surgery for the primary tumour and chemotherapy (and/or radiotherapy)	5	37	7	2	51
Chemotherapy (and/or radiotherapy) without surgery	1	9	6	4	20
Surgical metastectomy with or without chemotherapy	0	2	1	2	5
Total	43	61	15	9	128

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 21 Streamline C: breakdown of the primary treatment decision according to the WB-MRI staging pathway compared with the retrospective consensus panel optimal treatment decision (rectal cancer)

TABLE 59 Streamline C: breakdown of the primary treatment decision according to the WB-MRI staging pathway compared with the retrospective consensus panel optimal treatment decision (rectal cancer)

	Retrospective consensus panel optimal treatment decision (*n*)				
WB-MRI pathway	Surgery for the primary tumour but no chemotherapy	Surgery for the primary tumour and chemotherapy (and/or radiotherapy)	Chemotherapy (and/or radiotherapy) without surgery	Surgical metastectomy with or without chemotherapy	Total
Surgery for the primary tumour but no chemotherapy	37	13	1	0	51
Surgery for the primary tumour and chemotherapy (and/or radiotherapy)	5	36	7	3	51
Chemotherapy (and/or radiotherapy) without surgery	1	9	6	4	20
Surgical metastectomy with or without chemotherapy	0	3	1	2	6
Total	43	61	15	9	128

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 22 Streamline C: time to complete staging according to staging pathway: interquartile range

TABLE 60 Time to complete staging according to staging pathway: interquartile range

	Staging pathway, median days (IQR)	
Participants	WB-MRI	Standard
All participants	8 (4–13)	13 (8–25)

IQR, interquartile range.

Reproduced from Taylor *et al.*[25] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 23 Streamline C: per-participant sensitivity and specificity for metastatic disease – standard staging pathway versus WB-MRI alone against the consensus reference standard

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 61 Streamline C: per-participant sensitivity and specificity for metastatic disease – standard staging pathway vs. WB-MRI alone against the consensus reference standard

	Sensitivity				Specificity			
Outcome	Number of participants with metastatic disease[a]	WB-MRI alone,[b] % (95% CI)	Standard staging pathway, % (95% CI)	Difference: WB-MRI[b] vs. standard, % (95% CI)	Number of participants without metastatic disease[a]	WB-MRI alone,[b] % (95% CI)	Standard staging pathway, % (95% CI)	Difference: WB-MRI[b] vs. standard, % (95% CI)
Diagnostic accuracy[c]	68	70 (59 to 80)	63 (51 to 74)	7 (–3 to 18)	231	86 (81 to 90)	94 (90 to 96)	–8 (–13 to –3)

a Participants by consensus reference standard.
b WB-MRI as a single staging investigation.
c Equivocal results considered positive for colonic tumours and negative for rectal tumours.

Reproduced from Taylor *et al.*[25] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Appendix 24 Streamline C: per-organ sensitivity and specificity for metastatic disease – standard staging pathways versus WB-MRI alone against the consensus reference standard

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 62 Streamline C: per-organ sensitivity and specificity for metastatic disease – standard staging pathways vs. WB-MRI alone against the consensus reference standard

	Sensitivity				Specificity			
Site	Number of participants with metastatic disease[a]	WB-MRI alone[b], % (95% CI)	Standard staging pathway, % (95% CI)	Difference: WB-MRI[b] vs. standard, % (95% CI)	Number of participants without metastatic disease[a]	WB-MRI alone[b], % (95% CI)	Standard staging pathway, % (95% CI)	Difference: WB-MRI[b] vs. standard, % (95% CI)
Liver	48	77 (63 to 87)	73 (59 to 83)	4 (–8 to 16)	251	98 (95 to 99)	99 (96 to 100)	–1 (–4 to 1)
Lung	20	60 (39 to 78)	65 (43 to 82)	–5 (–28 to 18)	279	98 (96 to 99)	98 (96 to 99)	0 (–3 to 2)
Bone	3	0 (0 to 56)	0 (0 to 56)	0 (–33 to 33)	296	99 (96 to 100)	100 (98 to 100)	–1 (–2 to 1)
Mesentery/ peritoneum	7	14 (3 to 51)	14 (3 to 51)	0 (–34 to 34)	292	99 (96 to 99)	99 (97 to 100)	0 (–2 to 2)
Nodal (metastatic)	11	54 (28 to 79)	27 (10 to 57)	27 (3 to 52)	288	92 (88 to 94)	98 (96 to 99)	–6 (–10 to –3)
Other	2	0 (0 to 66)	0 (0 to 66)	0 (–50 to 50)	297	97 (95 to 99)	98 (96 to 99)	–1 (–3 to 2)

a Participants by consensus reference standard.
b WB-MRI as a single staging investigation.

Appendix 25 Streamline C: per-participant agreement for tumour T stage – standard staging pathways versus WB-MRI alone against the consensus reference standard – all participants

TABLE 63 Streamline C: per-participant agreement for tumour T stage – standard staging pathways vs. WB-MRI alone against the consensus reference standard – all participants

T stage	Number of participants[a]	Staging pathway, *n* (%)		Difference: WB-MRI[b] vs. standard, % (95% CI)
		WB-MRI alone[b,c]	Standard[c]	
T1	6	1 (16)	2 (33)	–17 (–63 to 30)
T2	40	23 (57)	23 (57)	0 (–19 to 19)
T3	148	71 (48)	108 (73)	–25 (–35 to –15)
T4	49	17 (35)	22 (45)	–10 (–22 to 2)
Total	243	112 (46)	155 (64)	–18 (–25 to –10)

a 54 and 3 participants missing data by WB-MRI alone and standard staging pathways, respectively.
b WB-MRI as a single staging investigation.
c 33 and 6 participants staged Tx by WB-MRI alone and standard staging pathways, respectively.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 26 Streamline C: per-participant agreement for tumour T stage – standard staging pathways versus WB-MRI alone against the consensus reference standard in participants with histological proof

TABLE 64 Streamline C: per-participant agreement for tumour T stage – standard staging pathways vs. WB-MRI alone against the consensus reference standard in participants with histological proof

T stage	Number of participants[a]	Staging pathway, *n* (%)		Difference: WB-MRI[b] vs. standard, % (95% CI)
		WB-MRI alone[b,c]	Standard[c]	
T1	6	1 (17)	2 (33)	–17 (–63 to 30)
T2	37	21 (57)	20 (54)	3 (–17 to 23)
T3	114	48 (42)	76 (67)	–25 (–37 to –13)
T4	37	7 (19)	10 (27)	–8 (–23 to 7)
Total	194	77 (40)	108 (56)	–16 (–25 to –7)

a 238 participants with histological proof by final consensus reference standard; 44 and one participants missing data for T stage by WB-MRI alone and standard staging pathways, respectively.
b WB-MRI as a single staging investigation.
c 29 and one participants staged Tx by WB-MRI alone and standard staging pathways, respectively.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 27 Streamline C: per-participant agreement for tumour N stage – standard staging pathways versus WB-MRI alone against the consensus reference standard – all participants

TABLE 65 Streamline C: per-participant agreement for tumour N stage – standard staging pathways vs. WB-MRI alone against the consensus reference standard – all participants

N stage	Number of participants[a]	Staging pathway, *n* (%)		Difference: WB-MRI[b] vs. standard, % (95% CI)
		WB-MRI alone[b,c]	Standard[c]	
N0	125	59 (47)	75 (60)	–13 (–22 to –4)
N1	84	32 (38)	43 (51)	–13 (–24 to –2)
N2	69	35 (51)	37 (54)	–3 (–17 to 11)
Total	278	126 (45)	155 (55)	–10 (–17 to –4)

a Information available for both standard pathway and WB-MRI alone pathway; three participants missing standard pathway N stage and 18 participants missing WB-MRI alone pathway N stage.
b WB-MRI as a single staging investigation.
c 37 have staged Nx by WB-MRI alone.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 28 Streamline C: per-participant agreement for tumour N stage – standard staging pathways versus WB-MRI alone against the consensus reference standard in participants with histological proof

TABLE 66 Streamline C: per-participant agreement for tumour N stage – standard staging pathways vs. WB-MRI alone against the consensus reference standard in participants with histological proof

N stage	Number of participants[a]	Staging pathway, *n* (%)		Difference: WB-MRI[b] vs. standard, % (95% CI)
		WB-MRI alone[b,c]	Standard	
N0	114	52 (46)	66 (58)	–12 (–22 to –3)
N1	61	21 (34)	27 (44)	–10 (–22 to 2)
N2	45	18 (40)	20 (44)	–4 (–22 to 13)
Total	220	91 (41)	113 (51)	–10 (–17 to –3)

a 233 participants with histological proof by final consensus reference standard; 12 and one participant(s) missing data for N stage by WB-MRI alone and standard staging pathways, respectively.
b WB-MRI as a single staging investigation.
c 28 participants have Nx stage by WB-MRI alone.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 29 Streamline C: per-participant sensitivity and specificity for metastatic disease: chest, abdomen and pelvis CT versus WB-MRI alone against the consensus reference standard

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 67 Streamline C: per-participant sensitivity and specificity for metastatic disease: chest, abdomen and pelvis CT vs. WB-MRI alone, against the consensus reference standard

Outcome	Sensitivity				Specificity			
	Number of participants with metastatic disease[a]	WB-MRI[b], % (95% CI)	Chest, abdomen and pelvis CT, % (95% CI)	Difference: WB-MRI[b] vs. standard, % (95% CI)	Number of participants without metastatic disease[a]	WB-MRI[b], % (95% CI)	Chest, abdomen and pelvis CT, % (95% CI)	Difference: WB-MRI[b] vs. standard, % (95% CI)
Diagnostic accuracy[c]	50	76 (63 to 86)	50 (37 to 63)	26 (13 to 39)	188	86 (81 to 90)	99 (97 to 100)	–13 (–18 to –9)

a Participants by consensus reference standard.
b WB-MRI as a single staging investigation.
c Equivocal results considered disease positive for non-rectal tumours and negative for rectal tumours.

Appendix 30 Streamline C: per-participant sensitivity and specificity for metastatic disease – standard staging pathways versus standard staging pathways plus WB-MRI against the consensus reference standard

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 68 Streamline C: per-participant sensitivity and specificity for metastatic disease – standard staging pathways vs. standard staging pathways plus WB-MRI against the consensus reference standard

	Sensitivity				Specificity			
Outcome	Number of participants with metastatic disease[a]	Standard staging pathway plus WB-MRI, % (95% CI)[b]	Standard staging pathway, % (95% CI)	Difference: WB-MRI[b] vs. standard, % (95% CI)	Number of participants without metastatic disease[a]	Standard staging pathway plus WB-MRI, % (95% CI)[b]	Standard staging pathway, % (95% CI)	Difference: WB-MRI[b] vs. standard
Diagnostic accuracy[c]	231	71 (59 to 80)	63 (51 to 73)	7 (0 to 14)	68	95 (91 to 97)	94 (90 to 96)	1 (–2 to 5)

a Participants by consensus reference standard.
b WB-MRI as an additional test to standard pathway.
c Equivocal results considered positive for colonic tumours and negative for rectal tumours.

Appendix 31 Streamline C: perceptual errors for metastatic disease according to staging pathway

TABLE 69 Streamline C: perceptual errors for metastatic disease according to staging pathway

Location	Perceptual errors[a] (*n*)		
	WB-MRI staging pathway[b]	Standard staging pathway	WB-MRI alone[c]
Per participant	16	15	16
Non-skeletal sites	18	15	18
Skeletal sites	0	1	0

a Perceptual errors are by definition only in disease-positive participants.
b WB-MRI plus additional generated tests.
c WB-MRI as a single staging investigation.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 32 Streamline C: perceptual errors for metastatic disease according to staging pathway and organ disease site

TABLE 70 Streamline C: perceptual errors for metastatic disease according to staging pathway and organ disease site

Site	Perceptual errors[a] (*n*)			
	Number of participants with metastatic disease[b]	WB-MRI staging pathway[c]	Standard staging pathway	WB-MRI alone[d]
Liver	48	7	5	7
Lung	20	3	4	3
Bone	3	0	1	0
Mesentery/peritoneum	7	4	1	4
Nodal (metastatic)	11	5	4	5
Other	2	0	1	0

a Perceptual error are by definition only in disease-positive participants.
b Participants by consensus reference standard.
c WB-MRI plus additional generated tests.
d WB-MRI as a single staging investigation.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 33 WB-MRI image quality scoring

TABLE 71 WB-MRI image quality scoring

Image quality score	Streamline C, *n* (%)	Streamline L, *n* (%)	Overall, *n* (%)
Optimal[a]	165 (51)	108 (39)	273 (46)
Suboptimal[b]	124 (39)	130 (47)	254 (43)
Degraded[c]	33 (10)	36 (13)	69 (11)
Total	322	274	596

a Technical quality score 5, anatomical coverage score 4.
b Technical quality score ≥ 3, anatomical coverage score ≥ 2.
c Technical quality score < 3, anatomical coverage score < 2.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 34 Detailed image quality score categories for scans rated as overall suboptimal category

TABLE 72 Detailed image quality score categories for scans rated as overall suboptimal category

Image quality score	Streamline C (*n*)	Streamline L (*n*)	Overall (*n*)
A4 T4	49	42	91
A4 T3	38	44	82
A3 T5	10	15	25
A3 T4	10	12	22
A3 T3	17	17	34

A, anatomical coverage score; T, technical quality score.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 35 Detailed image quality score categories for scans rated as overall degraded category

TABLE 73 Detailed image quality score categories for scans rated as overall degraded category

Image quality score	Streamline C (*n*)	Streamline L (*n*)	Overall (*n*)
A4 T2	3	1	4
A4 T1	3	1	4
A3 T2	0	1	1
A2 T5	5	4	9
A2 T4	0	7	7
A2 T3	8	2	10
A1 T5	5	6	11
A1 T4	2	2	4
A1 T3	6	4	10
A1 T2	0	2	2
A1 T1	4	3	7

A, anatomical coverage score; T, technical quality score.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 36 Streamline L: reasons for non-inclusion of screened patients

TABLE 74 Streamline L: reasons for non-inclusion of screened patients

Reason for non-inclusion	*n* (%)
Evidence of severe or uncontrolled systematic disease	113 (18)
Unequivocal metastatic or N3 disease	110 (18)
Patients did not want extra visit for WB-MRI	86 (14)
Patients felt WB-MRI would delay potential treatment	62 (10)
Histology other than NSCLC	53 (8)
Did not want WB-MRI scan	52 (8)
Contraindications to WB-MRI	47 (8)
Site error/missed opportunity to undergo WB-MRI	29 (5)
Further staging work up not indicated in opinion of care team	27 (4)
Too frail/not fit enough to participate	20 (3)
Language barrier	14 (2)
Psychiatric or other conditions likely to impact on informed consent	10 (2)
Total	623

NSCLC, non-small-cell lung cancer.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 37 Streamline L: reasons for withdrawal of recruited participants

TABLE 75 Streamline L: reasons for withdrawal of recruited participants

Reason for withdrawal	*n* (%)
Participant withdrawal owing to non-small-cell lung cancer diagnosis	84 (50)
Participant withdrawal owing to failure to undergo WB-MRI	75 (45)
Participant withdrew consent to follow-up	4 (2)
Participant lost to follow-up	1 (1)
Participant withdrawal owing to early release of WB-MRI findings	1 (1)
Participant died after registration	1 (1)
Total	166

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 38 Streamline L: primary tumour status at the time of consensus review

TABLE 76 Streamline L: primary tumour status at the time of consensus review

Participant status	*n* (%)
Participant died within 12 months[a]	31 (17)
Primary tumour removed within 3 months of diagnosis	93 (50)
Primary tumour in situ for > 3 months (or incomplete removal) – imaging capable of detecting metastatic disease within 6 months of diagnosis	80 (43)
Primary tumour in situ for > 3 months (or incomplete removal) – no imaging capable of detecting metastatic disease within 6 months of diagnosis	14 (7)
Other	0 (0)

a Participant death within 12-month follow-up recorded separately.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 39 Streamline L: final T and N stage based on consensus reference standard

TABLE 77 Streamline L: final T and N stage based on consensus reference standard

	N stage[a] (*n*)			
T stage[a]	N0	N1	N2	N3
T1a	29	1	1	0
T1b	12	2	4	1
T2a	33	7	10	2
T2b	15	2	5	1
T3	10	6	9	8
T4	11	3	7	8

a 87 participants with histological proof for both T and N stage, 13 participants with histological proof for T stage only and 24 participants with histological proof for N stage only.

Reproduced from Taylor *et al.*[26] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 40 Streamline L: final nodal disease sites based on the consensus reference standard

TABLE 78 Streamline L: final nodal disease sites based on the consensus reference standard

Site	Number of participants with metastatic disease,[a] *n*	Histological proof, *n* (%)
Supraclavicular, low cervical and sterna notch	8	1 (13)
Upper paratracheal (right)	15	2 (13)
Upper paratracheal (left)	4	0 (0)
Pre vascular	3	0 (0)
Retrotracheal	1	0 (0)
Lower paratracheal (right)	22	10 (45)
Lower paratracheal (left)	10	0 (0)
Subaortic	10	4 (40)
Para-aortic	11	2 (18)
Subcarnal	27	12 (44)
Paraesophageal (below carina)	1	1 (100)
Pulmonary ligament	3	3 (100)
Hilar (right)	29	7 (24)
Hilar (left)	25	13 (52)
Interlobar (right)	3	2 (67)
Interlobar (left)	4	3 (75)
Lobar (right)	3	2 (67)
Lobar (left)	2	0 (0)
Segmental (right)	0	–
Segmental (left)	2	2 (100)
Subsegmental (left)	0	–
Subsegmental (right)	0	–

a Participants may have more than one site of metastatic disease.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 41 Streamline L: detailed organs sites of metastatic disease

TABLE 79 Streamline L: detailed organs sites of metastatic disease

Site	Number of participants with metastatic disease[b]	Histological proof, *n* (%)	Imaging diagnosis without histological proof,[a] *n* (%)		
			Characteristic imaging appearances	Growth on follow-up	Response to therapy
Liver	9	2 (22)	6 (86)	4 (57)	1 (14)
Lung	14	0 (0)	14 (100)	11 (79)	3 (21)
Pleura	5	1 (20)	4 (100)	2 (50)	1 (25)
Brain	14	1 (7)	13 (100)	8 (62)	4 (31)
Spleen	0	–	–	–	–
Adrenal	7	0 (0)	6 (86)	5 (71)	1 (14)
Kidney	1	1 (100)	–	–	–
Pancreas	2	0 (0)	2 (100)	1 (50)	0 (0)
Mesentery/ peritoneum	0	–	–	–	–
Bowel	0	–	–	–	–
Neck/chest soft tissue	1	0 (0)	1 (100)	0 (0)	0 (0)
Abdomen/pelvis soft tissue	3	0 (0)	3 (100)	1 (33)	0 (0)
Limbs soft tissue	1	0 (0)	1 (100)	0 (0)	0 (0)
Nodal (metastatic)	3	1 (33)	2 (100)	2 (100)	0 (0)
Other[c]	2	0 (0)	3 (100)	1 (33)	0 (0)
Skull	1	0 (0)	1 (100)	1 (100)	0 (0)
Cervical spine	0	–	–	–	–
Thoracic spine	9	0 (0)	9 (100)	5 (56)	1 (11)
Lumbar spine	8	1 (13)	7 (100)	3 (43)	0 (0)
Pelvis	5	0 (0)	5 (100)	2 (40)	0 (0)
Sternum	2	0 (0)	2 (100)	0 (0)	0 (0)
Clavicle/scapula	0	–	–	–	–
Ribs	3	0 (0)	3 (100)	0 (0)	2 (67)
Upper limb	2	0 (0)	2 (100)	1 (50)	0 (0)
Lower limb	1	0 (0)	1 (100)	0 (0)	0 (0)

a Metastasis may fulfil more than one criterion for imaging diagnosis.
b Participants may have more than one site of metastatic disease.
c Two participants with metastasis in three other sites (right thyroid, left axilla and leptomeningeal).

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 42 Streamline L: number of metastatic deposits per organ

TABLE 80 Streamline L: number of metastatic deposits per organ

	Number of participants[a]					
Site	0	1	2	3	> 3	Unknown/not measurable
Liver	178	5	3	0	1	1
Lung	173	6	3	1	4	0
Pleura	182	1	3	0	1	3
Brain	173	9	3	1	1	0
Adrenal	180	5	2	0	0	0
Bone	172	7	0	2	5	1
Other	178	4	3	1	1	1

a Participants may have more than one site of metastatic disease.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 43 Streamline L: maximum size of metastatic deposits according to organ

TABLE 81 Streamline L: maximum size of metastatic deposits according to organ

Site	Number of participants with metastatic disease[a]	Number of participants[a]			
		< 1 cm	≥ 1 cm	Not visible on any staging investigation	Unknown/ not measurable
Liver	9	3	3	2	1
Lung	14	4	3	7	0
Pleura	5	1	1	0	3
Brain	14	3	7	4	0
Adrenal	7	1	3	3	0
Bone	15	3	11	0	1
Other	9	1	7	0	1

a Participants may have more than one site of metastatic disease.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 44 Streamline L: equivocal results for metastatic disease

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 82 Streamline L: equivocal results for metastatic disease

Location	Equivocal results							
	Participants with metastatic disease[a]				Participants without metastatic disease[a]			
	WB-MRI staging pathway[b]	Standard staging pathway	Standard staging pathway plus WB-MRI[c]	WB-MRI alone[d]	WB-MRI staging pathway [b]	Standard staging pathway	Standard staging pathway plus WB-MRI[c]	WB-MRI alone[d]
Per participant	1	4	2	2	1	2	1	17
Non skeletal sites[e]	3	11	N/A	8	8	17	N/A	17
Skeletal sites[e]	4	4	N/A	2	1	0	N/A	4

N/A, data were not collected.
a Participants by consensus reference standard.
b WB-MRI plus additional generated tests.
c WB-MRI as an additional test to standard pathway.
d WB-MRI as a single staging investigation.
e Per organ as multiple sites can have equivocal results.

Appendix 45 Streamline L: short-interval follow-up for equivocal findings

TABLE 83 Streamline L: short-interval follow-up for equivocal findings

	Number of participants		
Test	WB-MRI plus additional tests	Standard imaging	All tests combined
Chest CT	0	0	0
Neck, chest and abdomen CT	0	0	0
Liver MRI	0	0	0
Bone scan	0	0	0
Spine and bone MRI	0	1	1
Other	2	3	2

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 46 Streamline L: per-organ sensitivity and specificity for metastatic disease – standard staging pathways versus WB-MRI staging pathway against the consensus reference standard

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 84 Streamline L: per-organ sensitivity and specificity for metastatic disease – standard staging pathways vs. WB-MRI staging pathway against the consensus reference standard

	Sensitivity				Specificity			
Site	Number of participants with metastatic disease[a]	WB-MRI staging pathway[b], % (95% CI)	Standard staging pathway, % (95% CI)	Difference: WB-MRI[b] vs. standard, % (95% CI)	Number of participants without metastatic disease[a]	WB-MRI staging pathway[b], % (95% CI)	Standard staging pathway, % (95% CI)	Difference: WB-MRI[b] vs. standard, % (95% CI)
Liver	9	44 (19 to 73)	33 (12 to 65)	11 (–23 to 46)	178	99 (97 to 100)	100 (98 to 100)	–1 (–2 to 1)
Lung	14	36 (16 to 61)	36 (16 to 61)	0 (–22 to 22)	173	97 (94 to 99)	98 (95 to 99)	–1 (–4 to 3)
Pleura	5	80 (38 to 96)	40 (12 to 77)	40 (7 to 73)	182	97 (94 to 99)	98 (95 to 99)	–1 (–4 to 2)
Brain	14	57 (33 to 79)	44 (21 to 67)	13 (–5 to 33)	173	99 (97 to 100)	100 (98 to 100)	–1 (–2 to 1)
Adrenal	7	28 (8 to 64)	57 (25 to 84)	–29 (–71 to 14)	180	98 (95 to 99)	98 (95 to 99)	0 (–3 to 3)
Bone	15	60 (36 to 80)	67 (42 to 85)	–7 (–31 to 18)	172	99 (96 to 100)	99 (96 to 100)	0 (–3 to 3)
Other	9	22 (6 to 55)	44 (19 to 73)	–22 (–56 to 12)	178	97 (94 to 99)	93 (89 to 96)	4 (0 to 8)

a Participants by consensus reference standard.
b WB-MRI plus additional generated tests.

Reproduced from Taylor *et al.*[26] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Appendix 47 Streamline L: per-participant agreement for tumour T stage – standard staging pathways versus WB-MRI staging pathway against the consensus reference standard in participants with histological proof

TABLE 85 Streamline L: per-participant agreement for tumour T stage – standard staging pathways vs. WB-MRI staging pathway against the consensus reference standard in participants with histological proof

T stage	Number of participants[a]	Staging pathway, *n* (%)		Difference: WB-MRI[b] vs. standard, % (95% CI)
		WB-MRI[b,c]	Standard[c]	
T1a	21	20 (95)	20 (95)	0 (–16 to 16)
T1b	11	5 (45)	6 (54)	–9 (–31 to 13)
T2a	37	10 (27)	14 (38)	–11 (–26 to 5)
T2b	14	5 (36)	6 (43)	–7 (–39 to 24)
T3	11	4 (36)	3 (27)	9 (–24 to 42)
T4	2	0 (0)	2 (100)	–100 (–100 to –50)
Overall T stage	96	44 (46)	51 (53)	–7 (–17 to 2)

a Participants with histological proof by final consensus reference standard.
b WB-MRI plus additional generated tests.
c One participant each staged Tx by WB-MRI and standard staging pathways.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 48 Streamline L: per-participant agreement for tumour N stage – standard staging pathways versus WB-MRI staging pathway against the consensus reference standard in participants with histological proof

TABLE 86 Streamline L: per-participant agreement for tumour N stage – standard staging pathways vs. WB-MRI staging pathway against the consensus reference standard in participants with histological proof

N stage	Number of participants[a]	Staging pathway, *n* (%)		Difference: WB-MRI[b] vs. standard, % (95% CI)
		WB-MRI[b]	Standard[c]	
N0	70	56 (80)	60 (86)	–6 (–15 to 4)
N1	19	3 (16)	6 (32)	–16 (–41 to 9)
N2	15	6 (40)	10 (67)	–27 (–61 to 8)
N3	5	3 (60)	3 (60)	0 (–46 to 46)
Overall N stage	109	68 (62)	79 (72)	–10 (–19 to –1)

a Participants with histological proof by final consensus reference standard.
b WB-MRI plus additional generated tests.
c One participant staged Nx by standard staging pathway.

Reproduced from Taylor *et al.*[26] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 49 Streamline L: nodal site agreement – standard staging pathways versus WB-MRI staging pathway against the consensus reference standard in participants with histological proof

TABLE 87 Streamline L: nodal site agreement – standard staging pathways vs. WB-MRI staging pathway against the consensus reference standard in participants with histological proof

Site	Number of participants[a]	Staging pathway, *n* (%)		Difference: WB-MRI[b] vs. standard, % (95% CI)
		WB-MRI[b]	Standard	
Supraclavicular, low cervical and sterna notch	2	1 (50)	1 (50)	0 (–57 to 57)
Upper paratracheal (right)	12	2 (17)	2 (17)	0 (–22 to 22)
Upper Paratracheal (left)	2	0 (0)	0 (0)	0 (–50 to 50)
Pre vascular	2	0 (0)	0 (0)	0 (–50 to 50)
Retrotracheal	2	0 (0)	1 (50)	–50 (–169 to 69)
Lower paratracheal (right)	47	12 (25)	11 (23)	2 (–10 to 14)
Lower paratracheal (left)	3	0 (0)	1 (33)	–33 (–120 to 53)
Subaortic	29	5 (17)	4 (14)	3 (–12 to 19)
Para-aortic	9	5 (56)	6 (67)	–11 (–50 to 28)
Subcarinal	62	20 (32)	17 (27)	–5 (–4 to 14)
Paraesophageal	21	4 (19)	3 (14)	5 (–12 to 21)
Pulmonary ligament	30	6 (20)	6 (20)	0 (–15 to 15)
Hilar (right)	40	7 (18)	3 (8)	10 (–3 to 23)
Hilar (left)	32	6 (19)	8 (25)	–6 (–17 to 5)
Interlobar (right)	12	2 (17)	2 (17)	0 (–29 to 29)
Interlobar (left)	11	1 (9)	2 (18)	–9 (–37 to 19)
Lobar (right)	6	0 (0)	0 (0)	0 (–17 to 17)
Lobar (left)	3	0 (0)	0 (0)	0 (–33 to 33)
Segmental (left)	1	0 (0)	0 (0)	0 (–100 to 100)
Segmental (right)	3	0 (0)	0 (0)	0 (–33 to 33)
Subsegmental (left)	1	0 (0)	0 (0)	0 (–100 to 100)
Subsegmental (right)	2	0 (0)	0 (0)	0 (–50 to 50)

a Participants with histological proof by final consensus reference standard.
b WB-MRI plus additional generated tests.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 50 Streamline L: breakdown of the primary treatment decision according to the standard staging pathway compared with the retrospective consensus panel optimal treatment decision

TABLE 88 Streamline L: breakdown of the primary treatment decision according to the standard staging pathway compared with the retrospective consensus panel optimal treatment decision

	Retrospective consensus panel optimal treatment decision (*n*)		
Standard pathway	**Curative intent**	**Palliative intent**	**Total[a]**
Curative intent	123	25	148
Palliative intent	7	28	35
Total[a]	130	53	183

a Four participants were missing at least one type of participant management decision.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 51 Streamline L: breakdown of the primary treatment decision according to the WB-MRI staging pathway compared with the retrospective consensus panel optimal treatment decision

TABLE 89 Streamline L: breakdown of the primary treatment decision according to the WB-MRI staging pathway compared with the retrospective consensus panel optimal treatment decision

WB-MRI pathway including additional tests generated	Retrospective consensus panel optimal treatment decision (*n*)		
	Curative intent	Palliative intent	Total[a]
Curative intent	122	23	145
Palliative intent	8	30	38
Total[a]	130	53	183

a Four participants were missing at least one type of participant management decision.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 52 Streamline L: time to complete staging according to staging pathway

TABLE 90 Streamline L: time to complete staging according to staging pathway

	Staging pathway, median days (IQR)	
Participants	WB-MRI	Standard
All participants	13 (7–19)	19 (13–31)

IQR, interquartile range.

Reproduced from Taylor *et al.*[26] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 53 Streamline L: per-participant sensitivity and specificity for metastatic disease-standard staging pathway versus WB-MRI alone against the consensus reference standard

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 91 Streamline L: per-participant sensitivity and specificity for metastatic disease-standard staging pathway vs. WB-MRI alone, against the consensus reference standard

	Sensitivity				Specificity			
Outcome	Number of participants with metastatic disease[a]	WB-MRI alone[b], % (95% CI)	Standard staging pathway, % (95% CI)	Difference: WB-MRI[b] vs. standard, % (95% CI)	Number of participants without metastatic disease[a]	WB-MRI alone[b], % (95% CI)	Standard staging pathway, % (95% CI)	Difference: WB-MRI[b] vs. standard, % (95% CI)
Diagnostic accuracy	52	50 (37 to 63)	54 (41 to 67)	–4 (–16 to 8)	135	85 (78 to 90)	95 (91 to 98)	–10 (–17 to –4)

a Participants by consensus reference standard.
b WB-MRI as a single staging investigation.
Reproduced from Taylor *et al.*[26] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. The text below includes minor additions and formatting changes to the original text.

Appendix 54 Streamline L: per-organ sensitivity and specificity for metastatic disease – standard staging pathways versus WB-MRI alone against the consensus reference standard

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 92 Streamline L: per-organ sensitivity and specificity for metastatic disease-standard staging pathways vs. WB-MRI alone against the consensus reference standard

	Sensitivity				Specificity			
Site	Number of participants with metastatic disease[a]	WB-MRI alone,[b] % (95% CI)	Standard staging pathway, % (95% CI)	Difference: WB-MRI[b] vs. standard, % (95% CI)	Number of participants without metastatic disease[a]	WB-MRI alone,[b] % (95% CI)	Standard staging pathway, % (95% CI)	Difference: WB-MRI[b] vs. standard, % (95% CI)
Liver	9	33 (12 to 65)	33 (12 to 65)	0 (–38 to 38)	178	98 (94 to 99)	100 (98 to 100)	–2 (–5 to 1)
Lung	14	21 (8 to 48)	35 (16 to 61)	–14 (–37 to 9)	173	97 (93 to 98)	98 (95 to 99)	–1 (–5 to 3)
Pleura	5	60 (23 to 88)	40 (12 to 77)	20 (–20 to 60)	182	95 (91 to 97)	98 (95 to 99)	–3 (–7 to 0)
Brain	14	57 (33 to 79)	43 (21 to 67)	14 (–5 to 33)	173	99 (97 to 100)	100 (98 to 100)	–1 (–2 to 1)
Adrenal	7	43 (16 to 75)	57 (25 to 84)	–14 (–48 to 19)	180	97 (93 to 98)	98 (95 to 99)	–1 (–4 to 2)
Bone	15	60 (36 to 80)	67 (42 to 85)	–7 (–31 to 18)	172	98 (94 to 99)	99 (96 to 100)	–1 (–4 to 2)
Other	9	33 (12 to 65)	44 (19 to 73)	–11 (–38 to 16)	178	96 (92 to 98)	93 (89 to 96)	3 (–2 to 8)

a Participants by consensus reference standard.
b WB-MRI as a single staging investigation.

Appendix 55 Streamline L: per-participant agreement for tumour T stage – standard staging pathways versus WB-MRI alone against the consensus reference standard

TABLE 93 Streamline L: per-participant agreement for tumour T stage – standard staging pathways vs. WB-MRI alone against the consensus reference standard

T stage	Number of participants[a]	Staging pathway, *n* (%)		Difference: WB-MRI[b] vs. standard, % (95% CI)
		WB-MRI alone[b,c]	Standard[c]	
T1a	26	23 (88)	24 (92)	–4 (–22 to 14)
T1b	18	10 (55)	8 (44)	11 (–13 to 36)
T2a	50	16 (32)	24 (48)	–16 (–33 to 1)
T2b	22	7 (32)	9 (41)	–9 (–35 to 16)
T3	32	17 (53)	17 (53)	0 (–22 to 22)
T4	28	10 (36)	16 (57)	–21 (–47 to 4)
Overall T stage	176	83 (47)	98 (56)	–9 (–17 to 0)

a 11 participants were missing data.
b WB-MRI as a single staging investigation.
c 11 and one participant(s) staged Tx by WB-MRI alone and standard staging pathways, respectively.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 56 Streamline L: per-participant agreement for tumour T stage – standard staging pathways versus WB-MRI alone against the consensus reference standard in participants with histological proof

TABLE 94 Streamline L: per-participant agreement for tumour T stage – standard staging pathways vs. WB-MRI alone against the consensus reference standard in participants with histological proof

T stage	Number of participants[a]	Staging pathway, *n* (%)		Difference: WB-MRI[b] vs. standard, % (95% CI)
		WB-MRI alone[b,c]	Standard[c]	
T1a	19	17 (89)	19 (100)	–11 (–29 to 9)
T1b	11	5 (45)	6 (54)	–9 (–31 to 13)
T2a	36	8 (22)	15 (41)	–19 (–38 to –1)
T2b	14	6 (43)	6 (43)	0 (–28 to 28)
T3	10	3 (30)	3 (30)	0 (–36 to 36)
T4	2	0 (0)	2 (100)	–100 (–100 to –50)
Overall T stage	92	39 (42)	51 (55)	–13 (–23 to –3)

a Participants with histological proof by final consensus reference standard.
b WB-MRI as a single staging investigation.
c Five and one participant(s) staged Tx by WB-MRI alone and standard staging pathways, respectively.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 57 Streamline L: per-participant agreement for tumour N stage – standard staging pathways versus WB-MRI alone against the consensus reference standard – all participants

TABLE 95 Streamline L: per-participant agreement for tumour N stage – standard staging pathways vs. WB-MRI alone against the consensus reference standard – all participants

N stage	Number of participants[a]	Staging pathway, *n* (%)		Difference: WB-MRI[b] vs. standard, % (95% CI)
		WB-MRI alone[b,c]	Standard[c]	
N0	104	79 (76)	88 (85)	–9 (–17 to 0)
N1	21	4 (19)	8 (38)	–19 (–46 to 8)
N2	36	19 (53)	26 (72)	–19 (–42 to 3)
N3	20	6 (30)	12 (60)	–30 (–61 to 1)
Overall N stage	181	108 (60)	132 (74)	–14 (–22 to –6)

a 6 participants missing data.
b WB-MRI as a single staging investigation.
c 10 and one participant(s) staged Nx stage by WB-MRI alone and standard staging pathways, respectively.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 58 Streamline L: per-participant agreement for tumour N stage – standard staging pathways versus WB-MRI alone against the consensus reference standard in participants with histological proof

TABLE 96 Streamline L: per-participant agreement for tumour N stage – standard staging pathways vs. WB-MRI alone against the consensus reference standard in participants with histological proof

N stage	Number of participants[a]	Staging pathway, *n* (%)		Difference: WB-MRI[b] vs. standard, % (95% CI)
		WB-MRI alone[b,c]	Standard	
N0	69	54 (78)	58 (84)	–6 (–16 to 5)
N1	19	4 (21)	6 (32)	–11 (–37 to 16)
N2	15	5 (33)	10 (66)	–33 (–69 to 2)
N3	5	3 (60)	3 (60)	0 (–46 to 46)
Overall N stage	108	66 (61)	77 (71)	–10 (–20 to 0)

a Participants with histological proof by final consensus reference standard.
b WB-MRI as a single staging investigation.
c Four and one participant(s) staged Nx by WB-MRI alone and standard staging pathways, respectively.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 59 Streamline L: per-participant agreement for metastatic disease – PET-CT versus WB-MRI alone against the consensus reference standard

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 97 Streamline L: per-participant agreement for metastatic disease: PET-CT versus WB-MRI alone against the consensus reference standard

	Sensitivity				Specificity			
Outcome	Number of participants with metastatic disease[a]	WB-MRI alone,[b] % (95% CI)	PET-CT, % (95% CI)	Difference: WB-MRI[b] vs. PET-CT, % (95% CI)	Number of participants without metastatic disease[a]	WB-MRI alone,[b] % (95% CI)	PET-CT, % (95% CI)	Difference: WB-MRI[b] vs. PET-CT, % (95% CI)
Diagnostic accuracy[c]	47	47 (33 to 61)	53 (39 to 67)	–6 (–20 to 8)	124	85 (78 to 91)	90 (84 to 94)	–5 (–13 to 3)

a Participants by consensus reference standard.
b WB-MRI as a single staging investigation.
c Equivocal results considered disease positive.

Appendix 60 Streamline L: per-participant sensitivity and specificity for metastatic disease-standard staging pathways versus standard staging pathways plus WB-MRI against the consensus reference standard

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 98 Streamline L: per-participant sensitivity and specificity for metastatic disease-standard staging pathways vs. standard staging pathways plus WB-MRI against the consensus reference standard

	Sensitivity				Specificity			
Outcome	Number of participants with metastatic disease[a]	Standard staging pathway plus WB-MRI,[b] % (95% CI)	Standard staging pathway, % (95% CI)	Difference: WB-MRI[b] vs. standard, % (95% CI; *p*-value)	Number of participants without metastatic disease[a]	Standard staging pathway plus WB-MRI,[b] % (95% CI)	Standard staging pathway, % (95% CI)	Difference: WB-MRI[b] vs. standard, % (95% CI; *p*-value)
Metastatic disease	52	60 (46 to 72)	54 (41 to 67)	6 (–3 to 14; 0.375)	135	96 (91 to 98)	96 (91 to 98)	0 (–4 to 4; > 0.999)

a Participants by consensus reference standard.
b WB-MRI as an additional test to standard pathway.

Appendix 61 Streamline L: perceptual errors for metastatic disease according to staging pathway

TABLE 99 Streamline L: perceptual errors for metastatic disease according to staging pathway

Location	Perceptual errors[a] (*n*)		
	WB-MRI staging pathway[b]	Standard staging pathway	WB-MRI alone[c]
Per participant	18	7	19
Non skeletal sites	15	6	15
Skeletal sites	7	3	8

a Perceptual errors are by definition only in disease-positive participants.
b WB-MRI plus additional generated tests.
c WB-MRI as a single staging investigation.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 62 Streamline L: perceptual errors for metastatic disease according to staging pathway and organ disease site

TABLE 100 Streamline L: perceptual errors for metastatic disease according to staging pathway and organ disease site

Site	Perceptual errors[a] (*n*)			
	Number of participants with metastatic disease[b]	WB-MRI staging pathway[c]	Standard staging pathway	WB-MRI alone[d]
Liver	9	3	1	3
Lung	14	2	0	2
Pleura	5	0	2	0
Brain	14	3	0	3
Adrenal	7	1	1	1
Bone	15	7	3	8
Other	9	6	2	6

a Perceptual error are by definition only in disease-positive participants.
b Participants by consensus reference standard.
c WB-MRI plus additional generated tests.
d WB-MRI as a single staging investigation.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 63 Streamline C: concordance between treatment decisions based on standard and WB-MRI staging pathway (d1 and d2)

TABLE 101 Streamline C: concordance between treatment decisions based on standard and WB-MRI staging pathway (d1 and d2)

WB-MRI staging pathway (d2)	Standard staging pathway (d1) (%)				
	A	B	C	D	Total
A	10.47	0	0	0	10.47
B	0	26.69	3.04	0	29.73
C	0	1.69	52.7	0	54.39
D	0	0.34	0	5.07	5.41
Total	10.47	28.72	55.74	5.07	100

A, chemotherapy (and/or radiotherapy) without surgery; B, surgery for the primary tumour and chemotherapy (and/or radiotherapy); C, surgery for the primary tumour but no chemotherapy; D, surgical metastectomy with or without chemotherapy. Discordance = 1.69 + 0.34 + 3.04 = 5.07%.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 64 Streamline L: concordance between treatment decisions based on standard and WB-MRI staging pathway (d1 and d2)

TABLE 102 Streamline L: concordance between treatment decisions based on standard and WB-MRI staging pathway (d1 and d2)

	Standard staging pathway (d1) (%)		
WB-MRI staging pathway (d2)	X	Y	Total
X	78.69	0.55	79.23
Y	2.19	18.58	20.77
Total	80.87	19.13	100

X, treatments with curative intent; Y, treatments with non-curative intent.
Discordance = 2.19 + 0.55 = 2.74%.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 65 Streamline C: discordance between treatment decisions based on standard staging (d1) and the consensus reference panel optimal decision (d4)

TABLE 103 Streamline C: discordance between treatment decisions based on standard staging (d1) and the consensus reference panel optimal decision (d4)

Standard staging pathway (d1)	Panel consensus decision (d4) (%)				
	A	B	C	D	Total
A	4.73	3.72	0.68	1.35	10.47
B	3.04	20.27	4.39	1.01	28.72
C	1.35	12.84	40.54	1.01	55.74
D	1.01	1.69	0.00	2.36	5.07
Total	10.14	38.51	45.61	5.74	100

A, chemotherapy (and/or radiotherapy) without surgery; B, surgery for the primary tumour and chemotherapy (and/or radiotherapy); C, surgery for the primary tumour but no chemotherapy, D, surgical metastectomy with or without chemotherapy. Discordance = 3.04 + 1.35 + 1.01 + 3.72 + 12.84 + 1.69 + 0.68 + 4.39 + 1.35 + 1.01 + 1.01 = 32.09%.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 66 Streamline C: discordance between treatment decisions based on WB-MRI staging pathway (d1) and the consensus reference panel optimal decision (d4)

TABLE 104 Streamline C: discordance between treatment decisions based on WB-MRI staging pathway (d1) and the consensus reference panel optimal decision (d4)

WB-MRI staging pathway (d2)	Panel consensus decision (d4) (%)				
	A	B	C	D	Total
A	4.73	3.72	0.68	1.35	10.47
B	3.04	20.61	4.73	1.35	29.73
C	1.35	12.16	40.20	0.68	54.39
D	1.01	2.03	0.00	2.36	5.41
Total	10.14	38.51	45.61	5.74	100

A, chemotherapy (and/or radiotherapy) without surgery; B, surgery for the primary tumour and chemotherapy (and/or radiotherapy); C, surgery for the primary tumour but no chemotherapy; D, surgical metastectomy with or without chemotherapy. Discordance = 3.04 + 1.35 + 1.01 + 3.72 + 12.16 + 2.03 + 0.68 + 4.73 + 1.35 + 1.35 + 0.68 = 32.10%.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 67 Streamline L: discordance between treatment decisions based on standard staging (d1) and the consensus reference panel optimal decision (d4)

TABLE 105 Streamline L: discordance between treatment decisions based on standard staging (d1) and the consensus reference panel optimal decision (d4)

	Panel consensus decision (d4) (%)		
Standard staging pathway (d1)	X	Y	Total
X	67.21	13.66	80.87
Y	3.83	15.3	19.13
Total	71.04	28.96	100

X, treatments with curative intent; Y, treatments with non-curative intent.
Discordance = 3.83 + 13.66 = 17.49%.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 68 Streamline L: discordance between treatment decisions based on WB-MRI staging pathway (d1) and the consensus reference panel optimal decision (d4)

TABLE 106 Streamline L: discordance between treatment decisions based on WB-MRI staging pathway (d1) and the consensus reference panel optimal decision (d4)

	Panel consensus decision (d4) (%)		
WB-MRI staging pathway (d2)	X	Y	Total
X	66.67	12.57	79.23
Y	4.37	16.39	20.77
Total	71.04	28.96	100

X, treatments with curative intent; Y, treatments with non-curative intent.
Discordance = 4.37 + 12.57 = 16.94%.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 69 Interview topic guide

BOX 1 Interview topic guide

I understand that you have been having a number of scans to investigate your symptoms; can you tell me what's been happening?

Can you describe what it feels like physically to have the scan?

How did you feel during/after/about the test?

What were the staff like?

What did you do to cope with [problem identified by patient, e.g. noise, keeping still, etc.]?

Would you have another WB-MRI if the doctor recommended it?

Reproduced from Evans *et al.*[33] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 70 Demographic characteristics of interview participants by trial

TABLE 107 Demographic characteristics of interview participants by trial

Variable	Streamline L ($N = 25$)	Streamline C ($N = 26$)	Group differences
Sex			
Male, % (*n*)	60 (15)	61.5 (16)	$\chi^2 = 0.013$, df 1; $p = 0.910$
Female, % (*n*)	40 (10)	38.5 (10)	
Age (years), mean (SD)	65 (11)	64 (9)	$t = 0.261$, df 49; $p = 0.795$
Area deprivation score (quintile)			
1 (highest), % (*n*)	52.0 (13)	30.8 (8)	$\chi^2 = 5.459$, df = 2; $p = 0.065$
2, % (*n*)	32.0 (8)	23.1 (6)	
3–5 (mid to low), % (*n*)	16.0 (4)	46.2 (12)	
Interview method			
Face to Face, % (*n*)	20.0 (5)	26.9 (7)	$\chi^2 = 0.339$, df = 1; $p = 0.560$
Phone, % (*n*)	80.0 (20)	73.1 (19)	
Median interval between interview and WB-MRI, days (range)	15 (6–43)	20 (1–63)	$U = 270.00$; $p = 0.299$

Reproduced from Evans *et al.*[33] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 71 Comparison of demographic characteristics of participants sent both questionnaires, who were or were not included in the final analysis (study 2)

TABLE 108 Comparison of demographic characteristics of participants sent both questionnaires, who were or were not included in the final analysis (study 2)

Variable	Sent both questionnaires		Group difference
	Excluded from analysis ($n = 99$)	Included in final analysis ($n = 115$)	
Age (years), mean (range)	65.0 (30–86)	66.3 (31–89)	Mann–Whitney *U*-test; $p = 0.676$
Sex, % (*n*)			
Male	37.7 (29)	51.1 (70)	$\chi^2 = 3.578$, df 1; $p = 0.059$
Female	62.3 (48)	48.9 (67)	
Cancer type, % (*n*)			
Colorectal	48.3 (57)	43.8 (42)	$\chi^2 = 0.442$, df 1; $p = 0.506$
Lung	56.3 (54)	51.7 (61)	
Deprivation quintile, % (*n*)			
1 (highest)	60.9 (42)	39.1 (27)	$\chi^2 = 10.370$, df 4; $p = 0.035$
2	45.1 (23)	54.9 (28)	
3	32.4 (12)	67.6 (25)	
4	41.2 (14)	58.8 (20)	
5 (lowest)	34.8 (8)	65.2 (15)	

Reproduced from Evans *et al.*[34] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 72 Demographic and characteristics of participants who completed the post-staging questionnaire

TABLE 109 Demographic and characteristics of participants who completed the post-staging questionnaire

Variable	Overall ($N = 115$)	Participant cohort		Differences between participant cohorts
		Streamline L ($N = 54$)	Streamline C ($N = 61$)	
Demographic characteristics				
Age (years), median (range)	66.3 (31–89)	69.7 (50–89)	64.2 (31–85)	Mann–Whitney *U*-test; $p = 0.010$
Male sex,[a] % (*n*)	58.3 (67)	55.6 (30)	60.7(37)	$\chi^2 = 0.306$, df 1; $p = 0.580$
White ethnicity,[b] % (*n*)	91.8 (90)	93.8 (45)	90.0 (45)	Fisher's exact; $p = 0.715$
IMD deprivation,[a] % (n)				
1 (highest)	23.5 (27)	25.9 (14)	21.3 (13)	$\chi^2 = 0.3875$, df 4; $p = 0.423$
2	24.3 (28)	27.8 (15)	21.3 (13)	
3	21.7 (25)	24.1 (13)	19.7 (12)	
4	17.4 (20)	14.8 (8)	19.7 (12)	
5 (lowest)	13.0 (15)	7.4 (4)	18.0 (11)	
Physical and emotional well-being, % (n)				
Comorbidity (at least one comorbid illness reported)[b]	53.4 (55)	66.7 (34)	40.4 (21)	$\chi^2 = 7.147$, df 1; $p = 0.008$
Emotional distress (GHQ-12 score of ≥ 4)[b]	41.6 (42)	47.1 (24)	36.0 (18)	$\chi^2 = 1.271$, df 1; $p = 0.260$

a No missing data.
b Missing data > 5%.

Reproduced from Evans *et al.*[34] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 73 Comparative experience of WB-MRI vs. CT/PET-CT: scan information, communication and facilities

TABLE 110 Comparative experience of WB-MRI vs. CT/PET-CT: scan information, communication and facilities

Level of satisfaction	Overall, % (*n*)	Streamline L, % (*n*)	Streamline C, % (*n*)	Group differences, *p*-value
Satisfied with information received before scan[a]				
WB-MRI[b]				0.169,[c] 0.071[d]
Very satisfied	55.6 (60)	51.9 (27)[a]	58.9 (33)[b]	
Satisfied	37.0 (40)	40.4 (21)	33.9 (19)	
Dissatisfied	3.7 (4)	5.8 (3)	1.8 (1)	
Very dissatisfied	3.7 (4)	1.9 (1)	5.4 (3)	
CT/PET-CT[b]				
Very satisfied	57.5 (61)	49.0 (25)[a]	65.5 (36)[b]	
Satisfied	34.9 (37)	37.3 (19)	32.7 (18)	
Dissatisfied	0.9 (1)	2.0 (1)	0 (0)	
Very dissatisfied	6.6 (7)	11.8 (6)	1.8 (1)	
Satisfied with communication during scan[a]				
WB-MRI[b]				0.637,[c] 0.059[d]
Very satisfied	56.1 (60)	57.7 (30)[a]	54.5 (30)[b]	
Satisfied	39.3 (42)	34.6 (18)	43.6 (24)	
Dissatisfied	2.8 (3)	5.8 (3)	0 (0)	
Very dissatisfied	1.9 (2)	1.9 (1)	1.8 (1)	
CT/PET-CT[b]				
Very satisfied	64.2 (68)	62.7 (32)[a]	65.5 (36)[b]	
Satisfied	32.1 (34)	31.4 (16)	32.7 (18)	
Dissatisfied	1.9 (2)	3.9 (2)	0 (0)	
Very dissatisfied	1.9 (2)	2.0 (1)	1.8 (1)	
Satisfaction with facilities[a]				
WB-MRI[b]				0.225,[c] 0.480[d]
Very satisfied	45.8 (49)	49.0 (25)[a]	42.9 (24)[b]	
Satisfied	45.8 (49)	43.1 (22)	48.2 (27)	
Dissatisfied	4.7 (5)	2.0 (1)	7.1 (4)	
Very dissatisfied	3.7 (4)	5.9 (3)	1.8 (1)	

continued

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 110 Comparative experience of WB-MRI vs. CT/PET-CT: scan information, communication and facilities (*continued*)

Level of satisfaction	Overall, % (*n*)	Streamline L, % (*n*)	Streamline C, % (*n*)	Group differences, *p*-value
CT/PET-CT[b]				
Very satisfied	54.7 (58)	62.7 (32)[a]	47.3 (26)[b]	
Satisfied	38.7 (41)	33.3 (17)	43.6 (24)	
Dissatisfied	4.7 (5)	2.0 (1)	7.3 (4)	
Very dissatisfied	1.9 (2)	2.0 (1)	1.8 (1)	

a Streamline L, WB-MRI vs. PET-CT; Streamline C, WB-MRI vs. CT.
b Missing data > 5%; % is valid %.
c Comparison between WB-MRI and PET-CT (Streamline L).
d Comparison between WB-MRI and CT (Streamline C).

Reproduced from Evans *et al.*[34] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

Appendix 74 Demographic and psychological characteristics of the cohort completing the discrete choice experiment

TABLE 111 Demographic and psychological characteristics of the cohort completing the DCE

	All participants	Streamline L participants	Streamline C participants	Group differences (*p*-value)
Demographics				
Age (years),[a] mean (SD); *N*	64.7 (10.9); 138	66.0 (10.8); 72	63.2 (10.9); 66	0.122
Male sex,[a] *n* (%); *N*	84 (60.9); 138	42 (59.2); 72	41 (62.1); 66	0.723
White ethnicity,[b] *n* (%); *N*	112 (84.2); 133	59 (85.5); 69	53 (82.8); 64	0.670
Educational qualifications,[c] *N*	124	63	61	< 0.001
None, *n* (%)	35 (28.2)	28 (44.4)	7 (11.5)	
Below degree level, *n* (%)	39 (31.5)	17 (27.0)	22 (36.1)	
Degree level or equivalent, *n* (%)	50 (40.3)	18 (28.6)	32 (52.5)	
Home ownership (yes), *n* (%); *N*	82 (64.6); 127	34 (50.0); 68	48 (81.4); 59	< 0.001
Car ownership (yes), *n* (%); *N*	114 (87.0); 131	61 (88.4); 69	53 (85.5); 62	0.619
Marital status,[b] *N*	134	72	62	0.403
Married/cohabiting, *n* (%)	85 (63.4)	42 (58.3)	43 (69.4)	
Single, *n* (%)	22 (16.4)	13 (18.1)	9 (14.5)	
Divorced/separated/widowed, *n* (%)	27 (20.1)	17 (23.6)	10 (16.1)	
Employment status,[b] *N*	135	71	64	0.018
Employed full-time/part-time/self-employed/ fulltime homemaker, *n* (%)	46 (34.1)	17 (23.9)	29 (45.3)	
Retired, *n* (%)	74 (54.8)	43 (60.6)	31 (48.4)	
Unemployed, disabled or too ill to work, *n* (%)	15 (11.1)	11 (15.5)	4 (6.3)	
Health				
Self-rated health,[b] *N*	136	71	65	0.001
Very bad/bad/fair, *n* (%)	60 (44.1)	41 (57.7)	19 (29.2)	
Good/very good, *n* (%)	76 (55.9)	30 (42.3)	46 (70.8)	
Presence of comorbidities,[a] *n* (%); *N*	77 (55.8); 138	48 (66.7); 72	29 (43.9); 66	0.007
Psychological variables				
Negative mood,[b] *n* (%); *N*	18.01 (7.45); 136	18.76 (7.62); 71	17.20 (7.23); 65	0.224
Positive mood,[b] *n* (%); *N*	27.32 (7.92); 137	25.73 (7.72); 71	29.03 (7.84); 66	0.014

a No missing data.
b Missing data < 5%.
c Missing data > 5%.

Reproduced from Miles *et al.*[35] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 75 Results of conditional logit regression analysis: Streamline L participants stratified by home ownership

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 112 Results of conditional logit regression analysis: Streamline L participants stratified by home ownership

Attributes	Levels	All participants, coefficient (95% CI)	Home ownership, coefficient (95% CI)		*p*-value[b]
			Yes	No	
Time in scanner	Minutes	–0.008 (–0.014 to –0.002)	–0.010 (–0.020 to 0.0003)	–0.011 (–0.023 to 0.002)	0.89
Time to diagnosis	Weeks	–0.372 (–0.449 to –0.295)	–0.469 (–0.603 to –0.334)	–0.391 (–0.539 to –0.244)	0.45
Radiation dose	Increased risk of cancer (/1000)	–0.413 (–0.551 to –0.274)	–0.582 (–0.809 to –0.354)	–0.391 (–0.626 to –0.156)	0.25
Number of additional scans	Number	–0.179 (–0.330 to –0.028)[a]	–0.063 (–0.191 to 0.318)[a]	–0.354 (–0.654 to –0.054)[a]	0.03
Accuracy	Percentage	0.109 (0.079 to 0.138)[a]	0.181 (0.129 to 0.233)[a]	0.077 (0.035 to 0.119)[a]	0.01
Need for whole body and head to be in scanner	No	–	–	–	–
	Yes	0.017 (–0.190 to 0.224)	0.026 (–0.316 to 0.367)	0.146 (–0.201 to 0.493)	0.63
Observations/respondents		1230/72	582/34	576/33	0.02

a Coefficient not significantly different from zero; all other coefficients significant at p values < 0.05.
b p-values are from chi-squared tests that coefficients are equal for the two subgroups; p-values < 0.05 indicate coefficients that are significantly different between groups the; p-value in the bottom row is for a joint test across all coefficients.
Different attributes do not have the same unit of change so cannot be directly compared with one another.
Reproduced from Miles *et al.*[35] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

Appendix 76 Results of conditional logit regression analysis: Streamline L participants stratified by education

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 113 Results of conditional logit regression analysis: Streamline L participants stratified by education

Attributes	Levels	All participants, coefficient (95% CI)	Education, coefficient (95% CI)		*p*-value[b]
			Degree level or higher	No qualifications or qualification below degree level	
Time in scanner	Minutes	–0.008 (–0.014 to –0.002)	–0.002 (–0.014 to 0.011)	–0.011 (–0.020 to –0.003)	0.24
Time to diagnosis	Weeks	–0.372 (–0.449 to –0.295)	–0.369 (–0.541 to –0.198)	–0.432 (–0.540 to –0.324)	0.56
Radiation dose	Increased risk of cancer (/1000)	–0.413 (–0.551 to –0.274)	–0.522 (–0.834 to –0.210)	–0.356 (–0.532 to –0.180)	0.36
Number of additional scans	Number	–0.179 (–0.330 to –0.028)	–0.041 (–0.282 to 0.364)	–0.271 (–0.479 to –0.064)	0.11
Accuracy	Percentage	0.109 (0.079 to 0.138)[a]	0.190 (0.115 to 0.264)[a]	0.092 (0.054 to 0.130)[a]	0.01
Need for whole body and head to be in scanner	No	–	–	–	–
	Yes	0.017 (–0.190 to 0.224)	–0.252 (–0.695 to 0.190)	0.208 (–0.070 to 0.486)	0.09
Observations/respondents		1230/72	310/19	776/44	0.04

a Coefficient not significantly different from zero; all other coefficients significant at $p < 0.05$.
b *p*-values are from chi-squared tests that coefficients are equal for the two subgroups; $p < 0.05$ indicate coefficients are significantly different between groups; *p*-value in bottom row is for joint test across all coefficients.
Different attributes do not have the same unit of change so cannot be directly compared with one another.
Reproduced from Miles *et al.*[35] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

Appendix 77 Results of conditional logit regression analysis: Streamline L participants stratified by self-rated health

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 114 Results of conditional logit regression analysis: Streamline L participants stratified by self-rated health

Attributes	Levels	All participants, coefficient (95% CI)	Self-rated health, coefficient (95% CI)		*p*-value[b]
			Good or very good	Very poor, poor or fair	
Time in scanner	Minutes	–0.008 (–0.014 to –0.002)	–0.010 (–0.022 to 0.002)	–0.009 (–0.016 to –0.001)	0.81
Time to diagnosis	Weeks	–0.372 (–0.449 to –0.295)	–0.455 (–0.607 to –0.304)	–0.342 (–0.445 to –0.238)	0.22
Radiation dose	Increased risk of cancer (/1000)	–0.413 (–0.551 to –0.274)[a]	–0.119 (–0.348 to 0.109)[a]	–0.613 (–0.810 to –0.415)[a]	0.01
Number of additional scans	Number	–0.179 (–0.330 to –0.028)	–0.332 (–0.619 to –0.044)	–0.095 (–0.291 to 0.100)	0.18
Accuracy	Percentage	0.109 (0.079 to 0.138)[a]	0.151 (0.103 to 0.200)[a]	0.080 (0.043 to 0.116)[a]	0.02
Need for whole body and head to be in scanner	No	–	–	–	–
	Yes	0.017 (–0.190 to 0.224)	0.184 (–0.204 to 0.572)	–0.055 (–0.323 to 0.212)	0.32
Observations/respondents		1230/72	534/29	678/37	0.01

a Coefficient not significantly different from zero; all other coefficients significant at $p < 0.05$.
b *p*-values are from chi-squared tests that coefficients are equal for the two subgroups; $p < 0.05$ indicate coefficients are significantly different between groups; *p*-value in bottom row is for joint test across all coefficients.
Different attributes do not have the same unit of change so cannot be directly compared with one another.
Reproduced from Miles *et al.*[35] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

Appendix 78 Results of conditional logit regression analysis: Streamline L participants stratified by test preference

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

TABLE 115 Results of conditional logit regression analysis: Streamline L participants stratified by test preference

Attributes	Levels	All participants, coefficient (95% CI)	Test preference, coefficient (95% CI)		*p*-value[b]
			WB-MRI	Standard tests	
Time in scanner	Minutes	–0.008 (–0.014 to –0.002)[a]	0.008 (–0.008 to 0.024)[a]	–0.014 (–0.025 to –0.003)[a]	0.03
Time to diagnosis	Weeks	–0.372 (–0.449 to –0.295)	–0.340 (–0.523 to –0.157)	–0.262 (–0.409 to –0.115)	0.51
Radiation dose	Increased risk of cancer (/1000)	–0.413 (–0.551 to –0.274)	–0.294 (–0.809 to 0.221)	–0.251 (–0.692 to 0.190)	0.90
Number of additional scans	Number	–0.179 (–0.330 to –0.028)	–0.512 (–1.029 to 0.004)	–0.386 (–0.813 to 0.040)	0.71
Accuracy	Percentage	0.109 (0.079 to 0.138)	0.092 (–0.020 to 0.204)	–0.008 (–0.116 to 0.100)	0.21
Need for whole body and head to be in scanner	No	–	–	–	–
	Yes	0.017 (–0.190 to 0.224)[a]	0.359 (–0.152 to 0.870)[a]	–0.722 (–1.269 to –0.176)[a]	0.01
Observations/respondents		1230/72	234/15	292/18	0.01

a Coefficient not significantly different from zero; all other coefficients significant at $p < 0.05$.
b *p*-values are from chi-squared tests that coefficients are equal for the two subgroups; $p < 0.05$ indicate coefficients are significantly different between groups; *p*-value in bottom row is for joint test across all coefficients.
Different attributes do not have the same unit of change so cannot be directly compared with one another.
Reproduced from Miles *et al.*[35] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

Appendix 79 Participants unwilling to trade attributes off against one another

TABLE 116 Participants unwilling to trade attributes off against one another

Attribute	Participants, *n* (%)		
	All	Streamline L	Streamline C
Accuracy: non-traders are those who always prefer more accurate tests	18 (13)	7 (10)	11 (17)
Time to diagnosis: non-traders are those who always prefer shorter time to diagnosis	16 (12)	7 (10)	9 (14)
Radiation: non-traders are those who always prefer less radiation	11 (8)	7 (10)	4 (6)
Scan time: non-traders are those who always prefer shorter scan time	2 (1)	1 (1)	1 (2)
Scan number: non-traders are those who always prefer fewer scans	2 (1)	0 (0)	2 (3)
Enclosed: non-traders are those who always prefer non-enclosure of head and body	2 (1)	1 (1)	1 (2)
All	51 (37)	23 (32)	28 (42)

Reproduced from Miles *et al.*[35] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/. Includes minor additions and formatting changes to the original text.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

Appendix 80 WB-MRI interobserver variability (not considering agreement with the consensus reference standard): Streamline C and Streamline L

TABLE 117 WB-MRI interobserver variability (not considering agreement with the consensus reference standard): Streamline C and Streamline L

Outcome	Disease status	Total (*n*)	Agree (*n*)	Disagree (*n*)
***Streamline C* (N = *40*)**				
Primary[a]	Metastatic disease present	25	22	3
	Metastatic disease absent	15	13	2
	Overall	40	35	5
Equivocal lesions considered positive	Metastatic disease present	25	21	4
	Metastatic disease absent	15	11	4
	Overall	40	32	8
Equivocal lesions considered negative	Metastatic disease present	25	22	3
	Metastatic disease absent	15	13	2
	Overall	40	35	5
***Streamline L* (N = *43*)**				
Primary[a]	Metastatic disease present	27	17	10
	Metastatic disease absent	16	11	5
	Overall	43	28	15
Equivocal lesions considered negative	Metastatic disease present	27	21	6
	Metastatic disease absent	16	11	5
	Overall	43	32	11

a Per-participant sensitivity for metastatic disease. Equivocal results considered positive for colonic tumours and negative for rectal tumours (Streamline C). Equivocal results considered positive (Streamline L).

© Queen's Printer and Controller of HMSO 2019. This work was produced by Taylor *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.

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