



Health Technology Assessment

Volume 29 • Issue 8 • April 2025

ISSN 2046-4924

Evaluation of prognostic models to improve prediction of metastasis in patients following potentially curative treatment for primary colorectal cancer: the PROSPECT trial

Vicky Goh, Susan Mallett, Manuel Rodriguez-Justo, Victor Boulter, Rob Glynne-Jones, Saif Khan, Sarah Lessels, Dominic Patel, Davide Prezzi, Stuart Taylor, Steve Halligan on behalf of all the PROSPECT investigators





Extended Research Article

Evaluation of prognostic models to improve prediction of metastasis in patients following potentially curative treatment for primary colorectal cancer: the PROSPECT trial

Vicky Goh^{1*}, Susan Mallett², Manuel Rodriguez-Justo³, Victor Boulter⁴,
Rob Glynne-Jones⁵, Saif Khan⁶, Sarah Lessels⁷, Dominic Patel³, Davide Prezzi¹,
Stuart Taylor², Steve Halligan² on behalf of all the PROSPECT investigators

¹School of Biomedical Engineering and Imaging Sciences, King's College London, London, UK

²UCL Centre for Medical Imaging, London, UK

³Research Department of Pathology, UCL Cancer Institute, London, UK

⁴Patient Representative

⁵Mount Vernon Centre for Cancer Treatment, Northwood, UK

⁶Department of Pathology, London, UK

⁷Scottish Clinical Trials Research Unit (SCTRU), NHS National Services Scotland, Edinburgh, Scotland

*Corresponding author vicky.goh@kcl.ac.uk

Published April 2025

DOI: 10.3310/BTMT7049

This report should be referenced as follows:

Goh V, Mallett S, Rodriguez-Justo M, Boulter V, Glynne-Jones R, Khan S, *et al.* Evaluation of prognostic models to improve prediction of metastasis in patients following potentially curative treatment for primary colorectal cancer: the PROSPECT trial. *Health Technol Assess* 2025;29(8). <https://doi.org/10.3310/BTMT7049>

Abstract

Background: Despite apparently curative treatment, many patients with colorectal cancer develop subsequent metastatic disease. Current prognostic models are criticised because they are based on standard staging and omit novel biomarkers. Improved prognostication is an unmet need.

Objectives: To improve prognostication for colorectal cancer by developing a baseline multivariable model of standard clinicopathological predictors, and to then improve prediction via addition of promising novel imaging, genetic and immunohistochemical biomarkers.

Design: Prospective multicentre cohort.

Setting: Thirteen National Health Service hospitals.

Participants: Consecutive adult patients with colorectal cancer.

Interventions: Collection of prespecified standard clinicopathological variables and more novel imaging, genetic and immunohistochemical biomarkers, followed by 3-year follow-up to identify postoperative metastasis.

Main outcome: Best multivariable prognostic model including perfusion computed tomography compared with tumour/node staging. Secondary outcomes: Additive benefit of perfusion computed tomography and other biomarkers to best baseline model comprising standard clinicopathological predictors; measurement variability between local and central review; biological relationships between perfusion computed tomography and pathology variables.

Results: Between 2011 and 2016, 448 participants were recruited; 122 (27%) were withdrawn, leaving 326 (226 male, 100 female; mean \pm standard deviation 66 \pm 10.7 years); 183 (56%) had rectal cancer. Most cancers were locally advanced [\geq T3 stage, 227 (70%)]; 151 (46%) were node-positive (\geq N1 stage); 306 (94%) had surgery; 79 (24%) had neoadjuvant therapy. The resection margin was positive in 15 (5%); 93 (28%) had venous invasion; 125 (38%) had postoperative adjuvant chemotherapy; 81 (25%, 57 male) developed recurrent disease. Prediction of recurrent disease by the baseline clinicopathological time-to-event Weibull multivariable model (age, sex, tumour/node stage, tumour size and location, treatment, venous invasion) was superior to tumour/node staging: sensitivity: 0.57 (95% confidence interval 0.45 to 0.68), specificity 0.74 (95% confidence interval 0.68 to 0.79) versus sensitivity 0.56 (95% confidence interval 0.44 to 0.67), specificity 0.58 (95% confidence interval 0.51 to 0.64), respectively. Addition of perfusion computed tomography variables did not improve prediction significantly: c-statistic: 0.77 (95% confidence interval 0.71 to 0.83) versus 0.76 (95% confidence interval 0.70 to 0.82). Perfusion computed tomography parameters did not differ significantly between patients with and without recurrence (e.g. mean \pm standard deviation blood flow of 60.3 \pm 24.2 vs. 61.7 \pm 34.2 ml/minute/100 ml). Furthermore, baseline model prediction was not improved significantly by the addition of any novel genetic or immunohistochemical biomarkers. We observed variation between local and central computed tomography measurements but neither improved model prediction significantly. We found no clear association between perfusion computed tomography variables and any immunohistochemical measurement or genetic expression.

Limitations: The number of patients developing metastasis was lower than expected from historical data. Our findings should not be overinterpreted. While the baseline model was superior to tumour/node staging, any clinical utility needs definition in daily practice.

Conclusions: A prognostic model of standard clinicopathological variables outperformed tumour/node staging, but novel biomarkers did not improve prediction significantly. Biomarkers that appear promising in small single-centre studies may contribute nothing substantial to prognostication when evaluated rigorously.

Future work: It would be desirable for other researchers to externally evaluate the baseline model.

Trial registration: This trial is registered as ISRCTN95037515.

Funding: This award was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment programme (NIHR award ref: 09/22/49) and is published in full in *Health Technology Assessment*; Vol. 29, No. 8. See the NIHR Funding and Awards website for further award information.

Health Technology Assessment

ISSN 2046-4924 (Online)

Impact factor: 3.5

A list of Journals Library editors can be found on the [NIHR Journals Library website](#)

Launched in 1997, *Health Technology Assessment* (HTA) has an impact factor of 3.5 and is ranked 30th (out of 174 titles) in the 'Health Care Sciences & Services' category of the Clarivate 2022 Journal Citation Reports (Science Edition). It is also indexed by MEDLINE, CINAHL (EBSCO Information Services, Ipswich, MA, USA), EMBASE (Elsevier, Amsterdam, the Netherlands), NCBI Bookshelf, DOAJ, Europe PMC, the Cochrane Library (John Wiley & Sons, Inc., Hoboken, NJ, USA), INAHTA, the British Nursing Index (ProQuest LLC, Ann Arbor, MI, USA), Ulrichsweb™ (ProQuest LLC, Ann Arbor, MI, USA) and the Science Citation Index Expanded™ (Clarivate™, Philadelphia, PA, USA).

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@nhr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nhr.ac.uk/hta.

Criteria for inclusion in the *Health Technology Assessment* journal

Manuscripts 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 article

The research reported in this issue of the journal was funded by the HTA programme as award number 09/22/49. The contractual start date was in January 2011. The draft manuscript began editorial review in June 2022 and was accepted for publication in October 2022. 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' manuscript 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 article.

This article presents independent research funded by the National Institute for Health and Care 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, 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, the HTA programme or the Department of Health and Social Care.

This article was published based on current knowledge at the time and date of publication. NIHR is committed to being inclusive and will continually monitor best practice and guidance in relation to terminology and language to ensure that we remain relevant to our stakeholders.

Copyright © 2025 Goh *et al.* This work was produced by Goh *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: <https://creativecommons.org/licenses/by/4.0/>. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Published by the NIHR Journals Library (www.journalslibrary.nhr.ac.uk), produced by Newgen Digitalworks Pvt Ltd, Chennai, India (www.newgen.co).

Contents

List of tables	vii
List of figures	x
List of abbreviations	xii
Plain language summary	xiii
Scientific summary	xiv
Chapter 1 Introduction	1
Background	1
Treatment	1
Staging and prognostic models	1
Molecular biomarkers	2
Imaging	3
Perfusion computed tomography	3
Perfusion computed tomography in colorectal cancer	3
Objectives of the PROSPECT study	4
<i>Primary objective</i>	4
<i>Secondary objectives</i>	4
Chapter 2 Methods	5
Trial design	5
Patient and public involvement	5
Recruitment sites	5
Recruitment	6
Inclusion criteria	6
Exclusion criteria	6
Interventions	7
<i>Perfusion computer tomography imaging</i>	7
<i>Perfusion computed tomography analysis</i>	7
<i>Pathology</i>	8
Quality control of perfusion computed tomography	9
<i>Iodine phantom imaging</i>	9
<i>Clinical image quality</i>	9
Standard staging investigations and planned treatments	10
Standard pathological evaluation	10
Data collection and follow-up	10
Outcomes	10
Sample size estimation	10
Definition of end points for disease-free survival outcome	11
Analysis	11
<i>Summary</i>	11
<i>Study aim</i>	12
<i>Secondary outcomes</i>	16
Chapter 3 Results	17
Participants	17
Staging	17

Perfusion computed tomography	17
<i>Local site review</i>	17
Central review	17
Pathology	20
<i>Pathological evaluation</i>	20
<i>Immunohistochemical and somatic mutation analysis</i>	20
<i>Central review</i>	21
Quality control	21
<i>Iodine phantom imaging</i>	21
<i>Perfusion computed tomography image quality</i>	21
Follow-up and recurrences	22
Carcinoembryonic antigen levels during follow-up	22
Chapter 4 Prognostic modelling	23
Introduction	23
Methods	23
<i>Participants</i>	23
<i>Model development and assessment</i>	23
<i>Outcomes</i>	23
Results	24
<i>Current practice (American Joint Committee on Cancer staging: rule C)</i>	24
Model A (standard candidate variables)	24
<i>Model B (model A plus local perfusion computed tomography variables)</i>	26
<i>Model D (model A plus simplest local perfusion computed tomography variable)</i>	28
<i>Model E (model A plus central perfusion computed tomography variables)</i>	30
<i>Model F (model A plus additional pathology variables)</i>	32
<i>Sensitivities and specificities</i>	34
<i>Summary of model performance</i>	36
<i>Model A equation</i>	36
Chapter 5 Perfusion computed tomography: local versus central review	40
Introduction	40
Methods	40
<i>Participants</i>	40
<i>Perfusion computed tomography imaging and analysis</i>	40
<i>Statistical analysis</i>	41
Results	41
Chapter 6 Associations between perfusion computed tomography and pathological variables	44
Introduction	44
Methods	44
<i>Participants</i>	44
<i>Perfusion computed tomography imaging and analysis</i>	44
<i>Pathological analysis</i>	45
<i>Statistical analysis</i>	45
Results	46
<i>CD105 microvessel density</i>	46
<i>Vascular endothelial growth factor expression</i>	46
<i>Hypoxia-inducible factor-1 expression</i>	46
<i>Glucose transporter protein-1 expression</i>	46
<i>Mismatch repair status</i>	46
<i>Venous invasion</i>	47
<i>Tumour regression grade</i>	47

CONTENTS

Chapter 7 Discussion	48
Chapter 8 Overall conclusions and implications for practice	51
Additional information	52
References	54
Appendix 1 Supplemental tables	60
Appendix 2 Supplemental figures	79

List of tables

TABLE 1	Summary of overall study aim, primary and secondary outcomes	11
TABLE 2	Analyses undertaken for primary and secondary outcomes	12
TABLE 3	Characteristics of final participant cohort	19
TABLE 4	Summary of TN staging, by location (colon or rectum)	20
TABLE 5	Local site perfusion CT measurements for participants with and without recurrence	20
TABLE 6	Summary of pathological characteristics	21
TABLE 7	Summary of events during 3-year follow-up period	22
TABLE 8	Performance of AJCC staging for predicting recurrence	26
TABLE 9	Univariable hazard ratios (HR) for model variables assessed	27
TABLE 10	Multivariable-adjusted hazard ratios for model variables assessed	28
TABLE 11	C-index including internal validation and optimism adjustment	30
TABLE 12	Univariable hazard ratios for local hospital perfusion CT variables	31
TABLE 13	Multivariable-adjusted hazard ratios for model B with model A for reference	31
TABLE 14	Univariable hazard ratios for additional pathology variables	33
TABLE 15	Multivariable-adjusted hazard ratios for all modelled pathology variables	33
TABLE 16	Sensitivity and specificity analysis for model B compared with rule C	35
TABLE 17	Sensitivity and specificity analysis for model A compared with rule C	35
TABLE 18	Difference in sensitivity and specificity for model A compared with rule C	37
TABLE 19	Difference in sensitivity and specificity for model F compared with model A	37
TABLE 20	Summary of model performance	38
TABLE 21	Union for International Cancer Control/AJCC staging classification for colorectal cancer (8th edition)	60
TABLE 22	List of participating sites and site principal investigators	61
TABLE 23	Recruitment sites and multidetector CT capability	62
TABLE 24	Perfusion CT models and vascular parameters	62

TABLE 25	Definitions for end of period for disease-free survival	63
TABLE 26	Definitions employed for high-risk (stage III) disease	63
TABLE 27	Summary of treatments that participants received	63
TABLE 28	Local site perfusion CT measurements by tumour T stage	64
TABLE 29	Local site perfusion CT measurements for node-positive and node-negative cancers	65
TABLE 30	Central review of perfusion CT measurements by tumour T stage	65
TABLE 31	Central review of perfusion CT measurements for node-positive and node-negative cancers	65
TABLE 32	Central review of perfusion CT measurements for participants with and without recurrence	66
TABLE 33	CD105 expression (number of CD105 stained vessels/mm ² field) in participants with and without recurrence	66
TABLE 34	Immunohistochemical scores for HIF-1 α , VEGF and GLUT-1 in participants with and without recurrence	66
TABLE 35	Frequency of genetic mutations (MMR, KRAS, HRAS, NRAS and BRAF) for participants with and without recurrence	67
TABLE 36	Mean (SD) image noise obtained from the bladder ROI per participating centre and CT scanner model	68
TABLE 37	Summary of peak arterial iodine enhancement and full-width half maximum of the initial arterial enhancement peak	68
TABLE 38	Summary of timing of follow-up visits	68
TABLE 39	Location of disease recurrence	69
TABLE 40	Summary of serum CEA levels at baseline and subsequent follow-up clinic visits	69
TABLE 41	Summary of candidate variables included in model B	70
TABLE 42	Correlation coefficients between local hospital perfusion CT variables	71
TABLE 43	Univariable hazard ratios for perfusion CT variables from central review	71
TABLE 44	Multivariable-adjusted hazard ratios for model E with model A as reference	72
TABLE 45	Summary of model variables of participants in model F	72
TABLE 46	Multivariable-adjusted hazard ratios for somatic mutation analysis variables	73
TABLE 47	Difference in sensitivity and specificity for model B compared with rule C	73
TABLE 48	Sensitivity and specificity analysis for model B (alternative threshold) compared with Rule C	74

TABLE 49 Difference in sensitivity and specificity for model B (alternative threshold) compared with rule C	74
TABLE 50 Sensitivity and specificity analysis for model A (alternative threshold) compared with rule C	75
TABLE 51 Difference in sensitivity and specificity for model A (alternative threshold) compared with rule C	75
TABLE 52 Sensitivity and specificity analysis for model A compared with model B	76
TABLE 53 Difference in sensitivity and specificity for model A compared with model B	76
TABLE 54 Sensitivity and specificity analysis for model D compared with model B	77
TABLE 55 Sensitivity and specificity analysis for model E compared with model A	77
TABLE 56 Difference in sensitivity and specificity for model E compared with model A	78
TABLE 57 Sensitivity and specificity analysis for model F compared with model A	78
TABLE 58 Mean difference and 95% limits of agreement for perfusion CT variables between local hospital and central review	78

List of figures

FIGURE 1 Overview schema of the PROSPECT trial	6
FIGURE 2 Schema of CT perfusion acquisition	7
FIGURE 3 Example of parametric maps derived from perfusion CT image analysis	8
FIGURE 4 Participant flow through the trial	18
FIGURE 5 Figure showing analysis models	24
FIGURE 6 Kaplan–Meier curve for high- and low-risk patients as defined by AJCC stage group	25
FIGURE 7 Scatter plot showing time to recurrence with respect to risk groupings defined by clinical stage	25
FIGURE 8 Box plots showing distribution with respect to time to recurrence (in years) for the following candidate variables: age (a); tumour size (b); tumour [T] stage (c); nodal [N] stage (d); sex (e); tumour location as left or right colon (f); venous invasion (g); and treatment groups (h)	26
FIGURE 9 Kaplan–Meier plots for the different risk groupings	29
FIGURE 10 Scatter plot showing time to recurrence with respect to two risk groupings defined by model A prediction index (low/medium risk vs. high risk)	29
FIGURE 11 Scatter plot showing time to recurrence with respect to two risk groupings defined by model A prediction index (low vs. medium/high risk)	30
FIGURE 12 Box and whisker plots of perfusion CT variables by recurrence group	31
FIGURE 13 Scatter plot of prediction index (PI) for model A and model B	32
FIGURE 14 Scatter plot of prediction index (PI) for model A and model F3 (all additional pathology variables)	34
FIGURE 15 Plot showing net benefit gain with different true-positive/false-positive weightings for rule C (AJCC staging) and model A (based on standard candidate variables)	37
FIGURE 16 Forest plot of performance for all models	39
FIGURE 17 Scatter plots showing the values from local and central review on the same plot	41
FIGURE 18 Bland–Altman plots showing the difference plotted against the average measurement for blood flow, blood volume, mean transit time and permeability surface area product, respectively	42
FIGURE 19 Difference in size of ROI placed plotted against the difference in each vascular parameter between local and central review	43
FIGURE 20 Scatter plots of blood flow, blood volume, mean transit time and permeability surface area product against CD105 microvessel density	46

FIGURE 21 Scatter plots of blood flow, blood volume, mean transit time and permeability surface area product against MMR status	47
FIGURE 22 Kaplan–Meier plots for the different risk groupings	79
FIGURE 23 Scatter plot showing time to recurrence with respect to two risk groupings defined by model B prediction index (PI), low/medium risk vs. high risk	80
FIGURE 24 Scatter plot showing recurrences and time to recurrence with respect to two risk groupings defined by the model B prediction index (PI), high/medium vs. low risk	80
FIGURE 25 Scatter plot of prediction index (PI) for model A and model D	81
FIGURE 26 Box and whisker plots showing the distribution of perfusion CT variables by recurrence group	82
FIGURE 27 Kaplan–Meier plot for low/medium- vs. high-risk groups	82
FIGURE 28 Scatter plot showing time to recurrence with respect to two risk groupings defined by model E prediction index (PI), low/medium risk vs. high risk	83
FIGURE 29 Scatter plot of prediction index (PI) for models A and E	83
FIGURE 30 Kaplan–Meier plots (all variables, model F3) for the different risk groupings	84
FIGURE 31 Kaplan–Meier plot (somatic mutation analysis variables, model F1) for the different risk groupings	85
FIGURE 32 Scatter plot prediction index for model A and model F1 (somatic mutation analysis variables)	86
FIGURE 33 Difference in size of ROI placed plotted against the difference in each vascular parameter between local and central review with the scanner manufacturers highlighted by different colours	86
FIGURE 34 Scatter plots of CT blood flow, blood volume, mean transit time and permeability surface area product against VEGF score (0–4)	87
FIGURE 35 Scatter plots of CT blood flow, blood volume, mean transit time and permeability surface area product against HIF-1 α score (0–6)	88
FIGURE 36 Scatter plots of CT blood flow, blood volume, mean transit time and permeability surface area product against GLUT-1 score (0–8)	89
FIGURE 37 Scatter plots of CT blood flow, blood volume, mean transit time and permeability surface area product against the presence or absence of venous invasion	90
FIGURE 38 Scatter plots of CT blood flow, blood volume, mean transit time and permeability surface area product against the presence or absence of tumour regression	91

List of abbreviations

AIC	Akaike information criterion	MDT	multidisciplinary team
AJCC	American Joint Committee on Cancer	MMR	mismatch repair
BIC	Bayesian information criterion	MRI	magnetic resonance imaging
BRAF	v-raf murine sarcoma viral oncogene homolog B1	NICE	National Institute for Health and Care Excellence
CI	confidence interval	NRAS	neuroblastoma RAS viral oncogene homolog
CEA	carcinoembryonic antigen	PCA	principal components analysis
CRF	case report form	PET	positron emission tomography
df	degrees of freedom	RAF	rapidly accelerated fibrosarcoma
EGFR	epidermal growth factor receptor	RAS	rat sarcoma virus
18F-FDG	18 fluorodeoxyglucose	ROI	region of interest
FFPE	formalin-fixed paraffin embedded	TN	tumour and nodal
GE	General Electric	TNM	tumour node metastasis
GLUT-1	glucose transporter protein	TRG	tumour regression grade
HIF-1 α	hypoxia-inducible factor 1-alpha	TRIPOD	Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis
HR	hazard ratio	VEGF	vascular endothelial growth factor
IDMC	Independent Data Monitoring Committee	UICC	Union for International Cancer Control
KRAS	Kirsten rat sarcoma viral oncogene homolog		

Plain language summary

Bowel cancer is one of the most common United Kingdom cancers and a leading cause of death. Despite apparently curative treatment, up to half of patients ultimately die from their disease because the tumour subsequently spreads around the body, known as 'metastasis'. Patients are given chemotherapy upfront to prevent this spread, but predicting who will and will not develop metastasis is challenging, so it is difficult to know who to treat. Prediction is based on cancer 'stage', which describes how advanced the tumour is on imaging and under the microscope. A 'multivariable prognostic model' may improve prediction and is a combination of multiple factors known about the patient and their tumour that provides a score for the chance of future disease. However, multivariable models are not commonly used to predict recurrence for colorectal cancer and are criticised because they omit the latest 'cutting-edge' measurements (e.g. from scanning and genetic testing). To improve prediction of outcomes after bowel cancer, we performed a study in 13 National Health Service hospitals, where we collected both basic and more novel measurements from patients at the time of their diagnosis. We then followed patients for 3 years to determine who did and did not develop metastasis. From 2011 to 2016, we recruited 448 patients and used data from 326 to develop a multivariable model to predict metastasis. Our baseline model used a combination of basic factors, such as age, sex, tumour size and location, and treatment. This model predicted future disease significantly better than simple measurement of tumour stage. However, we found that the model did not improve when we added cutting-edge measurements. This suggests that these newer measurements are not useful to predict the chance of future disease. Our results suggest that researchers investigating prediction would be best served by concentrating on basic rather than more novel measurements.

Scientific summary

Background

Colorectal cancer accounts for 12% of all new UK cancers, with over 42,000 new patients diagnosed each year. Despite treatment with curative intent, up to 50% of colorectal cancer patients will develop subsequent recurrent disease, normally metastasis. Chemotherapy aims to combat metastasis but identification of who will and will not develop subsequent metastasis (i.e. who does and does not merit chemotherapy) is difficult. Currently, 'at-risk' patients are identified by tumour and nodal (TN) staging from diagnosis and surgery (when performed) but more accurate prognostication remains an unmet need. Multivariable models promise to improve prediction by combining multiple weighted predictor factors measured from the patient in question but are not used widely. A frequent criticism is that such models ignore 'cutting-edge' promising biomarkers, which are currently the subject of intense research and which appear to offer an opportunity to improve risk stratification at diagnosis. Also, the move in recent years from offering chemotherapy in the postoperative (adjuvant) to preoperative (neoadjuvant) setting has shifted the need for identification of high-risk patients from the post-surgery setting (i.e. by using pathological samples from the resected specimen) to the preoperative setting (which depends on imaging and biopsy samples of the primary tumour).

Objectives

Our primary objective was to improve prediction of outcomes from colorectal cancer by developing a multivariable prognostic model of disease-free survival. We aimed to develop a best baseline model using standard clinicopathological variables and to then improve its prediction significantly by incorporating cutting-edge, novel imaging [perfusion computed tomography (CT)], immunohistochemical and genetic biomarkers. Our primary outcome was prediction of the baseline model incorporating CT perfusion when compared with standard TN staging. Secondary outcomes included baseline model prediction when incorporating immunohistochemical or genetic biomarkers; assessment of measurement variability between local sites and central review; and to investigate the biological relationships between perfusion CT and pathology variables.

Methods

We conducted a prospective multicentre cohort trial. Participants were recruited from 13 representative NHS teaching and district general hospitals in England and Scotland. Participants were eligible if they had histologically proven or suspected primary colorectal cancer (mass suspicious on endoscopy or imaging). Exclusions included polyp cancers, unequivocal metastases at staging, patients aged < 18 years, contraindications to intravenous contrast, pregnancy, and final diagnosis not being cancer. All participants gave written informed consent.

Consecutive, unselected patients underwent perfusion CT in addition to standard staging CT. Standard investigations were interpreted locally by the usual clinical care team. Perfusion CT measurements were obtained locally by 26 radiologists. Central review was performed by three radiologists, who were blinded to all standard staging investigations. Treatment decisions were undertaken by the local multidisciplinary team as per usual practice. In patients undergoing surgery, central pathological review of the resected tumour was undertaken by two pathologists who performed additional pathological analysis, including immunohistochemistry for angiogenesis and hypoxia; microsatellite instability, mismatch repair (MMR), and somatic mutation analysis – Kirsten rat sarcoma viral oncogene homolog, neuroblastoma RAS viral oncogene homolog, v-raf murine sarcoma viral oncogene homolog B1 (BRAF). Participants were followed for 3 years and patients with recurrence were identified.

A best baseline multivariable prognostic model was developed from prespecified standard clinical (age, sex, treatment) and pathological (tumour location, size, presence of venous invasion) variables. All model variables were prespecified, based on existing literature and expert opinion; that is, univariable significance in the study data set was not used

to select any variable, either standard or novel. For our primary outcome, perfusion CT variables were added to the standard model as a composite (principal components) score. Prediction of this model was then compared with standard TN staging. The additive benefit (if any) of CT perfusion variables to the baseline model was calculated. For secondary outcomes, the additive benefit (if any) of immunohistochemical markers of angiogenesis, hypoxia and somatic mutation analysis was also calculated.

We calculated the extent of any variation between local and central perfusion CT measurements. We calculated correlations between perfusion CT measurements and histopathological variables to determine biological significance. We estimated a sample size of 320 patients with 80 events (i.e. metastasis) would have 80% power to detect a 15% difference in correct risk classification by the model, allowing for loss to follow-up. We reported our results according to Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis guidelines.

Results

Between 2011 and 2016, we recruited 448 participants; 122 (27%) were withdrawn (mostly due to additional cancer), leaving 326 for analysis [226 male, 100 female; mean \pm standard deviation (SD) 66 \pm 10.7 years]; a total of 183 (56%) had rectal cancer. Most cancers were locally advanced [\geq T3 stage, 227 (70%); 151 (46%) were node-positive (\geq N1 stage)]. Surgery was performed in 306 (94%). The resection margin was positive in 15 (5%). Venous invasion was present in 93 (28%). Neoadjuvant therapy was undertaken in 79 (24%) and adjuvant therapy in 125 (38%) participants. Eighty-one (25%, 57 male) developed recurrent disease over the 3-year follow-up period.

Perfusion CT measurements were available from local sites in 303 (93%) participants. Perfusion CT parameters did not differ between patients with and without positive local nodes (e.g. mean \pm SD blood flow: 64.5 \pm 25.2 vs. 75.0 \pm 44.1 ml/minute/100 ml) or with and without recurrence (e.g. mean \pm SD blood flow: 60.3 \pm 24.2 vs. 61.7 \pm 34.2 ml/minute/100 ml). Central review was undertaken in 291 (96%). Variability assessed by Bland–Altman plots was considerable between many local and central review perfusion CT measurements, most evident for permeability surface area product, where disagreement was greatest at higher permeability values. Although there were differences regarding where the region of interest was placed when local and central reviews were compared, this was not a major contributor to disagreement for vascular parameter values. Similarly, the individual CT scanner manufacturer did not impact substantially on disagreement, because all common manufacturers displayed large differences over all vascular parameters.

There was no clear relationship between perfusion CT variables and immunohistochemical markers of angiogenesis (CD105, vascular endothelial growth factor) or hypoxia (hypoxia-inducible factor-1, glucose transporter-1) in the primary tumour, suggesting that CT does not reflect angiogenesis precisely. There was no difference between perfusion CT variables and MMR deficient/MMR proficient tumours.

Prediction for the baseline clinicopathological model improved over standard TN staging due to significantly improved specificity: sensitivity 0.57 [95% confidence interval (CI) 0.45 to 0.68] and specificity 0.74 (95% CI 0.68 to 0.79) versus sensitivity 0.56 (95% CI 0.44 to 0.67) and specificity 0.58 (95% CI 0.51 to 0.64), respectively. The addition of perfusion CT variables to the baseline clinicopathological model did not improve prediction significantly: *c*-statistic 0.77 (95% CI 0.71 to 0.83) versus 0.76 (95% CI 0.70 to 0.82), respectively. Similarly, the addition of more novel histopathological variables (i.e. markers of angiogenesis, hypoxia, rat sarcoma virus, BRAF and MMR mutation status) to the baseline clinicopathological model did not improve model prediction significantly: *c*-statistic: 0.78 (95% CI 0.72 to 0.84) versus 0.76 (95% CI 0.70 to 0.82), respectively.

Limitations

The number of exclusions/withdrawals was higher than anticipated, mostly due to a higher prevalence of metastasis at baseline (possibly due to additional scans and multiple readers interpreting them). While prediction using our best baseline clinicopathologic model was significantly better than current practice, it may still be suboptimal for adoption in

day-to-day clinical practice and its clinical utility needs assessment. External evaluation (validation) of the model in an NHS setting was not performed. The number of patients undergoing additional histopathological analysis was relatively small, as the study was not specifically powered to detect an effect for these variables, but if a beneficial effect on prediction exists, it is likely to be small.

Conclusions

We developed a prognostic model to predict development of metastatic disease following apparently curative treatment for colorectal cancer. The best baseline model comprising prospectively collected prespecified clinicopathological variables improved over standard TN staging prediction significantly. However, the addition of perfusion CT, immunohistochemical or genetic variables was not able to improve prediction significantly.

Implications for health care

In the NHS setting, applying a prognostic model comprising standard clinicopathological variables achieves significantly greater specificity for predicting subsequent metastasis than does current TN staging, without any diminished sensitivity. If similar prediction is sustained in external validation, application of this model in clinical practice may have immediate beneficial implications for the care of patients presenting with apparently localised colorectal cancer.

Recommendations for future research

1. Model prediction should be externally evaluated in an NHS setting, preferably by authors unrelated to model development.
2. In addition to an external evaluation of its predictive accuracy, an evaluation should be made of the clinical utility to clinicians of the model in an NHS setting, including within neoadjuvant chemotherapy trials.
3. Venous invasion on pathological evaluation was a strong prognostic factor within the standard model; further research into preoperative imaging assessment of venous invasion on CT for colon cancer and magnetic resonance imaging for rectal cancer is warranted.
4. The fact that CT, immunohistochemistry and genetic markers of angiogenesis did not improve model prediction suggests that prior small, single-centre, retrospective studies including a benefit to these biomarkers are overoptimistic. This finding should be considered when contemplating funding future studies of such markers. Rather, our data suggest that future prognostic research should focus on standard clinicopathological variables.

Changes to protocol

1. Following interim presentation of trial data, the data monitoring committee increased recruitment from 370 to 448 patients, driven by a higher-than-expected baseline dropout rate due to metastasis at staging. The trial then continued recruitment until the original target of 80 evaluable participants with an event was achieved.
2. Intended model analysis adjusted by clustering by study site was removed due to methodological advances in the interim, showing that this adjustment can cause statistical model instability.
3. The proposed discrete choice study was not performed, so results for secondary outcome 7 were expressed as number of true-positive and false-positive patients, without a combined net benefit outcome (that would combine these metrics into a single measure).

Trial registration

This trial is registered as ISRCTN95037515.

Funding

This award was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment programme (NIHR award ref: 09/22/49) and is published in full in *Health Technology Assessment*; Vol. 29, No. 8. See the NIHR Funding and Awards website for further award information.

Chapter 1 Introduction

Background

Colorectal cancer is one of the most common cancers worldwide, with 1.8 million new patients diagnosed annually (10.2% of all new cancers). It is also a leading cause of cancer-related death, resulting in 881,000 deaths annually.¹ In the UK, the incidence of colorectal cancer has decreased by 6% between 2009 and 2019. Nevertheless, colorectal cancer still accounts for 12% of all new UK cancers, with more than 42,000 new patient diagnoses and 16,000 deaths per year.² Cancers of the sigmoid colon, rectosigmoid junction and rectum account for over half of diagnoses, that is 'left-sided' cancers.

Treatment

Surgery remains the mainstay of curative treatment. Refinements in surgical technique, notably total mesorectal excision,^{3,4} and the introduction of neoadjuvant radiotherapy with or without chemotherapy has decreased resection margin positivity and local recurrence rates for rectal cancer.⁵⁻⁷ Adjuvant chemotherapy for patients with nodal involvement has also lowered the risk of recurrence and death.⁸⁻¹¹ Adjuvant chemotherapy for node-negative cancers remains an individual decision, given the minimal benefit found by several trials for non-risk stratified stage II colon cancer in terms of both disease-free and overall survival.¹²⁻¹⁷ Currently, stage II cancers with: (1) bowel perforation or bowel obstruction, (2) pathological T4 stage, (3) lymphovascular or perineural invasion, or (4) < 12 lymph nodes examined by a pathologist would prompt a discussion regarding adjuvant chemotherapy. Nevertheless, despite treatment advances, cancer will recur in up to 50% of patients, who will die from their disease ultimately, predominantly due to distant metastases, most commonly liver and lung.¹⁸ Once metastasis is established, 5-year survival is around 13%.¹⁹

Staging and prognostic models

Accurate staging is required to optimise clinical management. The Union for International Cancer Control (UICC) tumour–node–metastasis (TNM) staging classification and corresponding American Joint Committee on Cancer (AJCC) stage grouping is the standard classification system used worldwide,²⁰ where 'T' represents the primary tumour characteristics (namely local spread through the bowel wall), 'N' represents the presence (and degree) or absence of nodal involvement and 'M' represents the presence or absence of distant metastases.

Staging also informs prognosis, although it has limitations because patients assigned the same stage can experience very different outcomes. For example, the 5-year survival rate varies from 63% to 87% for stage II cancers, which have spread beyond the bowel wall but not spread to nearby lymph nodes. This may reflect the fact that stage II amalgamates node-negative primary cancers that have not spread beyond the bowel wall extensively (T3) with more advanced node-negative primary cancers that invade adjacent organs (T4). Indeed, data from the Surveillance, Epidemiology, and End Results program and other studies have indicated that the survival of patients with stage IIB/IIC may be worse than those with stage IIIA,^{21,22} which may reflect an adverse biology associated with T4N0 cancers versus T1 or T2N1 cancers.

Tumour–node–metastasis revisions have aimed to improve prognostic performance and risk stratification for treatment. Currently, the TNM classification is in its eighth edition, introduced in 2020 (see [Appendix 1, Table 21](#)). More accurate identification of higher-risk patients would mean earlier intervention could be targeted more precisely, and remains an unmet clinical need.^{13,23} One approach to improving prognosis is via multivariable prognostic models. These models combine multiple factors to estimate the risk of a future outcome(s), with the aim of improving prediction by incorporating more individualised information than that available from simple TNM staging. However, while

models predicting recurrence from colorectal cancer are available, they are not widely used.²⁴ A major criticism is that they ignore 'novel' predictors in the face of extensive recent and topical research around imaging, genetic and immunohistochemical biomarkers. For example, recent years have seen the introduction of novel chemotherapeutic agents, including bevacizumab, targeted at vascular endothelial growth factor (VEGF),²⁵ cetuximab^{26,27} and panitumumab,^{28,29} targeted at epidermal growth factor receptor (EGFR), and more recently, programmed cell death protein 1 inhibitors, which are active against mismatch repair (MMR) deficient cancers,^{30,31} providing additional therapeutic options. Prognostic models that incorporate markers of angiogenesis and/or molecular characteristics may improve prediction. However, while novel biomarkers promise to improve prediction and 'personalise' treatment, their evaluation is hindered by small retrospective studies.

Molecular biomarkers

Currently, a number of molecular biomarkers are thought to be prognostic and/or predictive with respect to systemic anti-cancer therapy.^{32,33} We note here that oncologists use the term 'prognosis' to refer to outcomes independent of treatment and 'prediction' to refer to outcomes after treatment. In contrast, statisticians and methodologists avoid specific nomenclature for healthcare models of treatment outcomes because they consider all participants receive 'treatment' of some kind, even if that involves no active intervention.

The EGFR gene is frequently amplified in colorectal cancer and overexpressed at RNA and protein levels in most tumours.³⁴ Metastatic cancers with activating mutations of Kirsten rat sarcoma viral oncogene homolog (KRAS) affecting exon 2 codons 12 and 13 will not benefit from anti-EGFR monoclonal antibody therapy compared with wild-type KRAS tumours.³⁵⁻³⁷ For patients being considered for anti-EGFR therapy with or without chemotherapy, rat sarcoma virus (RAS) mutational testing, including KRAS and neuroblastoma RAS viral oncogene homologue (NRAS) codons 12, 13 of exon 2; 59, 61 of exon 3; and 117 and 146 of exon 4 is now performed. Downstream activating mutations in the EGFR signalling pathway [e.g. in the RAS–rapidly accelerated fibrosarcoma (RAF)–mitogen-activated protein kinase and phosphatidylinositol 3-kinase (PI3K) pathways] may also have a negative effect on response to anti-EGFR monoclonal antibody therapy. KRAS, NRAS and v-raf murine sarcoma viral oncogene homologue B1 (BRAF) mutations occur in more than half of colorectal cancers; KRAS or NRAS and BRAF mutations are inversely associated, though a small proportion of individual colorectal carcinomas show co-occurrence of RAS and RAF mutations. In one series, KRAS wild-type carriers of BRAF, NRAS and PIK3CA exon 20 mutations had a lower response rate to cetuximab than the wild-type population (36% vs. 41%).³⁸

Evidence from one observational study ($N = 783$) showed that in patients with stage II or III colorectal cancer treated with adjuvant chemotherapy, with or without anti-EGFR targeted therapy, those with KRAS mutations had poorer disease-free survival than patients with wild-type KRAS.

Microsatellite instability testing or immunohistochemistry testing for MMR proteins can identify people in whom colorectal cancer may have occurred because of Lynch syndrome.³⁹ Microsatellites are repetitive sequences of DNA that are at increased risk of copying errors during replication. Without an effective DNA MMR system, errors in copying microsatellite sequences cause them to vary in length. MMR proteins detected by immunohistochemistry testing are MLH1, MSH2, MSH6 and PMS2. Absent or reduced nuclear staining of one or more MMR proteins suggests that there may be a pathogenic mutation in a gene encoding these proteins. BRAF V600E mutational analysis should also be performed in deficient MMR tumours with loss of MLH1. Presence of a BRAF mutation strongly favours a sporadic pathogenesis versus Lynch syndrome.

Nevertheless, the move in recent years from offering chemotherapy in the postoperative (adjuvant) to preoperative (neoadjuvant) setting⁴⁰⁻⁴² has shifted the need for identification of high-risk patients from the post-surgery setting (i.e. by using pathological samples from the resected specimen) to the preoperative setting (which depends on imaging and biopsy samples of the primary tumour).⁴³

Imaging

The National Institute for Health Care and Excellence (NICE) clinical guidelines (CG131, updated 2014;⁴⁴ CG151, published 2020, updated 2021⁴⁵) recommend contrast-enhanced computed tomography (CT) of the thorax, abdomen and pelvis, with additional magnetic resonance imaging (MRI) of the pelvis for local staging when the primary tumour is a rectal cancer. Although not recommended by NICE, integrated positron emission tomography (PET)/CT with fluorine 18 fluorodeoxyglucose (¹⁸F-FDG) as the tracer may be performed additionally for patients where there is a high suspicion of distant metastases,^{46,47} or where there is oligometastatic disease and a curative strategy is being pursued.

Perfusion computed tomography

Perfusion CT is a simple addition to standard staging CT and is a surrogate marker of tumour perfusion, angiogenesis and hypoxia.⁴⁸⁻⁵⁰ By measuring regional blood flow, blood volume and vascular leakage rate, it complements the anatomical information provided by conventional CT.⁵¹⁻⁵³

Perfusion computed tomography in colorectal cancer

We searched PubMed and EMBASE for articles published between 1 January 1990 and 31 December 2020 without language restriction. We used medical subject headings and a full-text search for 'colorectal neoplasms', 'colorectal cancers', 'prognosis', 'X-ray computed tomography', 'computed tomography', and 'perfusion computed tomography', 'biomarkers, tumor/genetics'. We found limited data regarding the prognostic value of perfusion CT for colorectal cancer, with no systematic reviews or meta-analyses, and no data incorporating perfusion CT variables in prognostic models of recurrence.

Our retrospective, single-centre study of colorectal cancer incorporated perfusion CT and found a significant association between tumour perfusion and development of subsequent metastases;⁵⁴ this study was performed by some of the authors of the present work. These data appeared to show that poorly perfused colorectal cancers were more likely to metastasise following apparently curative surgery. We surmised that this may be related to tumour hypoxia.⁵⁵⁻⁵⁷ We elected to study perfusion CT as a novel imaging biomarker because it is believed to reflect tumour angiogenesis, a subject of intense research. Moreover, perfusion CT data can be acquired using conventional scanners, rendering it pragmatic for multicentre research and implementation if clinically useful.

Hayano *et al.*⁵⁵ published a review in 2014 highlighting the need for large multicentre studies, after presentation of only 6 identified studies including between 6 and 32 patients on the relationship between CT perfusion parameters and prognosis in colorectal cancer.

Hayano *et al.*⁵⁸ also reported a small study on 31 patients (11 with metastatic events) claiming to show a relationship but again highlighting the need for larger studies for a definitive investigation of the imaging parameters in prognosis.

More recently, a prospective observational study assessing vascular-metabolic imaging using ¹⁸F-FDG PET/CT perfusion has been published for colorectal cancer. Imaging was successful in 286 participants (184 males, mean age 70 ± 10 years; 84 deaths). Authors noted that a vascular-metabolic signature (high total lesion glycolysis or metabolic tumour volume and increased permeability surface area product/blood flow) was associated with poorer survival ($n = 61$ patients, number of deaths not reported) for both colon and rectal cancers with a hazard ratio of metabolic tumour volume 1.01 (95% CI: 1.04 to 1.02), whereas the hazard ratio from permeability surface area product/blood flow was based on such a small sample size that the hazard ratio cannot be estimated, as the 95% confidence interval (CI) ranged from 0.2 to 1000.⁵⁹

A multivariable prognostic model incorporating imaging and/or pathology prognostic biomarker variables (including histopathology and gene expression) may improve risk stratification for primary colorectal cancer, compared with current clinical practice. Because prior prognostic biomarker research is predominantly single-centre, retrospective,

and does not encompass the full range of potentially useful predictor factors, a prospective study that develops and evaluates current and novel biomarkers across multiple centres is required.

Objectives of the PROSPECT study

Primary objective

The PROSPECT study (improving PRediction Of metaStatic disease in Primary colorECTal cancer) aimed to improve prediction of metastasis in patients with colorectal cancer treated with curative intent. We accomplished this by developing a prognostic model of disease-free survival that incorporated conventional clinicopathological predictive variables and novel imaging (perfusion CT) and pathological variables. The primary outcome was the predictive performance of the best baseline standard model that also incorporated perfusion CT variables. We also aimed to examine overall survival at 5 years.

Secondary objectives

The following were the secondary objectives of the trial:

- To use robust statistical modelling to improve prediction by developing a 'best baseline' model that incorporates standard imaging and clinicopathological variables
- To assess the added value for prediction, if any, of perfusion CT parameters by comparing the primary outcome model (i.e. best baseline model plus perfusion CT variables) with the best baseline model alone
- To assess the detriment, if any, of simplifying the CT perfusion variables incorporated within the model (which would likely enhance generalisability)
- To assess the added value, if any, on model prediction by incorporating novel pathology biomarkers (i.e. markers of angiogenesis, hypoxia and MMR, RAS, RAF mutation status)
- To assess the detriment, if any, of simplifying the pathological variables incorporated within the model (which would likely enhance generalisability)
- To determine variability for perfusion CT measurements to estimate whether the limits of agreement are clinically acceptable at (1) local centres and (2) central review
- To assess if CT parameters from central review are significantly more predictive than those from local review
- To determine biological plausibility via exploration of associations between CT variables and pathology variables, including molecular markers of angiogenesis and hypoxia, and mutation status

Chapter 2 Methods

Trial design

PROSPECT (ISRCTN95037515) was designed as a prospective multicentre, observational cohort trial with 36 months of follow-up. The trial aimed to develop prognostic models to improve prediction of all recurrences in participants with colorectal cancer, incorporating both conventional predictive imaging and pathological variables and also novel perfusion CT and pathological variables.

Ethical permission was granted by Bloomsbury Research Ethics Committee (reference number 10/H0713/84) in 2011 and the trial was conducted in accordance with the principles of good clinical practice. The trial was co-ordinated by the Cancer Clinical Trials Unit, Scotland and overseen by an Independent Data Monitoring Committee (IDMC) and Trial Steering Committee.

All participants gave written informed consent prior to participation. Consecutive (i.e. unselected) eligible participants with known or suspected colorectal cancer underwent dynamic contrast enhanced CT imaging (perfusion CT) in addition to standard staging investigations.

As per normal practice, patient management was based on multidisciplinary team (MDT) decisions following standard staging examinations. Following surgery, standard pathological evaluation was performed as per the Royal College of Pathologists standards and data sets for reporting cancers: data set for colorectal cancer histopathology reports, July 2014, based on the fifth edition of the UICC TNM staging classification.

Participants' clinical course was followed for 36 months after the imaging intervention, including annual CT imaging for surveillance, and endoscopy at 3 years post treatment, as per usual practice. As per usual practice, investigations were performed in response to any unexpected/unanticipated symptoms. The date of any recurrence and/or death was recorded.

The trial was reported according to the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) guidelines for multivariable prediction models.

A summary of participant flow through the main trial is shown in [Figure 1](#). The full protocol for PROSPECT is available on the project web page <https://fundingawards.nihr.ac.uk/award/09/22/49>.

Patient and public involvement

The trial was developed in collaboration with patient representatives, who joined the trial team at project inception. The patient representatives helped to refine the research questions and devise the protocol, and helped with the funding application. All patient-facing materials were designed with patient representative input. Representatives also sat on the trial management and steering committees, providing guidance throughout the running of the trial (e.g. helping to refine recruitment strategies). Patient representatives have contributed to the project write-up, have seen this monograph and will aid dissemination of the work via patient forums.

Recruitment sites

Recruitment occurred at 13 NHS hospitals, a representative mix of district general hospitals and teaching hospitals, with multidetector CT scanners from the four major commercial vendors. Details are summarised in [Appendix 1](#), [Tables 22](#) and [23](#).

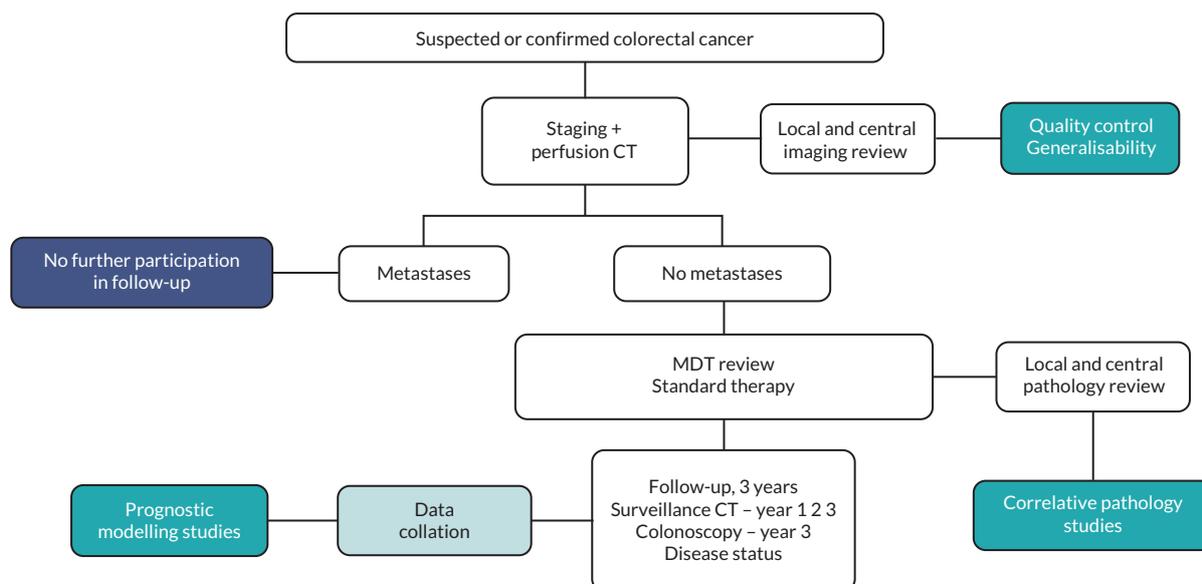


FIGURE 1 Overview schema of the PROSPECT trial.

Recruitment

Potentially suitable participants were identified from outpatient clinics, imaging requests and endoscopy lists, as well as MDT meetings, by members of the local clinical or research teams, who established whether or not the patient met trial entry criteria. All participants were given, e-mailed or posted a participant information sheet detailing the trial. The trial purpose and requirements were also explained to participants face to face with an appropriately trained member of the local research team. All participants gave written consent prior to participation. Participants retained a copy of their consent form and participant information sheet and were informed that they could withdraw at any time. Patient recruitment spanned 2011–6 with a 3-year follow-up for included patients to 2018.

Inclusion criteria

Inclusion criteria were as follows:

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

Exclusion criteria

Exclusion criteria were as follows:

- Contraindications to intravenous iodinated contrast agent administration, including renal impairment and prior contrast reaction.
- No mass visible with confidence on CT (i.e. intervention could not be applied).
- Stage IV disease at staging (i.e. metastasis already present).
- Previous colorectal or other cancer in the 5 years preceding potential recruitment.
- Diagnosis ultimately not colorectal cancer.
- Pregnancy.

Interventions

Perfusion computer tomography imaging

All eligible consenting participants underwent additional perfusion CT either at the same time as their staging CT, as an additional acquisition or on a different day if they could not be scheduled together. Perfusion CT was performed on multidetector CT scanners at all sites. At initial site set-up, phantom scanning was undertaken for quality assurance using a CATPHAN® 600 (Phantom Laboratory, Salem, NY, USA) and a water phantom with iodine inserts. Designated radiographers at each local site underwent dedicated training by the trial team prior to the commencement of patient scanning.

A typical acquisition for perfusion CT is shown in [Figure 2](#). First, a low-dose abdominal/pelvic CT scan was acquired to locate the colorectal cancer. If identified, this was followed by intravenous injection of iodinated CT contrast agent (> 300 mg/ml iodine concentration; 50 ml injected at 5 ml/second followed with a saline chaser 50 ml at 5 ml/second via a pump injector). Then, a dynamic scan centred on the primary tumour was acquired. The scan was obtained at 2.5-mm and 5-mm slice thickness, with the 5-mm thickness used for analysis due to lower noise. A tube voltage of 100 kV and tube current of 60–200 mAs was advocated. Five seconds following the start of intravenous injection, data were acquired every 1.5 seconds, for a total of 45 seconds, then every 15 seconds thereafter, for an additional 75 seconds. The dose constraint (limit) for the perfusion CT acquisition was set at 20 mSv to ensure that good image quality could be achieved with the different CT scanners located at the various participating centres.

If the perfusion CT acquisition was undertaken at the same visit as the staging CT, the perfusion CT scan was performed first, followed by standard staging CT (acquired following a further injection of the standard volume of contrast as per local practice). Information regarding perfusion CT was noted on a case report form (CRF), including scan date, technical details, completion or otherwise, and any complications related to perfusion CT. CRFs are available on the project web page <https://fundingawards.nihr.ac.uk/award/09/22/49>.

Perfusion computed tomography analysis

Local review and radiologist training

Perfusion CT was interpreted and analysed by designated trial radiologists at each local site, 25 radiologists in total. All received on-site training from the central trial team with respect to data acquisition and analysis at site set-up and proceeded with analysis following completion of a test set of cases on the local centre's software platform. All were general radiologists with a subspecialty interest in gastrointestinal imaging or subspecialty gastrointestinal radiologists, mirroring NHS practice should perfusion imaging be adopted in the NHS ultimately.

Image analysis used the local sites' commercial software platform provided by the CT scanner manufacturer [i.e. General Electric (GE), Siemens, Phillips, Toshiba]. Commercial software platforms were based on different kinetic analysis models, depending on the CT manufacturer, and included the distributed parameter model: Patlak analysis, deconvolution and maximum slope method (see [Appendix 1, Table 24](#)).

The steps undertaken during image analysis included defining the arterial input function [by placing a region of interest (ROI) within the arterial lumen], defining the end of the first pass of contrast agent, and defining the tumour ROI. This then generated the following tumour vascular parameters: regional blood flow, blood volume, mean transit time or permeability surface area product (depending on the software platform; [Figure 3](#)). Perfusion CT measurements were

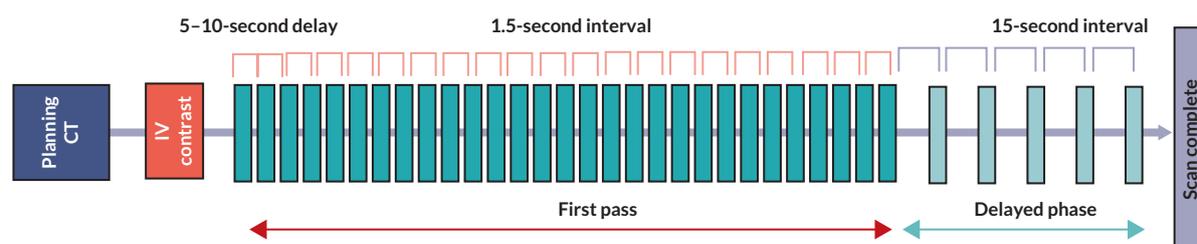


FIGURE 2 Schema of CT perfusion acquisition. IV, intravenous.

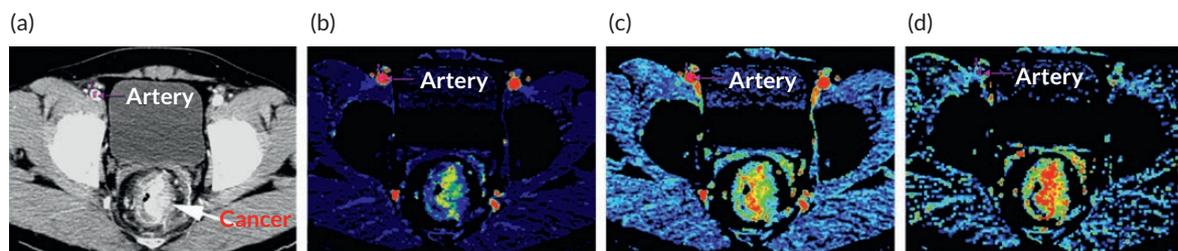


FIGURE 3 Example of parametric maps derived from perfusion CT image analysis. A vascular rectal cancer (a), white arrow, is shown with heterogeneous increased blood flow (b), blood volume (c) and permeability surface area product (d).

recorded on the CRF. Additional information captured on the CRF included tumour size (maximal cross-sectional diameter), location, TNM stage by standard CT criteria, venous invasion and technical aspects related to analysis and vascular parameters from the tumour ROI. CRFs are available on the project web page <https://fundingawards.nihr.ac.uk/award/09/22/49>.

Central review

Following imaging data transfer, perfusion CT was also reviewed centrally by three radiologists with 5–18 years of experience in perfusion CT. Radiologists undertaking central review were unaware of local measurements, and of findings from standard imaging investigations and outcomes. Image analysis was performed using the commercial software platform (GE, Siemens, Phillips or Toshiba) used at the local sites. In addition to image quality review (described in the quality assurance section), image analysis was undertaken as described previously to generate the same tumour vascular parameters, which were then recorded on the CRF.

Pathology

Immunohistochemistry

Tissue blocks from the local specimen were transferred centrally for further processing and assessment. Tissue sections 3 μ thick were obtained from each submitted tissue block and prepared for immunohistochemistry. The following were assessed:

- **DNA MMR protein status:** MMR status was determined by assessing expression of MLH1, MSH2, MSH6 and PMS2. MMR-deficient (equivalent to microsatellite instability in the majority of the cases) have a better prognosis than MMR-proficient patients (particularly in stage II colorectal cancer). MMR status is also used to inform clinical decisions (e.g. lack of response to 5-fluorouracil, potential benefit of immune checkpoint inhibitor therapy).
- **CD105 microvessel density:** CD105 is a proliferation and hypoxia-inducible protein associated marker expressed in angiogenic endothelial cells.
- **VEGF expression:** VEGF is produced by colorectal tumour cells, stromal cells and tumour infiltrating macrophages and is a key event for angiogenesis.
- **Glucose transporter protein (GLUT-1) expression:** GLUT-1 mediates cellular uptake of glucose and is upregulated under hypoxic conditions via the hypoxia-inducible factor 1-alpha (HIF-1 α) pathway to enable anaerobic glycolysis, providing an indirect marker of hypoxia.
- **HIF-1 expression:** HIF-1 is upregulated in hypoxic conditions, providing an indirect marker of hypoxia.

The following concentrations of antibodies were used: CD105 (Novocastra, Leica Biosystems, Nussloch, Germany; 1/200 dilution, discontinued); VEGF (Dako, Agilent Technologies, Santa Clara, CA, USA; concentration: 0.45 μ g/ml); GLUT-1 (Millipore, Merck, Darmstadt, Germany; concentration: 2.5 μ g/ml); HIF-1 α (Abcam, Cambridge, UK; concentration: 1.94 μ g/ml); MLH1 (Novocastra; concentration: 19.5 μ g/ml); MSH2 (Novocastra; concentration: 1.94 μ g/ml); MSH6 (Dako; ready-to-use antibody in 0.015 mol/l sodium azide) and PMS2 (BD Pharmingen, BD Biosciences, San Diego, CA, USA; concentration: 1.6 μ g/ml).

Sections were stained in batches on the fully automated BOND-MAX system (Leica Biosystems), which was used in conjunction with the BOND Polymer Refine detection system (Leica Biosystems).

All slides were scanned at 20 × magnification using a Nanozoomer 2.0 RS (Hamamatsu Photonics, Herrsching, Germany) and images were exhibited in a liquid crystal display monitor under contrast, focus, saturation and white balance standardisation.

Mismatch repair protein expression was assessed by determining retained expression or lack of staining in tumour areas of the sample. 'Internal' controls (i.e. lymphocytes, stromal cells, non-tumour crypts) were used in the tissue sections as markers of 'normal' (retained) expression. Generally, MLH1 and PMS2 work as a pair, as do MSH2 and MSH6. Therefore, lack of staining in MLH1, for example, results in lack of expression of PMS2. However, PMS2 expression loss can occur in isolation, and this will raise the possibility of the patient having Lynch syndrome as a result of a germline mutation in PMS2. All antibodies were localised to the cell nucleus with some faint cytoplasmic occasionally seen.

CD105-stained vessels with a clearly defined lumen or well-defined linear vessel shape were considered for microvessel assessment. The invasive front of each sample was selected for assessment. The two areas of highest vascularisation (hot spots) were averaged and given as a count per mm².

Scores for VEGF, GLUT-1 and HIF-1 α were based on staining intensity and the proportion of positively stained cells, according to previously published systems. VEGF and GLUT-1 expression was calculated by combining the intensity of stained cells (0–3) with the proportion of positive cells (0–4), and HIF-1 α expression on the combined cytoplasmic and nuclear staining (range 0–6).

Image analysis software (Visiopharm®, Hoersholm, Denmark) was used to evaluate CD105 staining. A histopathologist with more than 15 years' experience of gastrointestinal pathology performed semiquantitative analysis of immunoreactivity of the other markers.

Somatic mutation analysis

Formalin-fixed paraffin embedded (FFPE) tissue blocks were cut at 10 μ m and extracted using a Qiagen FFPE DNA extraction kit (Qiagen NV, Venlo, Netherlands). DNA quality and quantification were assessed using an Agilent TapeStation 2200 System (Agilent Technologies, Santa Clara, CA, USA). Amplification and library preparation of the samples was done using the Roche High Fidelity PCR Master system (Roche Diagnostics, Mannheim, Germany) and custom designed primer targets (KRAS, NRAS, BRAF, PIK3CA, pTEN, APC, HRAS), which were run on the Biomark HD system (Fluidigm Corporation, San Francisco, CA, USA). Sequence preparation and sequencing was performed using the Life Technologies Ion Torrent system (Thermo Fisher Scientific, Waltham, MA, USA). Sequencing data were analysed on an Integrative Genomics Viewer.

Quality control of perfusion computed tomography

Iodine phantom imaging

A cylindrical water phantom containing different iodine inserts was scanned by 7 CT platforms from the 13 participating hospitals. The relationship between measured CT number (Hounsfield units) and iodine concentration (milligrams per millilitre) within the inserts was established and contrast-to-noise ratios calculated. Radiation doses (CT dose index, dose-length product) of the acquisition were compared across all sites.

Clinical image quality

Central review of perfusion image quality was undertaken in a subset of participants (up to 20% of those recruited) to ensure consistency of image quality across the different scanners and sites. Consecutive participants' scans were triaged to select consecutive perfusion CT scans of the lower pelvis that included a distended bladder. Image quality evaluation included image noise [assessed by placing a ROI within the distended bladder and measuring the standard deviation (SD)], tumour contrast-to-noise ratio (assessed by placing a ROI within the tumour, measuring tumour enhancement and calculating the ratio of enhancement to noise) and quality of the arterial input function (by placing a ROI within the arterial lumen and measuring peak enhancement and the full width half maximum of the enhancement-time curve).

Standard staging investigations and planned treatments

Recruited participants underwent standard staging investigations locally, according to local protocols and clinical care team requirements. All standard investigations were performed and interpreted by the usual local care team. The type and date of all standard imaging investigations and MDT staging decisions were recorded on the CRF. Planned management was as per local policy, guided by NICE recommendations and decided at MDT meetings. There was no change to usual patient treatment trajectory contingent on trial participation. For participants undergoing curative treatment, this included curative primary tumour resection; \pm neoadjuvant (chemo)radiation; \pm adjuvant chemotherapy. Treatment decisions were recorded on the CRF.

Standard pathological evaluation

Following surgical resection, specimens were fixed in 10% neutral buffered formalin solution. Standard pathological evaluation and reporting was performed by local consultant pathologists as per the Royal College of Pathologists Standards and data sets for reporting cancers: data set for colorectal cancer histopathology reports, July 2014. Tumour blocks were processed conventionally and embedded in paraffin wax. Sections 4- μ thick were obtained from each block and stained with haematoxylin and eosin. Staging was based on the fifth edition of the UICC TNM staging classification. CRFs are available on the project web page <https://fundingawards.nihr.ac.uk/award/09/22/49>. For quality assurance, central review of pathological evaluation was undertaken.

Data collection and follow-up

Data collation via CRFs was co-ordinated by the Clinical Trials Unit. Initial clinical data included participant age, sex, type and date of staging investigations, carcinoembryonic antigen (CEA), MDT stage, MDT outcome and treatment(s) undertaken. All CRFs were collated by local research nurses/practitioners and transferred to the clinical trial unit by post or fax. Form contents were then entered into a bespoke trial database and any missing fields or apparent data inaccuracies queried with the local site to optimise and maximise data collection.

Participants were followed for 36 months (or until date of death, if sooner) to develop the prognostic model. During this time, findings from annual CT surveillance (or arising from any CT performed at other intervals in response to symptoms) were recorded, together with outpatient visits and CEA levels. A colonoscopy at year 3 was optional, according to local practice.

Evidence of recurrence (new metastasis, local recurrence, new tumour) alongside any histological findings from any further surgical resections or biopsies and date of death were recorded.

Outcomes

A summary of the primary and secondary outcomes is shown in [Table 1](#).

Sample size estimation

We estimated that with 10 centres each recruiting 3 participants/month (representing around 50% of potentially eligible participants, a recruitment rate achieved in the previous single-centre study⁵⁴), over 12–15 months, a prospective cohort trial could recruit 370 participants with a median follow-up of 40 months. We estimated 30% of participants would develop metastasis subsequent to treatment with curative intent, with most events occurring within 36 months.⁶⁰ This gave an effective sample size of approximately 80 participants for the primary outcome, using the end point of uncensored time to metastasis [taking into account participants with metastases identified at primary staging (up to 20%) who would then be excluded]. A second end point of time to death from any cause would also be collected. Based on a reclassification index similar to Pencina *et al.*⁶¹ of participants at high risk for metastasis (top 30%) compared

TABLE 1 Summary of overall study aim, primary and secondary outcomes

Aim/outcome	Description
Overall study aim	To improve prediction of metastatic disease in participants with colorectal cancer, treated with curative intent, by developing a prognostic model based on DFS, that is superior to current practice, via prospective evaluation of both conventional predictive variables (comprising the 'baseline' model) and novel imaging and pathological variables
Primary outcome	Best prognostic model for DFS, that combines the baseline model and local perfusion CT variables, compared with clinical standard of TN staging
Secondary outcome 1	Best baseline model (i.e. standard variables alone) compared with TN staging
Secondary outcome 2	Added value (if any) when local perfusion CT variables are combined with the baseline prognostic model
Secondary outcome 3	Added value (if any) of alternative/simplified scores for perfusion CT variables
Secondary outcome 4	Added value (if any) when novel pathology variables are combined with the baseline prognostic model
Secondary outcome 5	Comparison of perfusion CT variables measured locally and centrally, and of overall measurement variability
Secondary outcome 6	Added value (if any) when central perfusion CT variables are combined with the baseline prognostic model
Secondary outcome 7	Impact on model performance of extra participants diagnosed with true-positive or false-positive results. <i>Comparisons of model performances at different relative clinical weightings</i>
Secondary outcome 8	Exploratory analysis investigating potential relationships between perfusion CT variables and pathology characteristics, to determine biological plausibility. <i>Comparisons based on tumour characteristics and not prognostic models</i>

DFS, disease-free survival; TN, tumour and nodal.

between the two models, 300 participants with 80 events would have 80% power to detect a 15% difference in correct risk classifications,⁶² with allowance for loss to follow-up (estimated at < 10% from previous study experience⁵⁴).

During trial accrual and following interim presentation of trial data to the IDMC, the IDMC recommended the sample size be increased to 445 participants to address a higher-than-anticipated number of participants withdrawn due to metastasis at staging. The trial then recruited until the original target of evaluable participants with an event was achieved.

Definition of end points for disease-free survival outcome

The start of the disease-free survival period was defined as the date of primary CT staging. Definitions for the end of the period are summarised in [Appendix 1, Table 25](#).

Treatment of multiple end points within a patient was as follows: the timing of the first event for each category was collected but not the second event within the same definition. Timing of survival in the same patient was recorded. Events detected from tests arising from the clinical visit were classified as simultaneous. In the case of simultaneous events, the 'worst' event was defined as the first.

Recurrence was defined as the first event of distal recurrence, development of new primaries, or death from any cause, within 3 years of diagnosis.

Analysis

Summary

A series of time-to-event nested prognostic models were developed and models were compared with current clinical practice (i.e. TN staging) and to each other. Analyses for the primary and secondary outcomes are summarised in [Table 2](#). Bland–Altman analysis was performed to assess perfusion CT agreement between local and central

measurements. Potential relationships between perfusion CT variables and pathology characteristics were assessed using descriptive analysis.

Study aim

To improve prediction of metastatic disease in participants with colorectal cancer treated with curative intent by developing a prognostic model based on disease-free survival that is superior to current practice, via prospective evaluation of both conventional predictive variables and novel variables derived from perfusion CT.

Definition of current clinical practice

Current clinical practice was defined as participants at high risk for recurrence as defined by stage III disease using the AJCC staging of colon and rectal cancer (see [Appendix 1, Table 26](#)). The entire AJCC staging schema is shown in [Appendix 1, Table 21](#).

Method for model development

We developed a time to event prognostic model using measurements from perfusion CT combined with standard prognostic variables known already to be predictive of risk of disease recurrence. For the primary outcome, the event of interest was recurrence and death during the 3-year follow-up period.

Interval censoring was present in this trial because recurrence was assessed by periodic testing. Although there are specialised statistical methods we could have employed to account for interval censoring, we did not use these as we considered them unlikely to impact on trial results. This was because the appearance of symptoms suggestive of

TABLE 2 Analyses undertaken for primary and secondary outcomes

Outcome	Summary	Methods compared in outcome	
		Method 1	Method 2
Primary	Best model including perfusion CT compared with current method: risk threshold (high vs. medium/low risk)	Prognostic model for DFS, including standard variables and CT perfusion variables	Current clinical practice
	Sensitivity analysis: risk threshold (high/medium vs. low risk)	Model B: model A as linear predictor and in addition perfusion CT variables (PCA variables or up to two variables)	Rule C: based on current practice variables of T and N stage only
Secondary 1	Best model with standard variables compared with current method	Prognostic model for DFS with standard variables only (no perfusion CT variables)	Current clinical practice
		Model A: best model based on all standard clinical variables	Rule C
Secondary 2	Added value of perfusion CT variables in prognostic model	Prognostic model including standard variables and perfusion CT variables	Prognostic model with standard variables only (no perfusion CT variables)
	Sensitivity analysis: different risk threshold (high/medium risk vs. low risk)	Model B	Model A

TABLE 2 Analyses undertaken for primary and secondary outcomes (continued)

Outcome	Summary	Methods compared in outcome	
		Method 1	Method 2
Secondary 3	Added value of alternative scores for perfusion CT variables	Composite single score of four parameters in prognostic model Model B	Simpler scores for perfusion CT variables (e.g. single or pairs of parameters) Model D: based on model A as linear predictor and in addition simpler score (e.g. using single or pair of perfusion CT variables)
Secondary 4	Added value of pathology variables in prognostic model	Preferred prognostic model from trial with pathology variables Model F: model A plus pathology variables	Preferred prognostic model from trial without prognostic variables Model A or B
Secondary 5	Comparison of perfusion CT variables and variability; Bland–Altman analysis	Local hospital measurements	Central review measurements
Secondary 6	Added value of perfusion CT, based on central review data	Prognostic model including standard variables and perfusion CT variables from central review data Model E: model A plus perfusion CT variables from central review (PCA variables or up to two variables)	Prognostic model with standard variables only (no perfusion CT variables) Model A
Secondary 7	Impact on model performance of extra participants diagnosed with true-positive or false-positive results	Comparison of model performances at different relative clinical weightings Use in primary outcome and secondary outcome 1	
Secondary 8	Exploratory analysis investigating potential relationships between perfusion CT variables and pathology characteristics in the tumours; descriptive analysis	Comparisons based on tumour characteristics and not prognostic models	

DFS, disease-free survival; PCA, principal components analysis.

recurrence would trigger immediate clinic visits and rapid testing thereafter. The time for the event of recurrence was only known to be between the current and previous examination.

We used flexible parametric modelling so that the baseline model was based on a parametric model using `stpm2` in STATA 14.1 (StataCorp LP, College Station, TX, USA). This was a flexible parametric approach that uses cubic splines to model the baseline hazard, where the splines are a piecewise function with boundaries defined by knots, where the degrees of

freedom (df) are indicative of the number of knots. We compared the baseline survival hazard function with 1 df to the semiparametric Cox regression baseline, and 2–6 df. We note that Cox models do not estimate directly the baseline hazard as is the case with flexible parametric and Weibull models. Higher df did not increase fit [Akaike information criterion (AIC), Bayesian information criterion (BIC)] of baseline survival, so we used parametric modelling with a Weibull model with no spline knots at 1 df. We checked our modelling by comparing the c-index with semi-parametric Cox models. The Weibull model gave a more credible baseline than the Cox model, reflecting the biological rate of recurrence as increasing monotonically, as opposed to peaks that artificially reflect timing of regular annual CT surveillance, which was when most recurrences were detected. The baseline model included a constant term and a shape parameter (rsc) which was multiplied by the natural log of time in years. Models were fitted on the log cumulative hazard scale $\{\ln[-\ln S(t)]\}$ with proportional hazards. Nested models of model A coefficients were fitted using a fixed offset of the model A linear predictor, allowing the effect of additional variables to be assessed. Continuous variables were centred on values close to the median; therefore, in this trial, age was centred for 65 years and tumour size was centred at 40 mm.

Included variables

We prespecified the following standard clinicopathological variables for inclusion in the full model: N stage, T stage, age, sex, tumour size, tumour location, venous invasion, treatment. For the primary outcome, to this model we added a perfusion CT summary score, developed from principal components analysis (PCA). Standard prognostic variables were selected based on clinical consensus among experts in the study team based on published studies in colorectal cancer. Categorical variables were coded in the modelling as follows: N stage N0 no lymph node involvement found, N1 cancer cells in one to three nearby lymph nodes, N2 cancer cells found in four or more nearby lymph nodes; T stage T1, T2, T3 and T4 based on American Cancer Society staging of colon and rectum cancer, T staging from TNM staging; sex as binary male/female; tumour location as left (descending colon, sigmoid, rectum) or right (colon, ascending colon), where middle tumours (transverse colon) were grouped with left; venous invasion as binary yes/no; treatment four groups – immediate surgery with no chemotherapy, immediate surgery with chemotherapy, late surgery, no surgery; HIF1a as binary variable 0 = score of 0–2, 1 = score of 3–6; MMR as binary with MMR proficient compared with a reference standard of deficient corresponding to tumour with mutation; the gene mutations BRAF not 600, BRAF 600. KRAS, HRAS, NRAS were categorised as binary corresponding to patient has mutation or has wild-type gene, regardless of the number of mutations, except for BRAF, where mutations were counted at two gene locations separately.

Tumour location (colon vs. rectum, the latter including rectosigmoid) was originally included in the statistical analysis plan as a subgroup analysis but was included instead as a variable to improve statistical power of comparing these tumour locations.

The following variables were based on pathology of the primary tumour at baseline, where available, or CT imaging at baseline if neoadjuvant treatment was given: T stage, N stage, venous invasion, tumour size. Tumour location was based on diagnostic CT imaging. No variable selection by significance testing was undertaken, to conserve statistical power.

We prespecified incorporating perfusion CT measurements from individual centres for model building for the primary outcome (i.e. 'local' measurements), since this would better reflect likely clinical practice should the model be adopted. We elected to investigate central review perfusion CT measurements within models as part of secondary analyses (important, because improvements in data transfer mean that central review is now a realistic proposition for day-to-day clinical practice). Continuous data were retained for all the included continuous variables, to retain statistical power.⁶³

We checked the relationship of continuous variables (age, tumour size, CD105) with outcome, using flexible parametric models. Linear relationships were appropriate for both variables, based on automatic selection of the best fitting model and manual confirmation of best fit using AIC and BIC (using STATA commands `fp`, using methods from *Flexible Parametric Survival Analysis Using STATA: Beyond the Cox Model*).⁶⁴ Assumptions of proportional hazards were evaluated graphically and there was no evidence of assumptions being violated, based on Schoenfeld residuals. No interactions between variables were considered.

Methods to generate principal components score from perfusion computed tomography parameters

From analysis of prior study data,⁵⁴ we identified that several perfusion CT parameters were correlated, so prior to model building we prespecified development of a summary score using PCA. A summary score is recommended to

summarise collinear variables because it is more powerful and stable than arbitrary selection of one variable from the group.^{65,66} We used PCA to build a score based on the perfusion CT parameters (STATA PCA commands). Each perfusion CT parameter was standardised to a mean of 0 and a variance of 1, to avoid undue influence due to different measurement scales. We retained principal component composite variables, including components with eigenvalues of greater than one, which corresponded to the first two principal component scores.

Imputation to account for missing data

After data cleaning, 324 of 326 included participants had complete data for conventional clinical variables. Therefore, analysis was conducted using these 324 participants for model A and rule C. For perfusion CT and pathology measurements, more participants had missing data. However, as comparisons of different models were based on the additive effect within patients of prespecified novel variables to a model already containing conventional clinicopathological variables (by comparison of nested models), there was no statistical advantage to using multiple imputation of missing data. This is because the imputation would be based on data from all patients and so incorporate between patient variation into the within patient analysis, where all data were present. This meant that using imputation would not increase statistical power in our model comparisons.

Outcomes presented from model

Our primary measure of model performance was based on the improvement in correct predictions for individual participants between model B and rule C.

We generated three risk groups ('high', 'medium' and 'low') from model B based on tertiles splitting trial participants into three equal-sized groups, based on their risk predicted from the model. For the primary outcome, two risk groups were used (high and medium/low). A sensitivity analysis was conducted using two risk groups (high/medium vs. low). Rule C is standard clinical practice and generates two risk groups (high and low). Approximately 30% of participants were anticipated to develop recurrence.

Model performance was presented using Kaplan–Meier plots according to risk groups, plotted with 95% CIs based on each time point estimate. These CIs are only valid for comparing the risk groups at particular time points, such as at 3 years. Nelson–Aalen cumulative hazard estimates at 3 years with 95% CI were used to identify whether there was a statistically significant difference between risk groups in a model. A 2 × 2 table to estimate the number of participants for each risk group with and without recurrence at 3 years was calculated using Nelson–Aalen cumulative hazard estimates. This used a non-parametric method for estimating probability of an event within a group of participants, making no assumptions about the distribution of events and appropriately taking censoring into account. To be able to compare between models based on all recurrences in participants, the number of participants with recurrence was proportionally adjusted to the total of 81 events in 324 participants at 3 years, to allow better comparison of sensitivity and specificity outcomes in models A, B and E for the same participants.

Results were expressed in terms of comparison between models for predicting participants at high risk of recurrence within 3 years in terms of: (1) difference in sensitivity and specificity; (2) described based on a hypothetical population of 1000 participants diagnosed with colorectal cancer; and (3) difference in proportion of participants reclassified.

Other model performance measures

We also presented standard measures of prognostic model performance, such as discrimination and calibration, including c-index and calibration slope. Comparison of model fit between nested models was evaluated based on log likelihood of model fit, AIC and BIC. Calibration plots were not included as there was no data set available for external evaluation.

Internal evaluation

All variables were prespecified and no variable selection based on statistical significance was used for model development. Internal evaluation was used to assess any over-optimism of c-index, R^2 and adjusted R^2 using bootstrapping (100 repeats).

Model shrinkage

Shrinkage was used to improve future model fit, by reducing bias due to minimisation of the mean squared error in regression methods used for model development. We used uniform shrinkage of regression coefficients after estimation using heuristic shrinkage methods, based on the formula:

$$s = \frac{(\text{model } \delta^2 - df)}{\text{model } \delta^2}$$

where δ^2 is the likelihood ratio squared of the fitted model (i.e. the difference in 2log likelihood between the model with and without predictors) and df indicates the degrees of freedom based on the number of pre-specified predictors.

Secondary outcomes

Secondary outcomes are reported as per [Table 1](#). Analysis methods were identical to those described for the primary outcome except where noted in this section.

Methods for comparing models

To establish whether the addition of novel predictor variables (i.e. perfusion CT, immunohistochemistry, genetic variables) would improve prediction of patient outcomes, we compared nested models using standard clinicopathological variables to the same models following addition of novel variables, in the same participants. The linear predictor for the clinical variables was included in modelling as a fixed offset (constraint) with or without additional predictors.

We compared model predictions for high recurrence risk in terms of: (1) difference in sensitivity and specificity; (2) difference in model fit based on AIC and BIC; and (3) visual examination of a scatter plot of prediction index comparisons from two models.

Secondary outcome 3

Graphs and correlation matrices were used to present the correlation of perfusion CT parameters. Simplified methods for including perfusion CT parameters in model D were evaluated using single or pairs of perfusion CT parameters. Comparison of central and local perfusion CT variables identified those variables with better reliability, for inclusion in the simpler model D versions.

Secondary outcome 5

Bland–Altman methods were used to determine limits of agreement for perfusion CT measurements between local and central review. Measurement methods were considered interchangeable, where the limits of agreement are narrower than (or the same as) differences that might be considered within clinically acceptable variability. We presented graphs to examine potential sources of variation related to the technical acquisition of CT data (e.g. scanner type, analysis software) and ROI area.

Secondary outcome 7

We presented the impact on the model operational cut-point for risk, of different weightings of the clinical and patient assigned values for: (1) correct prediction of an additional patient with metastasis, and (2) one less patient given a false prediction of metastasis.

Secondary outcome 9

We explored the potential relationships between perfusion CT parameters and pathology characteristics in the tumours including tumour stage; immunohistochemistry (CD105, VEGF, GLUT-1), MMR mutation status, extramural venous invasion, and tumour regression score post chemoradiation.

Chapter 3 Results

Participants

Recruitment commenced 2011 and completed 2016. Participant flow through the trial is shown in [Figure 4](#). Of 448 participants recruited, 122/448 (27%) were withdrawn for the following reasons: metastases identified at staging ($n = 48/122$, 39%); non-cancer or non-colorectal diagnosis ($n = 23/122$, 19%); additional cancer at diagnosis ($n = 9/122$, 7%); previous diagnosis of cancer ($n = 4/122$, 3%); perfusion CT not performed (23/122, 19%) or technically unsuccessful ($n = 11/122$, 9%); other ($n = 4/122$, 3%). A screening log was not kept locally.

The final cohort consisted of 326 participants (226 male, 100 female), mean \pm SD age of 66 ± 10.7 years. Participant characteristics are summarised in [Table 3](#).

Cancers were located in the colon in 143/326 (44%) participants and the rectum (including rectosigmoid) in 183/326 (56%) participants. Median for tumour size was 40 mm [interquartile range (IQR) 30–50 mm].

Treatments that participants received are summarised in [Appendix 1, Table 27](#). Neoadjuvant therapy was administered in 79/326 (24%). Surgery was performed ultimately in 306/326 (94%). Adjuvant therapy was administered in 125/326 (38%). Twelve out of 326 (3%) participants, all with rectal cancer, did not proceed to surgery following neoadjuvant therapy due to complete response. Five out of 326 (2%) participants with rectal, transverse colon, ascending colon and caecal cancers, respectively, received no treatment.

Staging

Baseline tumour stage for all participants is summarised in [Table 4](#). The majority of cancers were locally advanced (i.e. \geq T3 stage, 227/326, 70%); 151/326 (46%) of participants were node positive (i.e. \geq N1 stage). By definition, no recruited participant had metastases (i.e. all were M0 at baseline).

Perfusion computed tomography

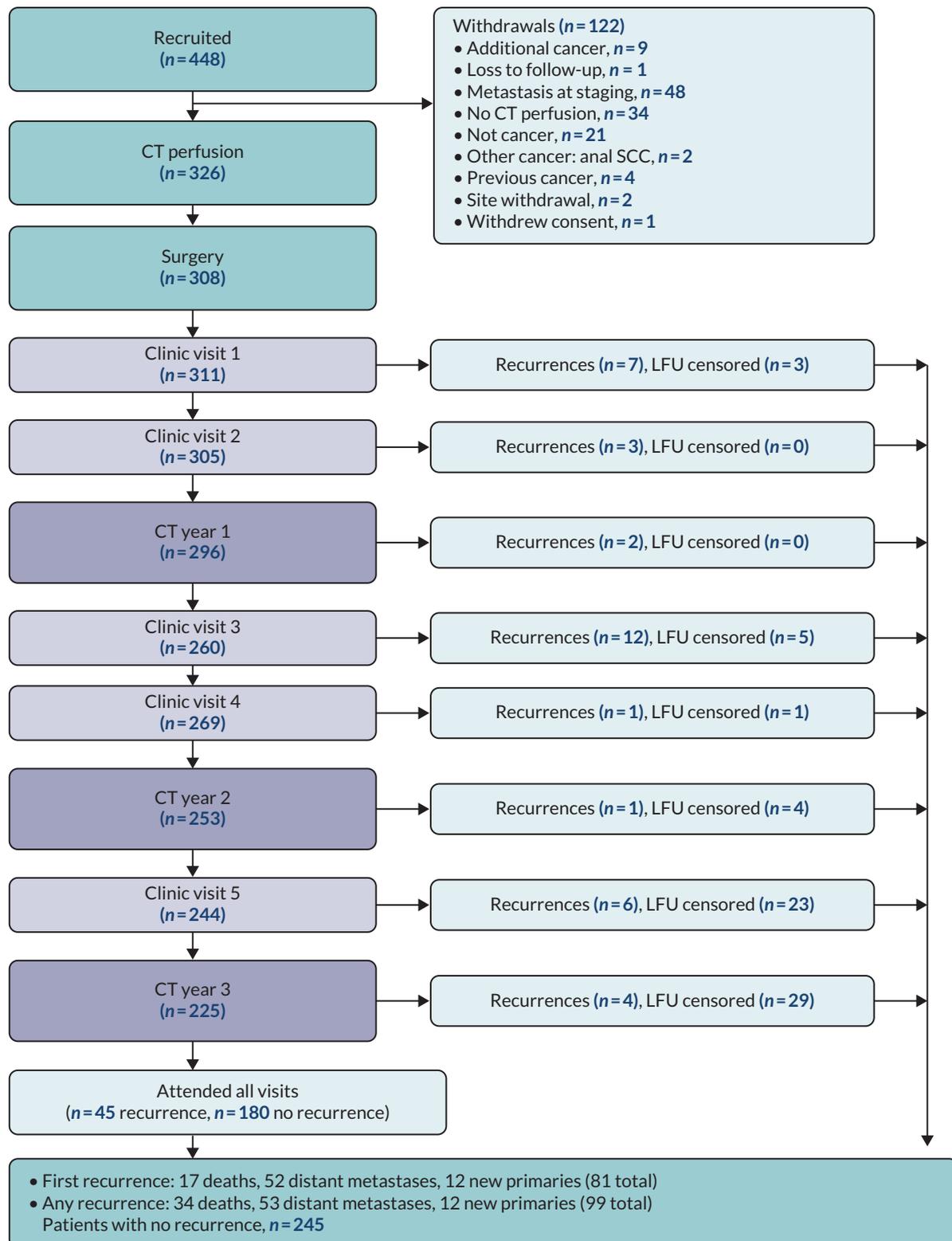
Local site review

Perfusion CT measurements obtained locally were available for 303/326 (93%) participants. Perfusion variables split by TN status are summarised in [Appendix 1, Tables 28 and 29](#). No significant difference in perfusion measurements was noted between T1/T2 tumours versus T3/T4 tumours, and N0 versus N1/N2.

Perfusion variables obtained locally, split by recurrence status are summarised in [Table 5](#). Blood flow, blood volume, permeability surface area product and mean transit time were not significantly different when participants with and without recurrence were compared.

Central review

Perfusion CT measurements obtained centrally were available for 291/303 (96%) participants who also had corresponding local perfusion CT measurements available. Perfusion variables obtained centrally, split by TN status are summarised in [Appendix 1, Tables 30 and 31](#). Perfusion variables obtained centrally, split by recurrence status, are summarised in [Appendix 1, Table 32](#). No significance difference in blood flow, blood volume, permeability surface area product and mean transit time was found between groups.



Note that for recurrences the flow chart reports the clinic visits and scans where information is available, not the time of recurrence.

FIGURE 4 Participant flow through the trial. LFU, loss to follow-up; SCC, squamous cell carcinoma.

TABLE 3 Characteristics of final participant cohort

Variable	Recurrence	No recurrence	Total
Sex	N (%)	N (%)	N (%)
Female	25 (31)	75 (31)	100 (31)
Male	57 (69)	170 (69)	226 (69)
Age (years)			
Mean (SD)	70.3 (9.6)	64.6 (10.7)	66.0 (10.7)
Median (IQR)	71.0 (65.0–75.5)	66.0 (59.0–72.0)	67.0 (60.0–74.0)
Range	37.0–92.0	28.0–90.0	28.0–92.0
Tumour characteristics			
Tumour location	N (%)	N (%)	N (%)
Caecum	11 (14)	18 (7)	29 (9)
Ascending colon	10 (12)	15 (6)	25 (8)
Transverse colon	3 (4)	22 (9)	25 (8)
Descending colon	1 (1)	12 (5)	13 (4)
Sigmoid colon	11 (14)	40 (16)	51 (16)
Rectosigmoid	2 (3)	11 (5)	13 (4)
Rectum	43 (52)	127 (52)	170 (51)
Tumour type	N (%)	N (%)	N (%)
Adenocarcinoma	80 (99)	244 (99)	324 (98)
Other	1 (1)	1 (1)	2 (2)
Tumour size (mm)			
Median (IQR)	40.0 (30.0–50.0)	40.0 (30.0–50.0)	40.0 (30.0–50.0)
Range	18.0–75.0	10.0–150.0	10.0–150.0
TNM staging			
T stage	N (%)	N (%)	N (%)
T1	1 (1)	11 (5)	12 (4)
T2	12 (15)	75 (31)	87 (27)
T3	45 (56)	138 (55)	183 (55)
T4	23 (28)	21 (9)	44 (14)
N stage	N (%)	N (%)	N (%)
N0	38 (47)	137 (56)	175 (54)
N1	26 (32)	73 (30)	99 (30)
N2	17 (21)	35 (14)	52 (16)
M stage	N (%)	N (%)	N (%)
M0	81 (100)	245 (100)	326 (100)

TABLE 4 Summary of TN staging, by location (colon or rectum).

Stage	N0	N1	N2	Total
	N (%)	N (%)	N (%)	N (%)
Colon				
T1	3 (100)	0 (0)	0 (0)	3 (100)
T2	24 (77)	5 (16)	2 (7)	31 (100)
T3	48 (59)	24 (30)	9 (11)	81 (100)
T4	8 (29)	11 (39)	9 (32)	28 (100)
Total	83 (58)	40 (28)	20 (14)	143 (100)
Rectum^a				
T1	8 (89)	1 (11)	0 (0)	9 (100)
T2	39 (70)	13 (23)	4 (7)	56 (100)
T3	40 (39)	38 (37)	24 (24)	102 (100)
T4	5 (31)	7 (44)	4 (25)	16 (100)
Total	92 (50)	59 (32)	32 (18)	183 (100)

^a Includes rectal and rectosigmoid cancers.

Note

By definition, all patients were M0 at baseline.

TABLE 5 Local site perfusion CT measurements for participants with and without recurrence

Local review	Recurrence				No recurrence			
	Participants with data (n)	Mean (SD)	Median (IQR)	Range	Participants with data (n)	Mean (SD)	Median (IQR)	Range
Blood flow (ml/minute/100 ml or 100 g)	78	72.9 (40.4)	62.5 (52.6–85.1)	27.5–350.8	225	69.2 (35.7)	63.1 (47.3–81.9)	0–248.0
Blood volume (ml/100 ml or 100 g)	76	13.1 (8.4)	11.3 (6.5–16.3)	0.6–46.7	220	12.9 (7.4)	12.5 (6.8–16.5)	0–45.5
Permeability surface area product (ml/minute/100 ml or 100 g)	70	18.1 (14.9)	13.8 (8.8–19.9)	0.3–66.8	199	16.5 (13.4)	12.8 (8.9–18)	0–72.1
Mean transit time (seconds)	68	13.2 (5.8)	11.6 (9–17.7)	4.4–29.7	186	13.8 (5.8)	13.3 (9.2–18.1)	3.4–33.6

Pathology

Pathological evaluation

Information regarding venous invasion, tumour involvement of the resection margin, and tumour regression grade (TRG) following chemoradiation in rectal or rectosigmoid cancers is summarised in [Table 6](#). Venous invasion was present in a higher proportion of participants with recurrent disease [36/81, 44% (95% CI 43% to 55%)] than without [57/233, 24% (95% CI 19% to 30%)], a difference of 20% (95% CI 8% to 32%, *p*-value for unpaired comparison 0.0007). No participant who achieved complete/near-complete regression following neoadjuvant therapy (TRG 1) recurred within the 3-year follow-up period.

Immunohistochemical and somatic mutation analysis

Tumour blocks were provided for further immunohistochemical and genomic analysis in 270/326 (83%) participants. CD105 expression in participants with and without recurrence is summarised in [Appendix 1, Table 33](#). The number of CD105 stained vessels/mm² field was similar for patients who did and did not recur.

TABLE 6 Summary of pathological characteristics

Variable	With recurrence	Without recurrence	Total
Venous invasion	N (%)	N (%)	N (%)
Present	36 (44)	57 (23)	93 (28)
Absent	45 (56)	188 (77)	233 (72)
Resection margin	N (%)	N (%)	N (%)
Involved by tumour	3 (3.7)	12 (4.9)	15 (4.6)
Clear of tumour	54 (66.7)	183 (74.7)	237 (72.7)
Missing data	24 (29.6)	50 (20.4)	74 (22.7)
TRC^a	N (%)	N (%)	N (%)
1	0 (0)	9 (25)	9 (19)
2	4 (33)	13 (36)	17 (36)
3	6 (50)	11 (28)	17 (33)
Missing data	2 (17)	4 (11)	6 (12)

^a Applies to cancers treated with preoperative (chemo)radiation and surgery.

Immunohistochemical scores for HIF-1 α , VEGF and GLUT-1 are summarised for participants with and without recurrence in [Appendix 1, Table 34](#). The distribution of HIF-1 α , VEGF and GLUT-1 scores were similar across patients who did and did not recur.

Mutations in DNA MMR, RAS oncogene family (KRAS, HRAS, NRAS) and BRAF proto-oncogene serine/threonine kinase (BRAF) genes are summarised for participants with and without recurrence in [Appendix 1, Table 35](#). The proportion of participants with KRAS wild type was higher in participants with recurrence [34/62, 55% (95% CI 43% to 67%)] than without [96/208, 46% (95% CI 40% to 53%)] with a non-significant difference of 9% (95% CI 5% to 23%).

Central review

Central review of standard pathological reporting performed locally was conducted for 30/270 (11%) participants. Discrepancies were noted in 5/30 (17%) cases. In two cases, T stage was upstaged centrally from T2 to T3 (major discrepancy) and T4a to T4b (minor discrepancy). In the remaining 3/30 cases, discrepancies were related to additional findings of small vessel invasion ($n = 2$) and intramural venous invasion ($n = 1$) related to changes in recommended Royal College of Pathologists reporting guidance since January 2018.⁶⁷

Quality control

Iodine phantom imaging

As expected, scanning a cylindrical water phantom containing inserts with different iodine concentrations confirmed a linear relationship between iodine density and CT values (Hounsfield unit number). Across the 13 participating centres, iodine enhancement varied by a factor of up to 1.1. At an iodine concentration of 1 mg/ml, image contrast to noise ranged from 3.6 to 4.8 in the 220-mm phantom, dropping to 1.4–1.9 in the 300-mm phantom but remaining adequate. Radiation dose varied by a factor of up to 2.4 across centres but remained within the study constraints of 20 mSv maximum. Iterative reconstruction algorithms did not alter CT values significantly but resulted in reduced image noise by a factor of up to 2.2.

Perfusion computed tomography image quality

Image quality was assessed for 86 scans, representing imaging acquired by 5 different CT scanner models from 9 participating centres.

Image noise

Mean noise value was 20.6 HU (SD 4.6 HU). For the majority of participating centres, mean noise was below 21 HU, with the exception of three centres ([Appendix 1, Table 36](#)), where this was related to a suboptimal protocol (PC01), which was adjusted after the initial three participants scanned, older CT scanner model and participants with a higher body mass index scanned (PC03) and sharper image reconstruction kernel (PC11).

Contrast-to-noise ratio

Mean contrast-to-noise value was 1.6 (SD 0.04) for all imaging. Mean contrast-to-noise value was 1.7 (SD 0.66) for Siemens scanner models and 1.5 (SD 0.47) for GE scanner models, both within acceptable limits.

Arterial input function

Values for peak arterial iodine enhancement and full width half maximum of the initial arterial enhancement peak for each centre are shown in [Appendix 1, Table 37](#). There was a difference between GE and Siemens scanners, which, in part, is due to GE analysis software using a single pixel value to generate the arterial time–density curve, whereas the Siemens analysis software uses the mean value from an ROI generated by the user on a temporal maximum intensity projection image, resulting in higher mean values for GE scanners.

The full width half maximum of the initial arterial enhancement peak was generally between 8 and 12 seconds, indicating a good-quality bolus injection, with the exception of one site (PC04) where this was > 12 seconds in 50% of participants.

Follow-up and recurrences

Median, IQR and range of follow-up visits, including clinic visits, CT imaging, and colonoscopy are summarised in [Appendix 1, Table 38](#).

There were 81 events during the 3-year follow-up period, 31 (39%) in year 1, 29 (36%) in year 2 and 21 (25%) in year 3; 52 (64%) participants developed metastatic disease and 12 (14%) developed new primary tumours. In 17 (22%) participants, death occurred as the first event ([Table 7](#)). The location of recurrences is summarised in [Appendix 1, Table 39](#).

A total of 34 deaths (including the 17 deaths as first event) occurred in the 3-year follow-up period, 8 (24%) in year 1, 14 (41%) in year 2 and 12 (35%) in year 3. There were no serious adverse events related to CT scanning reported.

Carcinoembryonic antigen levels during follow-up

Serum CEA levels are summarised in [Appendix 1, Table 40](#).

TABLE 7 Summary of events during 3-year follow-up period

Recurrence	Year 1	Year 2	Year 3	Total
Type	N (%)	N (%)	N (%)	N (%)
Distant metastases	20 (25)	21 (26)	11 (13)	52 (64)
New primary	5 (6)	2 (2)	5 (6)	12 (14)
Death (first event)	6 (8)	6 (8)	5 (6)	17 (22)
Total	31 (39)	29 (36)	21 (25)	81 (100)

Chapter 4 Prognostic modelling

Introduction

Multivariable prognostic models combine multiple prognostic factors to estimate the risk of a future outcome(s) in individuals. Models aim to inform clinical decisions and by considering multiple sources of information simultaneously, aim to facilitate a more personalised approach to clinical management.⁶⁸ Prognostic models are typically developed using a multivariable regression framework. This generates an equation that estimates an individual's expected outcome value or outcome risk. The equation combines weighted values from multiple prognostic factors (e.g. age, sex, tumour stage, imaging, genetic information).

To date, multivariable models for colorectal cancer have not been used widely. This may be related to significant study heterogeneity, lack of internal and/or external model evaluation (validation), or a perception that they ignore potentially important novel biomarkers.^{23,24} Most models have been developed for metastatic colorectal cancer, predominantly metastatic liver disease. Of those addressing stage I to III cancers, models have either been developed retrospectively with single-centre data⁶⁹⁻⁷² or used prospectively collected data from randomised controlled trials, with development of web-based calculators (e.g. NUMERACY, ACCENT).^{11,24}

With advances in imaging and a better understanding of cancer biology, there is an opportunity to assess the contribution of novel biomarkers to a prognostic model for colorectal cancer. Accordingly, we prospectively developed models to predict recurrence following primary colorectal cancer from prospective multicentre data and then estimated the additive effect of incorporating novel factors from perfusion CT imaging and pathology, including immunohistochemistry and genetic biomarkers.

Methods

Participants

All participants were included in the analysis, equivalent to an intention to treat analysis.

Model development and assessment

The following models were developed and assessed ([Figure 5](#)):

- **Model A**, comprising generally accepted standard clinicopathological candidate variables, including demographic, tumour and treatment variables. Model A was then nested within the following models:
- **Model B**, comprising model A plus local hospital perfusion CT variables derived from PCA (and comprising the primary outcome for the PROSPECT trial).
- **Model D**, comprising model A plus the simplest single local hospital perfusion CT variable.
- **Model E**, comprising model A plus central review perfusion CT variables derived from PCA.
- **Model F**, comprising model A plus additional pathology variables.

Methods used for model development and assessment have been described previously in [Chapter 2](#). Model performance was compared with current standard clinical practice (i.e. AJCC staging, rule C – 'clinical'), which defines participants at high risk for recurrence as participants with stage III disease (i.e. those with node-positive cancers and participants with stage I and II disease as low risk (AJCC stage IV disease is defined by metastasis and such patients were excluded by definition).

Outcomes

Our primary outcome was comparison of prognostic model B for 3-year disease-free survival to standard AJCC staging (i.e. standard clinical variable baseline model with CT perfusion variables).

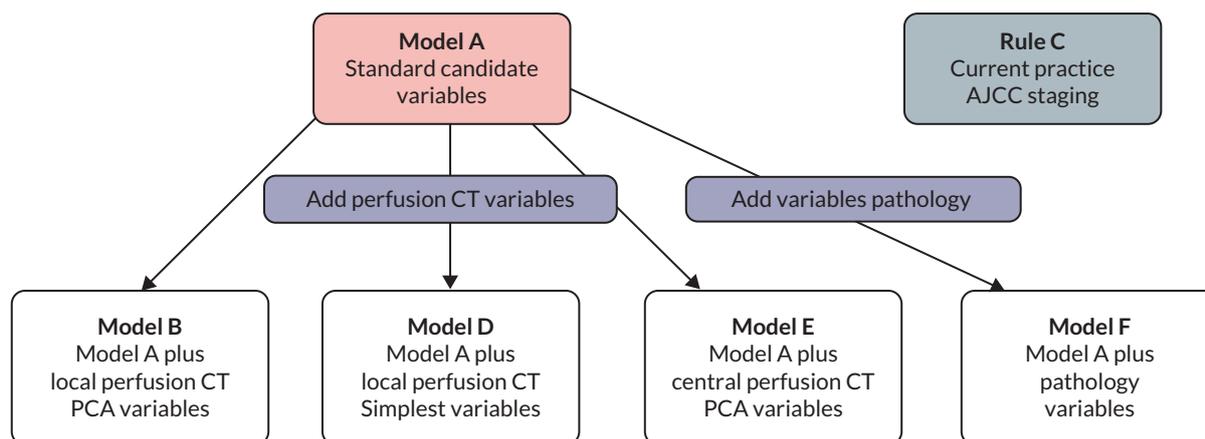


FIGURE 5 Figure showing analysis models. Arrows indicate that model A is nested within models B, D, E and F.

Secondary outcomes for prognostic modelling were: (1) prognostic model A (standard clinical variables) compared with AJCC staging; (2) added value (if any) when perfusion CT variables were added to model A; (3) added value (if any) when novel pathology variables were added to model A; (4) added value (if any) when alternative scores for perfusion CT were added to model A; (5) added value (if any) of perfusion CT to model A, based on central review data; (6) comparisons of model performance at different relative clinical weightings.

We reported our findings according to TRIPOD recommendations.⁷³

Results

Current practice (American Joint Committee on Cancer staging: rule C)

The Kaplan–Meier curve for low- (stage I/II) versus high-risk (stage III) participants is shown in [Figure 6](#).

The scatter plot shown in [Figure 7](#) provides a visual representation of recurrences and time to recurrence with respect to the low- and high-risk groupings.

Sensitivity and specificity (based on Nelson–Aalen estimates at 3 years) for predicting recurrence are summarised in [Table 8](#). With high-risk participants defined by stage III disease, sensitivity for recurrence was 0.56 (95% CI 0.44 to 0.67) while specificity was 0.58 (95% CI 0.51 to 0.64).

Model A (standard candidate variables)

The distribution of the candidate variables included in model A with respect to time to recurrence are shown in [Figure 8](#).

Model A: model development

Univariable analysis

The univariable hazard ratios for the prespecified variables are summarised in [Table 9](#). Higher T stage, right-sided location, presence of venous invasion, no surgery and older age were associated with a higher risk of recurrence.

Multivariable analysis

The adjusted hazard ratios for the prespecified variables are summarised in [Table 10](#). Again, higher T stage, presence of venous invasion, no surgery and older age were associated with a higher risk of recurrence.

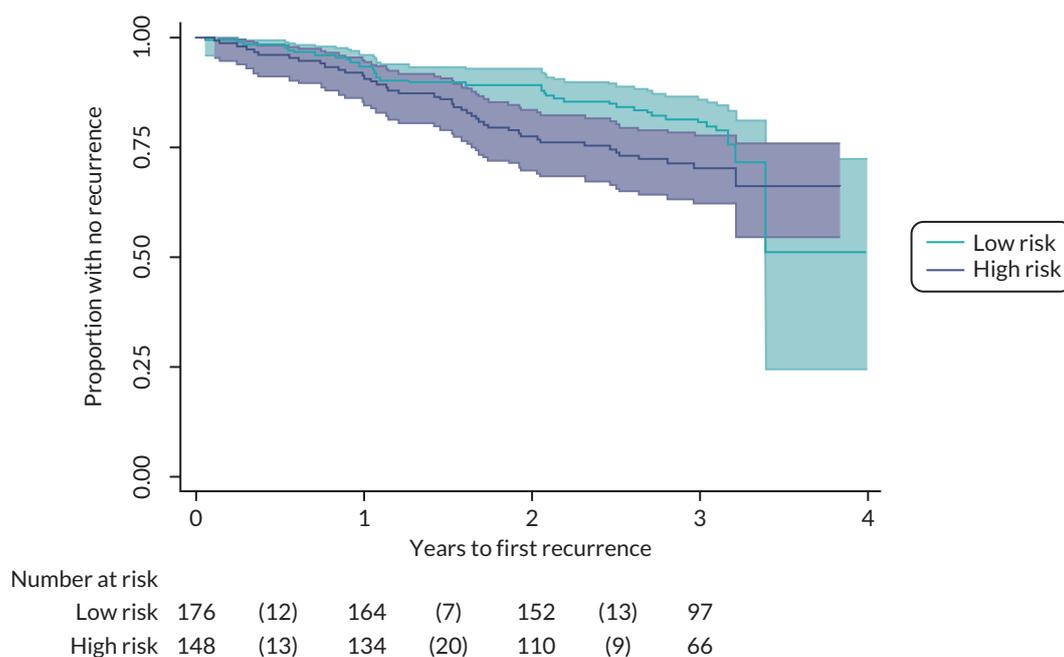


FIGURE 6 Kaplan-Meier curve for high- and low-risk patients as defined by AJCC stage group. Note that the study end was 3 years. Data beyond 3 years will be sparse and should not be overinterpreted.

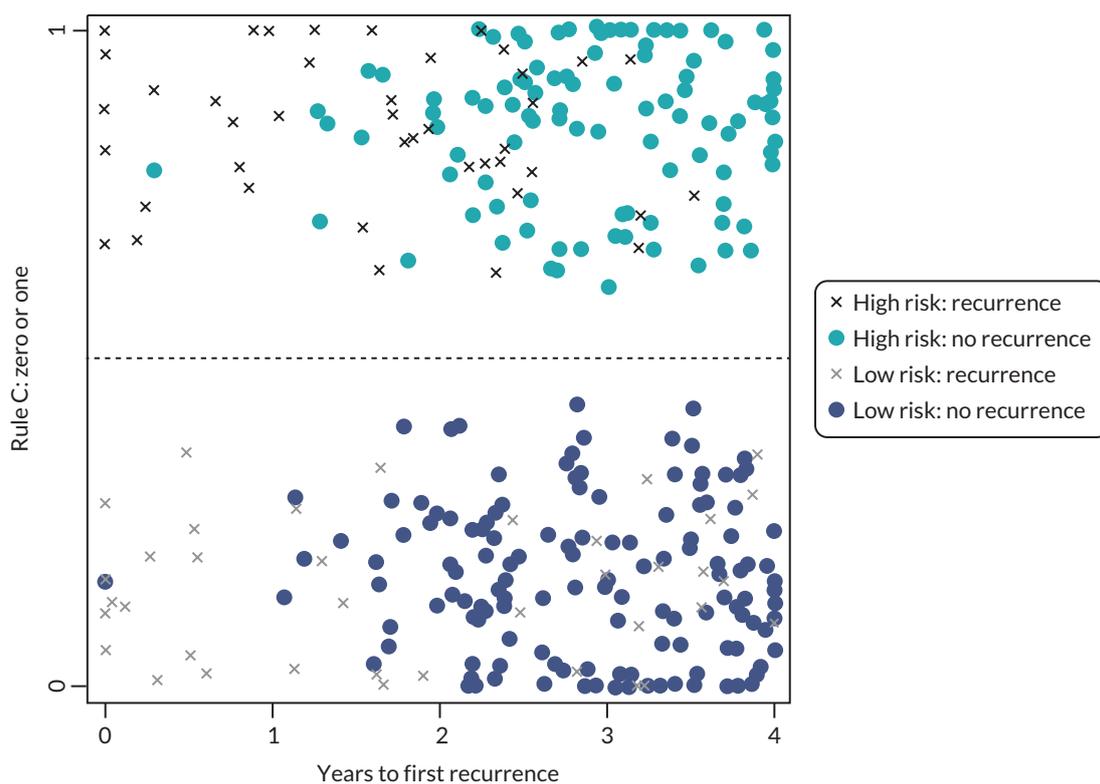


FIGURE 7 Scatter plot showing time to recurrence with respect to risk groupings defined by clinical stage.

Model A: model performance

Kaplan-Meier plots

Figure 9 summarises the Kaplan-Meier plots for the three different model A risk groupings defined by the prediction index. The high-risk group consisted of the 33% of participants with the highest prediction.

TABLE 8 Performance of AJCC staging for predicting recurrence

Model	Risk group	Survival (95% CI)	Total	Recurrence	No recurrence	Sensitivity (95% CI) (n/N)	Specificity (95% CI) (n/N)
Rule C, group, 2 : 2 risk groups	High risk	0.71 ^a (0.62 to 0.77)	148	45	103	0.56 (0.44 to 0.67) (45/81)	0.58 (0.51 to 0.64) (140/243)
	Low risk	0.81 ^a (0.74 to s0.86)	176	36	140		
	Missing		2	0	2		
	Total		326	81	245		

^a Survival calculated from Nelson–Aalen cumulative hazard at 3 years, as these estimates allow for censored data due to loss to follow-up. Survival probability was multiplied for both high- and low-risk groups to standardise to a common value of 81 recurrences multiplied by 0.985.

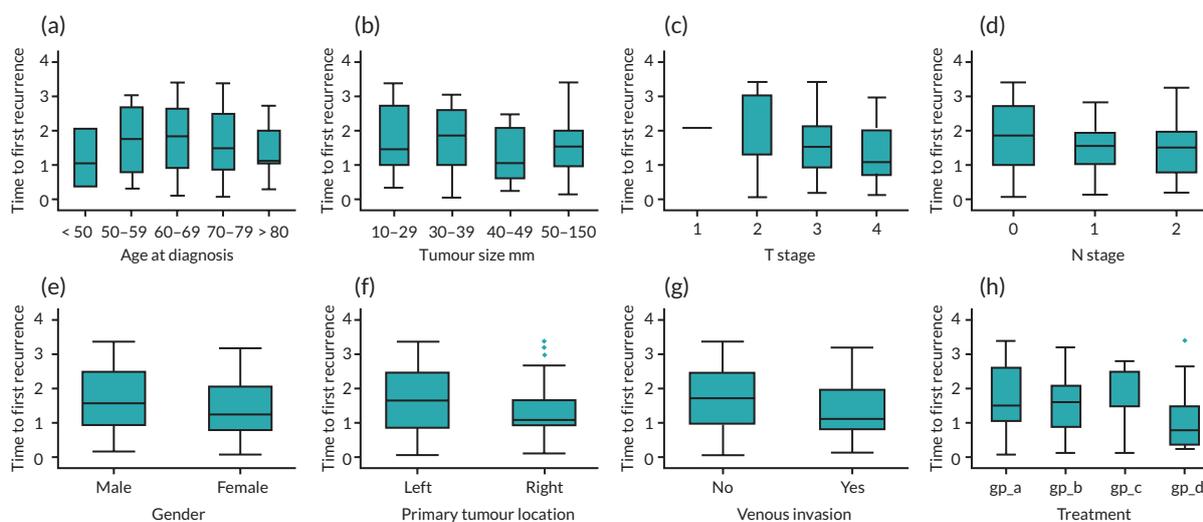


FIGURE 8 Box plots showing distribution with respect to time to recurrence (in years) for the following candidate variables: age (a); tumour size (b); tumour [T] stage (c); nodal [N] stage (d); sex (e); tumour location as left or right colon (f); venous invasion (g); and treatment groups (h), where (gp_a, surgery only; gp_b, surgery and adjuvant therapy; gp_c, neoadjuvant therapy and surgery; gp_d, no surgery).

Scatter plots

The scatter plots shown in [Figures 10](#) and [11](#) provide a visual representation of recurrences and time to recurrence with respect to the different model A risk groupings. Combining high- and medium-risk groups would increase the number of true positives but at the expense of a higher number of false positives.

C-index

[Table 11](#) summarises the results from assessment of model performance with internal validation. There was no difference between the development and internal validation model, demonstrating that the model was not overfitted (no change in c-index using tertile risk groups).

Model B (model A plus local perfusion computed tomography variables)

Model variables

Standard candidate variables for participants included/excluded in model B by recurrence event are summarised in [Appendix 1, Table 41](#).

Box plots demonstrating distribution of perfusion CT variables by presence or absence of recurrence are shown in [Figure 12](#). There was no significant difference between groups.

TABLE 9 Univariable hazard ratios (HR) for model variables assessed

Variable	N ^a	HR (95% CI)	p-value ^b
Tumour (T) stage (Reference group: T2)			
T1	324	0.61 (0.08 to 4.65)	< 0.001
T3		2.01 (1.06 to 3.79)	
T4		5.48 (2.72 to 11.03)	
Nodal (N) stage (Reference group: N0)			
N1 (1 to 3 LN positive)	324	1.22 (0.74 to 2.03)	0.11
N2 (4 or more LN positive)		1.82 (1.04 to 3.19)	
Treatment group (Reference group: surgery only)			
Surgery and adjuvant therapy	324	1.11 (0.65 to 1.91)	< 0.001
Neoadjuvant therapy and surgery	324	1.03 (0.56 to 1.89)	
No surgery	324	6.11 (3.14 to 11.88)	
Tumour location			
Rectal (Reference group: colon)	324	0.95 (0.61 to 1.47)	0.81
Right colon (Reference group: left colon)	324	1.81 (1.10 to 2.98)	0.02
Sex			
Reference group: male	324	1.01 (0.63 to 1.62)	0.96
Tumour size (mm)	324	1.00 (0.99 to 1.01)	0.96
Venous invasion	324	2.37 (1.53 to 3.67)	<0.001
Age (years)	324	1.06 (1.03 to 1.08)	<0.001

a Two patients excluded: no event time.

b Overall Wald test for categorical data of more than two categories.

Associations between perfusion CT variables are summarised in [Appendix 1, Table 42](#) and show correlation coefficients below 0.50 between blood flow, volume, permeability and transit time.

Model B: model development

Univariable analysis

The univariable hazard ratios for perfusion CT variables from (1) local hospital review and (2) PCA are summarised in [Table 12](#). None of the measurements were associated with a higher risk of recurrence.

Multivariable analysis

The multivariable model is summarised in [Table 13](#). Addition of perfusion CT variables as PCA scores did not improve prediction of model A significantly.

Model B: model performance

Kaplan–Meier plots

The Kaplan–Meier plots for the different model B risk groupings shown in [Appendix 2, Figure 22](#) have a similar distribution to model A indicating that the addition of perfusion CT variables as PCA scores to the baseline model did not benefit prediction significantly.

TABLE 10 Multivariable-adjusted hazard ratios for model variables assessed

Variable	N ^a	Hazard ratios (95% CI)	p-value ^b
T stage (Reference group: T2)			
T1	324	0.84 (0.11 to 6.52)	< 0.001
T3	324	1.83 (0.90 to 3.70)	
T4	324	5.70 (2.52 to 12.84)	
N stage (Reference group: N0)			
N1 (1–3 positive node)	324	1.10 (0.60 to 2.01)	0.69
N2 (4 or more positive nodes)	324	1.32 (0.70 to 2.51)	
Treatment group (Reference group: surgery only)			
Surgery and adjuvant therapy	324	0.90 (0.42 to 1.92)	
Neoadjuvant therapy and surgery	324	0.68 (0.35 to 1.30)	
No surgery	324	4.87 (2.37 to 10.0)	< 0.001
Location (Reference group: left colon)			
Right colon location	324	1.52 (0.88 to 2.65)	0.14
Tumour size (mm)^c	324	0.99 (0.98 to 1.01)	0.19
Venous invasion	324	1.93 (1.17 to 3.19)	0.01
Age (years)^c	324	1.05 (1.02 to 1.07)	< 0.001
Sex (Reference group: male)	324	0.77 (0.47 to 1.26)	0.30
Rcs1	324	2.300 (1.951 to 2.711)	
Constant	324	0.073 (0.038 to 0.140)	

a Overall Wald test for categorical data of more than two categories;
 b Two patients excluded: no event time.
 c Age is centred at 65 years and tumour size is centred at 40mm. Constant term when Model A is not centred is 0.006.

Scatter plots

The scatter plot of prediction index for the different risk groupings by year to recurrence is shown in [Appendix 2, Figures 23 and 24](#).

The scatter plot of the prediction indices with and without perfusion CT variables confirmed that there was no substantial difference between model B and model A (i.e. the CT variables contributed nothing useful to prediction; [Figure 13](#)).

Model D (model A plus simplest local perfusion computed tomography variable)

The scatter plots of prediction indices with and without the simplest perfusion CT variables are shown in [Appendix 2, Figure 25](#), again indicating no substantial difference between model D and model A.

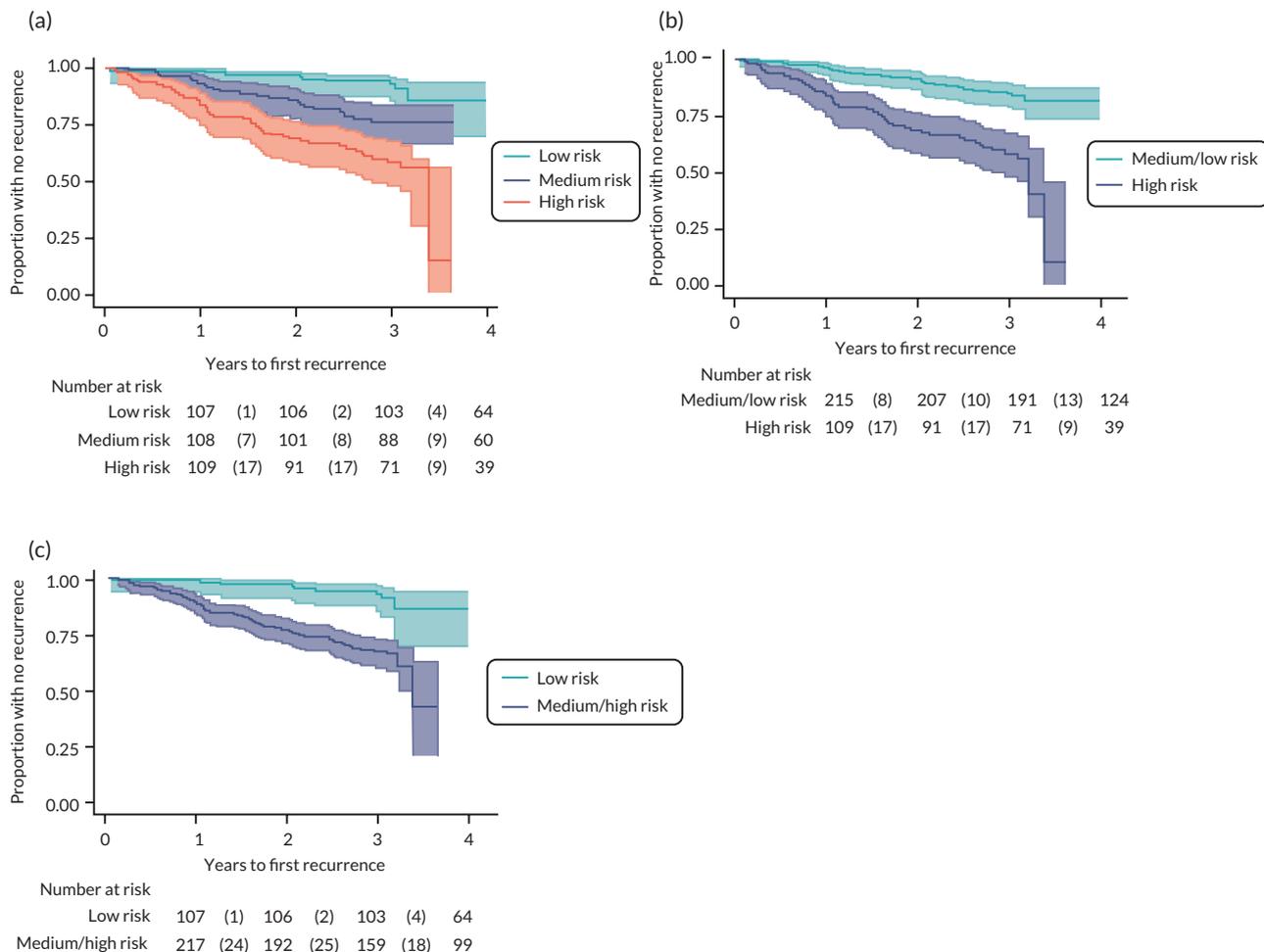


FIGURE 9 Kaplan-Meier plots for the different risk groupings: (a) all three risk groups; (b) two risk groups, low/medium vs. high; and (c) low vs. medium/high groups. Note: study end was 3 years; data beyond 3 years will be sparse and should not be overinterpreted.

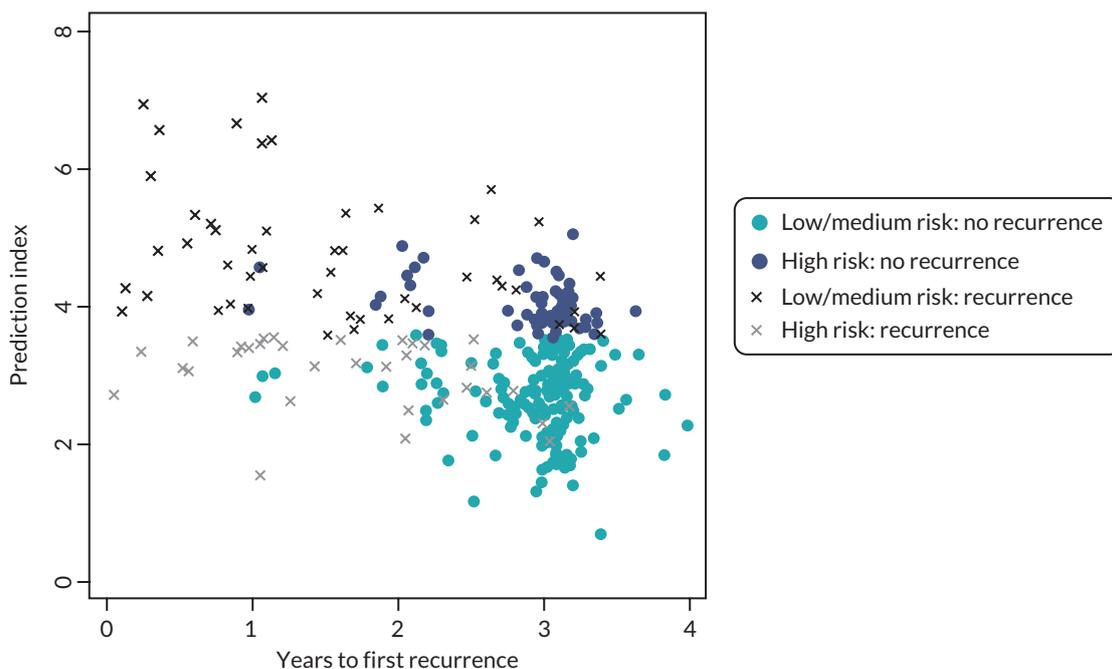


FIGURE 10 Scatter plot showing time to recurrence with respect to two risk groupings defined by model A prediction index (low/medium risk vs. high risk).

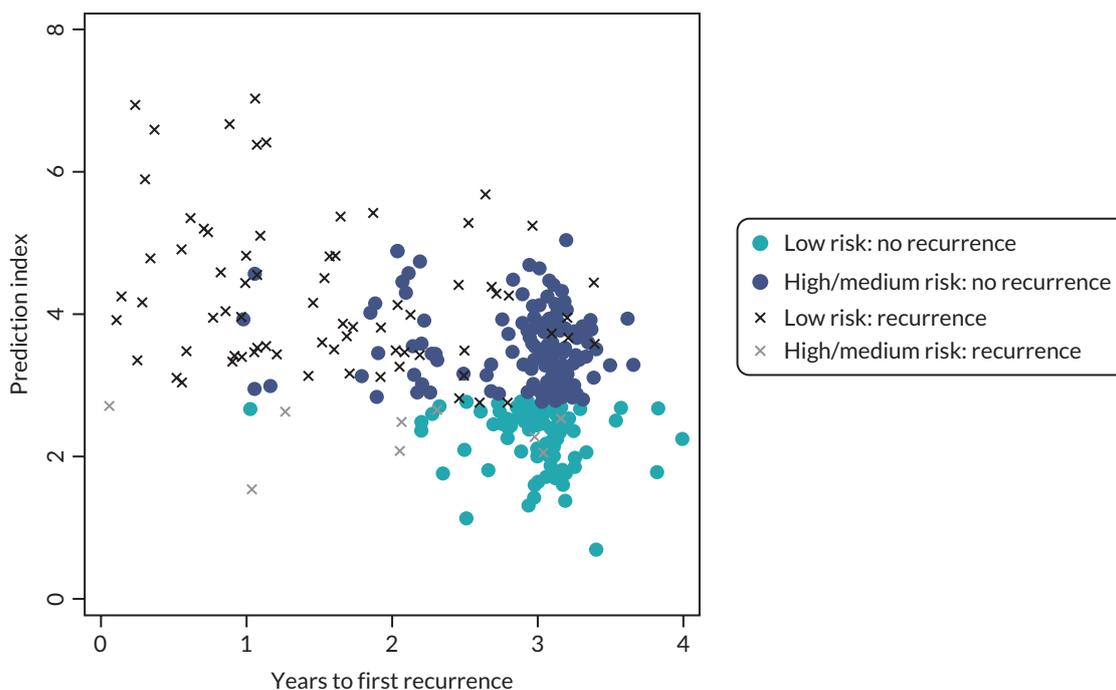


FIGURE 11 Scatter plot showing time to recurrence with respect to two risk groupings defined by model A prediction index (low vs. medium/high risk).

TABLE 11 C-index including internal validation and optimism adjustment

Model A (N = 324)	Original apparent (95% CI)	Optimism (95% CI)	Optimism adjusted (N = 100 bootstrap) (95% CI)
c-statistic (Harrell's)	0.75 (0.70 to 0.80)	0.000 (-0.002 to 0.002)	0.75 (0.70 to 0.80)
R ²	0.43 (0.31 to 0.54)		
Adjusted R ²	0.35 (0.22 to 0.47)	0.00	0.35 (0.26 to 0.56)
Calibration slope	1.00 (0.80 to 1.20)	0.001 (-0.001 to 0.000)	1.00 (0.99 to 1.00)

Model E (model A plus central perfusion computed tomography variables)

Central perfusion computed tomography variables

Appendix 2, Figure 26 shows that the distribution was similar for patients with and without recurrence for all four perfusion CT variables.

Model E: model development

Univariable analysis

The univariable hazard ratios for (1) individual perfusion CT variables from central review and, (2) PCA based on these variables are summarised in Appendix 1, Table 43. None of the measurements were significantly associated with a higher risk of recurrence.

Multivariable analysis

The multivariable model, with fitting of model A variables with the same coefficients, is summarised in Appendix 1, Table 44. Addition of central perfusion CT variables (as PCA scores) to the baseline clinical model did not improve prediction accuracy significantly.

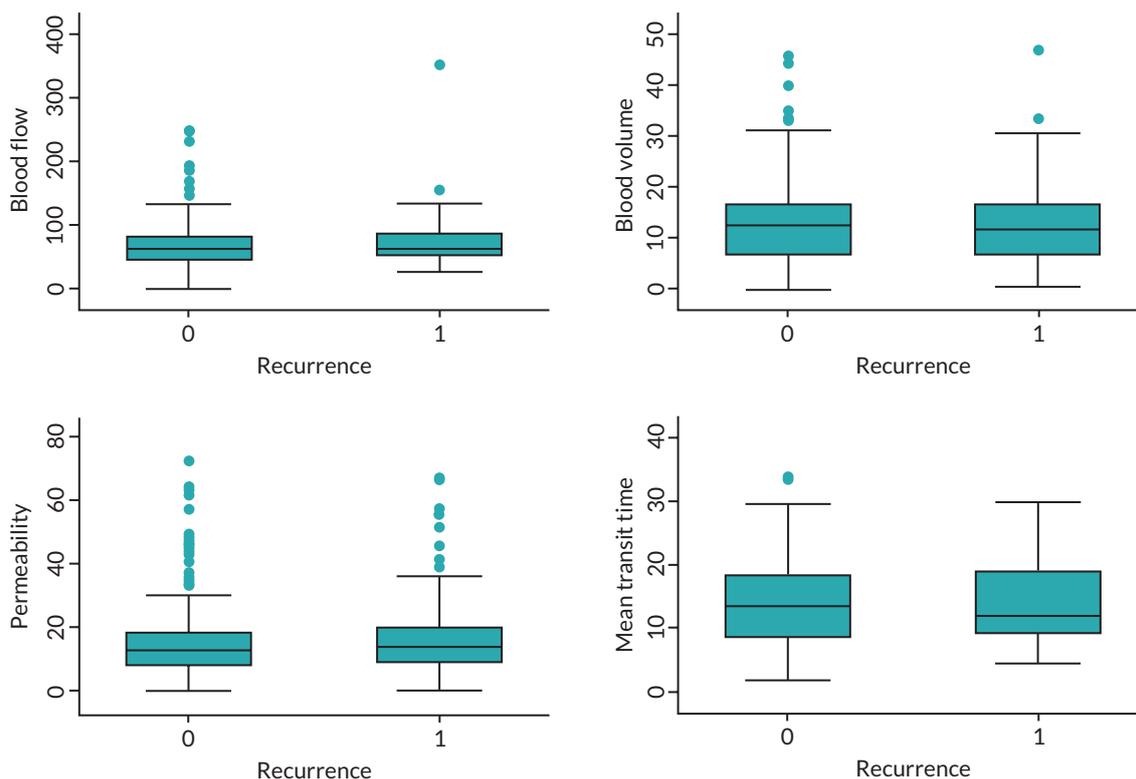


FIGURE 12 Box and whisker plots of perfusion CT variables by recurrence group (1 = recurrence; 0 = no recurrence).

TABLE 12 Univariable hazard ratios for local hospital perfusion CT variables

Variable	N	Hazard ratio (95% CI)	p-value
Blood volume	311	1.00 (0.98 to 1.03)	0.83
Blood flow	313	1.00 (0.99 to 1.01)	0.50
Permeability	277	1.01 (0.99 to 1.02)	0.44
Mean transit time	277	1.00 (0.96 to 1.05)	0.97
PCA local variables			
PCA1	277	1.05 (0.88 to 1.24)	0.60
PCA2	277	1.03 (0.86 to 1.25)	0.72

Note

PCA1 and PCA2 represent first and second eigenvalue from PCA.

TABLE 13 Multivariable-adjusted hazard ratios for model B with model A for reference

Models	Model A (clinical)				Model B (clinical + perfusion CT)			
	HR	HR, 95% CI lower	HR, 95% CI upper	p-value	HR	HR, 95% CI lower	HR, 95% CI upper	p-value
Model A_324	1.000				1.000			
PCA1 score					0.006	-0.154	0.167	0.94
PCA2 score					0.038	-0.143	0.219	0.68
_rcs1 ^a	0.723	0.562	0.885	< 0.001	0.724	0.562	0.886	< 0.001
Constant ^a	-0.359	-0.599	-0.119	0.003	-0.361	-0.602	-0.120	0.003
Log likelihood	-238.24				-238.15			
N in model	277				277			

^a _rcs (baseline restricted spline coefficient) and constant terms of the baseline survival function of the flexible parametric model at df 1.

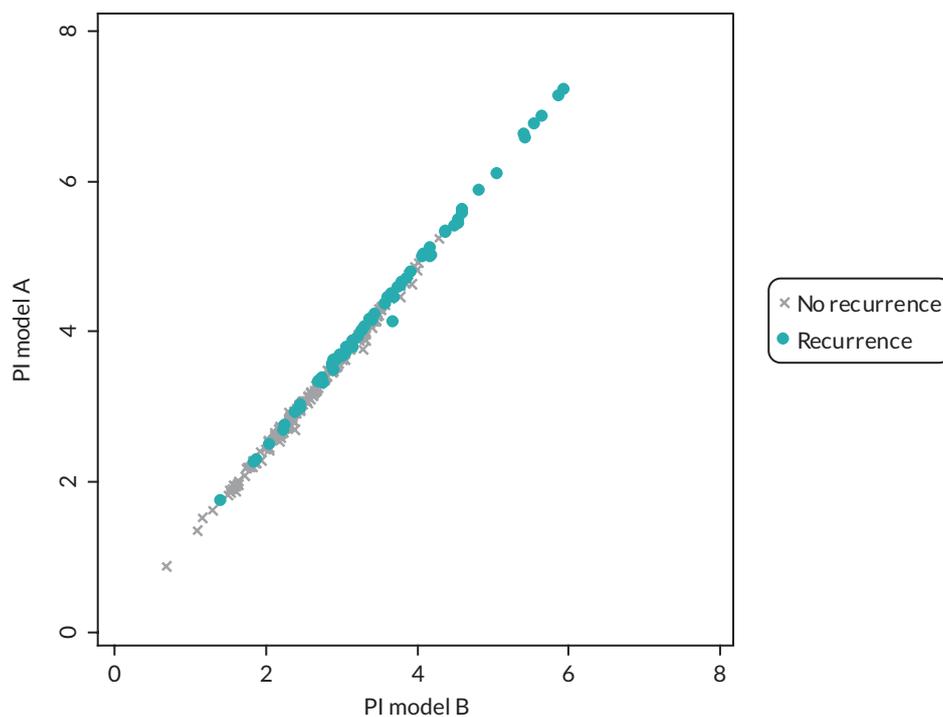


FIGURE 13 Scatter plot of prediction index (PI) for model A and model B.

Model E: model performance

Kaplan–Meier plot

The Kaplan–Meier plots for the model E risk groupings had a similar distribution to model A and model B, indicating that addition of central review perfusion CT variables (as PCA scores) to the clinical variables did not improve prediction accuracy significantly. [Appendix 2, Figure 27](#) illustrates this for low-/medium- versus high-risk groupings.

Scatter plots

The scatter plot showing time to recurrence with respect to the risk groupings defined by model E (see [Appendix 2, Figure 28](#)) was similar to that achieved by model A or B.

The scatter plots of prediction indices with and without central perfusion CT variables are shown in [Appendix 2, Figure 29](#), indicating no substantial difference between models E and A.

Model F (model A plus additional pathology variables)

Participants

Standard candidate variables for participants included/excluded in model F by recurrence event are summarised in [Appendix 1, Table 45](#).

Model F: model development

Univariable analysis

The univariable hazard ratios for the additional novel pathology biomarkers are summarised in [Table 14](#). None of the measurements were associated with a significantly higher risk of recurrence.

Multivariable analysis

The multivariable hazard ratios for the novel pathology biomarkers modelled are summarised in [Table 15](#). In the multivariable analysis, none of the variables had a statistically significant relationship with recurrence.

Multivariable hazard ratios for somatic mutation variables alone are summarised in [Appendix 1, Table 46](#). When somatic mutation variables alone were included, KRAS mutation was statistically significant, although interpretation is constrained by the small numbers with KRAS mutations.

Model performance

Kaplan–Meier plots

The Kaplan–Meier plots for all pathology variables and somatic mutation analysis variables alone are shown in [Appendix 2, Figures 30](#) and [31](#), respectively. The plots for model F had a similar distribution to models A and B, indicating that the addition of novel pathology biomarkers did not change prediction accuracy significantly.

TABLE 14 Univariable hazard ratios for additional pathology variables

Variable	N	Hazard ratio (95% CI)	p-value
CD105 ^a	253	1.000 (0.996 to 1.004)	0.83
HIF-1a ^b	253	1.022 (0.610 to 1.711)	0.94
MMR ^c	253	0.845 (0.338 to 2.114)	0.72
BRAF_ other	253	0.670 (0.304 to 1.475)	0.32
BRAF 600	253	0.689 (0.215 to 2.201)	0.53
KRAS	253	0.708 (0.422 to 1.187)	0.19
HRAS	253	1.338 (0.634 to 2.822)	0.44
NRAS	253	1.326 (0.602 to 2.922)	0.48

a Continuous variable.
b Binary measure: 0 = 0–2, 1 = 3–6.
c MMR-deficient is the reference standard.

TABLE 15 Multivariable-adjusted hazard ratios for all modelled pathology variables

Variables	N	Hazard ratio (95% CI)	p-value
Model A	212	2.718 constrained	
CD105 ^a	212	1.001 (0.996 to 1.005)	0.56
HIF-1a ^b	212	1.329 (0.741 to 2.386)	0.34
MMR ^c	212	0.635 (0.239 to 1.683)	0.36
BRAF_other	212	0.566 (0.229 to 1.398)	0.22
BRAF 600	212	0.338 (0.076 to 1.504)	0.16
KRAS	212	0.570 (0.306 to 1.060)	0.08
HRAS	212	1.996 (0.875 to 4.552)	0.10
NRAS	212	2.321 (0.931 to 5.781)	0.07
Baseline survival (rcs1)	212	2.190 (1.785 to 2.685)	< 0.001
Constant	212	0.014 (0.004 to 0.046)	< 0.001

a Continuous variable.
b Binary measure: 0 = 0–2, 1 = 3–6.
c MMR deficient is the reference standard.

Scatter plots

The scatter plots in [Figure 14](#) and [Appendix 2, Figure 32](#) show no substantial improvement when compared with model A.

Dotted lines show the thresholds for high-risk group.

Sensitivities and specificities

Primary outcome: model B versus American Joint Committee on Cancer staging (rule C)

Sensitivity and specificity, based on top tertile group, to predict recurrence

Sensitivity and specificity based on Nelson–Aalen estimates at 3 years for the model including standard candidate plus perfusion CT variables are summarised in [Table 16](#) and compared with AJCC staging.

There was a non-significant 5% increase in sensitivity (95% CI –9% to 18%) and statistically significant 18% increase in specificity (9% to 27%) when model B was compared with AJCC staging (see [Appendix 1, Table 47](#)).

Sensitivity and specificity, based on top two tertiles, to predict recurrence

Sensitivity and specificity based on Nelson–Aalen estimates at 3 years for the best model including standard candidate plus perfusion CT variables at a different threshold are summarised in [Appendix 1, Table 48](#) and compared with AJCC staging.

Using this alternative threshold, sensitivity increased significantly by 36% (95% CI 21% to 50%) but at the cost of a significant 16% decrease in specificity (95% CI 6% to 26% decrease) (see [Appendix 1, Table 49](#)).

Secondary outcome: model A versus American Joint Committee on Cancer staging (rule C)

Sensitivity and specificity, based on top tertile group, to predict recurrence

Sensitivity and specificity based on Nelson–Aalen estimates at 3 years for the model with standard clinical variables are summarised in [Table 17](#) and compared with current clinical practice.

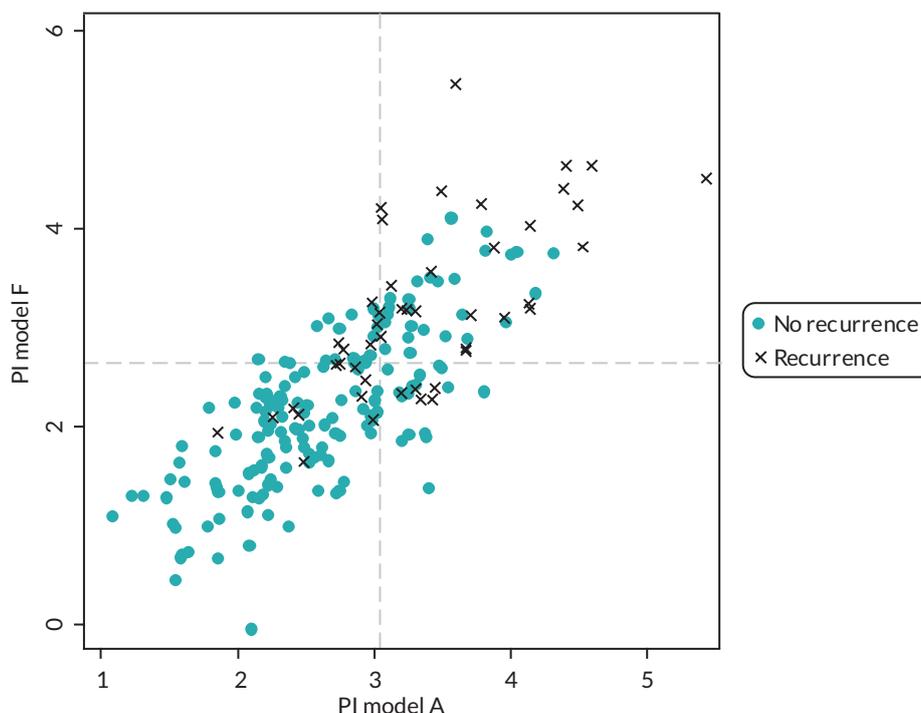


FIGURE 14 Scatter plot of prediction index (PI) for model A and model F3 (all additional pathology variables).

TABLE 16 Sensitivity and specificity analysis for model B compared with rule C

Model	Risk group	Survival (95% CI)	Total (n)	Recurrence (n)	No recurrence (n)	Sensitivity		Specificity	
						(95% CI)	(n/N)	(95% CI)	(n/N)
Model B, group 2 : 2 risk groups	High risk (top 33%)	0.573 ^a (0.462 to 0.670)	93	41	52	0.59 (0.46 to 0.70)	(41/70)	0.75 (0.68 to 0.81)	(155/207)
	Medium/low risk	0.857 ^a (0.795 to 0.901)	184	29	155				
	Total model B		277	70	207				
	Missing PCT		49	11	36				
	Total			326	81	247			
Rule C, group 2 : 2 risk groups	High risk		127	38	89	0.54 (0.42 to 0.66)	(38/70)	0.57 (0.50 to 0.64)	(118/207)
	Low risk		150	32	118				
	Total model B		277	70	207				
	Missing PCT		49	11	36				
	Total			326	81	247			

a Survival calculated from Nelson–Aalen cumulative hazard at 3 years, as these estimates allow for censored data due to loss to follow-up. Survival probability multiplied for both high- and low-risk groups to standardise to a common value of 81 recurrences (a) multiplied by 0.985 (b) multiplied by 0.985. PCT, perfusion CT.

TABLE 17 Sensitivity and specificity analysis for model A compared with rule C

Model	Risk group	Survival (95% CI survival)	Total (n)	Recurrence (n)	No recurrence (n)	Sensitivity		Specificity	
						(95% CI)	(n/N)	(95% CI)	(n/N)
Model A, group 2 : 2 risk groups	High risk (top 33%)	0.589 ^a (0.488 to 0.678)	109	46	63	0.57 (0.45 to 0.68)	(46/81)	0.74 (0.68 to 0.79)	(180/243)
	Medium/low risk	0.859 ^a (0.792 to 0.892)	215	35	180				
	Missing		2	0	2				
	Total		326	81	245				
Rule C, group 2 : 2 risk groups	High risk	0.706 ^a (0.624 to 0.774)	148	45	103	0.56 (0.44 to 0.67)	(45/81)	0.58 (0.51 to 0.64)	(140/243)
	Low risk	0.809 ^a (0.740 to 0.861)	176	36	140				
	Missing		2	0	2				
	Total		326	81	245				

a Survival calculated from Nelson–Aalen cumulative hazard at 3 years, as these estimates allow for censored data due to loss to follow-up. Survival probability multiplied for both high- and low-risk groups to standardise to a common value of 81 recurrences multiplied by 0.985.

The difference in sensitivity and specificity is summarised in [Table 18](#). There was a non-significant 1% increase in sensitivity (95% CI -15% to 17%) and statistically significant 16% increase in specificity (95% CI 8% to 25%) with model A compared with AJCC staging.

Sensitivity and specificity, based on top two tertile groups to predict recurrence

Sensitivity and specificity based on Nelson–Aalen estimates at 3 years for the model with standard clinical variables at a different threshold are summarised in [Appendix 1, Table 50](#) and compared with current clinical practice.

The difference in sensitivity and specificity is summarised in [Appendix 1, Table 51](#). There was a 33% increase (95% CI 21% to 46%) in sensitivity at the expense of an 18% decrease (95% CI 8% to 26% decrease) in specificity.

Secondary outcome: model B versus model A

Sensitivity and specificity based on Nelson–Aalen estimates at 3 years for the model with both standard candidate plus perfusion CT variables (model B) are compared with the model with standard candidate variables only (model A) in [Appendix 1, Table 52](#). There was no difference in sensitivity and specificity between these two models (see [Appendix 1, Table 53](#)) confirming that inclusion of CT perfusion variables did not improve model performance.

Secondary outcome: model D versus model B

Sensitivity and specificity for model D (simplest perfusion CT score) and for model B (with standard and perfusion CT variables) are compared in [Appendix 1, Table 54](#). Again, there was no difference in sensitivity and specificity between these two models.

Secondary outcome: model E versus model A

Sensitivity and specificity for model E (central perfusion CT score) and for model A (standard candidate variables) are compared in [Appendix 1, Table 55](#). Again, there was no substantial difference in sensitivity (3% increase with 95% CI from -4% to 9%) and specificity (no change) between these two models ([Appendix 1, Table 56](#)).

Secondary outcome: model F versus model A

Sensitivity and specificity based on Nelson–Aalen estimates at 3 years for model F (where all pathological variables are presented as a single PCA score) and model A (the baseline model with standard variables) are summarised in [Appendix 1, Table 57](#). Again, there was no substantial difference in sensitivity (8% increase, 95% CI -9% to 26%) and specificity (3% increase, 95% CI -5% to 10%) when the two models were compared ([Table 19](#)).

Secondary outcome: impact of weighting on best model performance

The impact of different relative weightings on model performance is shown in [Figure 15](#). Different trade-offs of false positives for each true positive were plotted for model A at two different risk thresholds, and for AJCC staging alone. Here, there was a higher net benefit for model A compared with AJCC staging alone, of up to 10 false-positive diagnoses.

Summary of model performance

Model performance is summarised in [Table 20](#) and [Figure 16](#). All models were compared in the same 212 patients to facilitate comparison of model performance. In summary, model A (baseline standard clinicopathological variables) provided the best model fit, with no improvement in model prediction when additional novel imaging, immunohistochemical or genetic variables were added.

Model A equation

Model A log cumulative hazard = $-2.617 + 0.833 \cdot \ln(\text{time in years})$

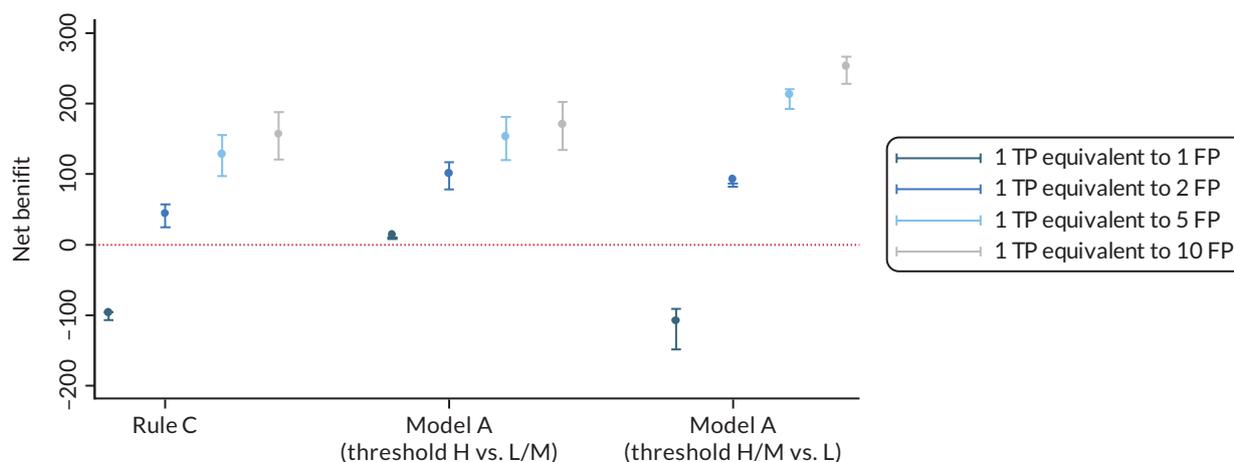
$$\begin{aligned}
 & - 0.174 \cdot T1 + 0 \cdot T2 + 0.604 \cdot T3 + 1.739 \cdot T4 + 0 \cdot N0 + 0.604 \cdot N1 + 0.279 \cdot N2 \\
 & + 0 \cdot \text{male} - 0.262 \cdot \text{female} + 0 \cdot \text{No EMV} + 0.657 \cdot \text{EMV} + 0 \cdot \text{left_colon} + 0.421 \cdot \text{right_colon} \\
 & + 0 \cdot \text{surgery_only} - 0.393 \cdot \text{Neoadjuvant_surgery} - 0.110 \cdot \text{Surgery_adjuvant} + 1.582 \cdot \text{No_surgery} \\
 & + 0 \cdot 65 \text{ years} + 0.045 \cdot \text{increased_years_above_65 years} \\
 & + 0 \cdot 40 \text{ mm_tumour_size} - 0.010 \cdot \text{each_1-mm_increased_size_over_40 mm}
 \end{aligned}$$

TABLE 18 Difference in sensitivity and specificity for model A compared with rule C

Model (<i>n</i> = 324)	Sensitivity (95% CI)	Specificity (95% CI)	True positives (of 330 with recurrence) (95% CI)	True negatives (of 670 with no recurrence) (95% CI)
Model A (high vs. medium/low)	0.57 (0.45 to 0.68)	0.74 (0.68 to 0.79)	188 (149 to 224)	496 (456 to 529)
Rule C	0.56 (0.44 to 0.67)	0.58 (0.51 to 0.64)	185 (145 to 221)	389 (342 to 429)
Difference	0.01 (-0.15 to 0.17)	0.16 (0.08 to 0.25)	3 (-50 to 56)	107 (54 to 167)

TABLE 19 Difference in sensitivity and specificity for model F compared with model A

Model, <i>n</i> = 212	Sensitivity (95% CI)	Specificity (95% CI)	True positives (of 330 with recurrence) (95% CI)	True negatives (of 670 with no recurrence) (95% CI)
Model F, high vs. medium/low	0.68 (0.53 to 0.81)	0.76 (0.68 to 0.82)	225 (175 to 267)	509 (456 to 549)
Model A, high vs. medium/low	0.60 (0.44 to 0.74)	0.73 (0.66 to 0.80)	197 (145 to 244)	489 (442 to 536)
Difference	0.08 (-0.09 to 0.26)	0.03 (-0.05 to 0.10)	26 (-30 to 86)	20 (-34 to 67)

**FIGURE 15** Plot showing net benefit gain with different true-positive/false-positive weightings for rule C (AJCC staging) and model A (based on standard candidate variables). PI, prediction index.

Note that model A used flexible parametric survival analysis based on `stmp2` in STATA 14.1. The choice of model was selected based on comparing the smoothed baseline hazard from a Cox model to parametric baselines based on 1–6 df, which included 0–5 spline knots in the baseline. Modelling using 1 df was chosen, which is a Weibull model with no spline knots, based on the lowest AIC and BIC. This was also a more credible baseline reflecting the biological rate of recurrence, as opposed to peaks that artificially reflect timing of regular yearly CT imaging when the majority of recurrences were detected. The baseline model included a constant term and a shape parameter (*rsc*), which is multiplied by the natural log of time in years.

Models were fitted on the log cumulative hazard scale $\{\ln[-\ln S(t)]\}$ with proportional hazards. Nested models of model A coefficients were fitted using a fixed offset of the model A linear predictor, allowing the effect of additional variables to be assessed. Continuous variables were centred on values close to the median: age in years was centred for 65 years and tumour size was centred at 40 mm.

TABLE 20 Summary of model performance

Model (n = 212)	Log likelihood	C-statistic (Harrell's)	R ²	Adjusted R ²	D statistic (SE)	Adjusted D statistic (SE)	Calibration slope	Model df	AIC ^a	BIC ^a	Interpretation
A: baseline clinical variables	-128.39	0.759 (0.698 to 0.821)	0.38	0.37	1.606 (0.258)	1.572 (0.259)	1.269 (0.889 to 1.649)	2	260.8	267.5	Best model fit (both AIC and BIC are lowest)
B: model A plus local PCT using 2 PCA scores	-128.37	0.768 (0.708 to 0.829)	0.38	0.35	1.614 (0.258)	1.509 (0.260)	1.277 (0.894 to 1.661)	4	264.8	278.2	No increase in model fit
E: model A plus central PCT using 2 PCA scores	-127.67	0.777 (0.718 to 0.836)	0.38	0.35	1.599 (0.253)	1.494 (0.254)	1.197 (0.844 to 1.550)	4	263.3	276.8	No increase in model fit
D: model A plus local PCT using 2 individual PCT (BF and BV)	-128.32	0.761 (0.699 to 0.822)	0.37	0.34	1.568 (0.253)	1.462 (0.255)	1.267 (0.888)	4	264.6	278.1	No increase in model fit
D2: model A plus local PCT using all 4 individual PCT (BF, BV, PS, MTT)	-128.11	0.765 (0.705 to 0.824)	0.39	0.33	1.630 (0.260)	1.451 (0.264)	1.319 (0.919 to 1.719)	6	268.2	288.4	No increase in model fit
F1: model A plus 4 main genes only	-122.87	0.766 (0.705 to 0.826)	0.47	0.41	1.937 (0.282)	1.717 (0.279)	1.251 (0.907 to 1.596)	7	259.7	283.2	No increase in model fit
F2: model A plus IHC (CD105, HIF-1a, MMR)	-127.16	0.777 (0.718 to 0.836)	0.41	0.37	1.695 (0.263)	1.553 (0.264)	1.311 (0.924 to 1.698)	5	264.3	281.1	No increase in model fit
F3: model A plus 4 main genes and IHC (CD105, HIF-1a, MMR)	-121.88	0.777 (0.718 to 0.836)	0.50	0.41	2.040 (0.296)	1.696 (0.290)	1.258 (0.917 to 1.600)	10	263.8	297.3	No increase in model fit

BF, blood flow; BV, blood volume; HIF, hypoxia-inducible factor; IHC, immunohistochemical; MTT, mean transit time; PCT, perfusion CT; PS, permeability surface area product; SE, standard error.

^a Lower AIC and BIC indicate a better model fit.

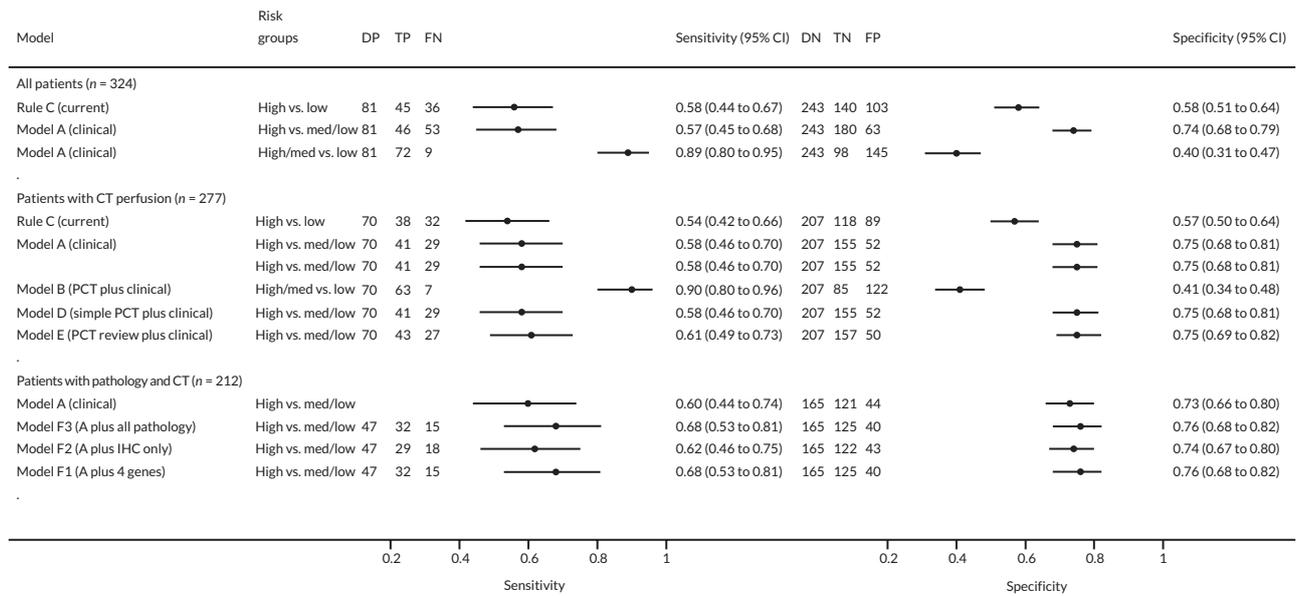


FIGURE 16 Forest plot of performance for all models. DN, disease negative; DP, disease positive; FN, false negative; FP, false positive; PCT, perfusion CT; TN, true negative; TP, true positive.

Chapter 5 Perfusion computed tomography: local versus central review

Introduction

Perfusion CT analysis relies predominantly on commercial CT software linked to the scanner acquisition protocol and is based on differing mathematic analysis methods, depending on the CT manufacturer. Applied mathematical models include deconvolution, slope, Patlak and distributed parameter models, which provide quantitative measurements including tissue blood flow, blood volume, mean transit time and capillary permeability.^{52,55}

Measurements must be accurate and reproducible, or any clinical utility will inevitably suffer. Measurements using such software are often assumed to be reliable but previous single-centre studies have identified measurement variability related to a number of factors. These may be related to the observer, mathematical model used, software version (for the same model), tumour motion and tumour coverage, in both colorectal⁷⁴⁻⁷⁸ and other cancers.⁷⁹⁻⁸²

To date, measurement variability has not been assessed in a multicentre setting across multiple different readers, who would be typical of those using the technique in day-to-day clinical practice. Here, we aimed to determine the variability of perfusion CT measurements made at local centres, compare them to those made centrally, and to estimate whether the limits of agreement are clinically acceptable.

Methods

Participants

As part of the consent process, all participants recruited to the main trial gave permission for their data sets to be used for this substudy.

Perfusion computed tomography imaging and analysis

As described previously, perfusion CT imaging was performed across 13 local hospital sites, either at the same time as the staging CT or on a different day if they could not be scheduled together. An initial low-dose abdominopelvic CT acquisition to locate the primary colorectal tumour was followed by intravenous injection of iodinated CT contrast agent (> 300 mg/ml iodine concentration; 50 ml injected at 5 ml/second). Then, a dynamic acquisition, centred on the tumour, was undertaken. These dynamic data were acquired for every 1.5 seconds for a total of 45 seconds, then every 15 seconds thereafter for an additional 75 seconds.

The perfusion CT imaging at local hospital sites was analysed and interpreted by 25 radiologists in total with 0–3 years' experience of perfusion CT analysis. All were general radiologists with a subspecialty interest in gastrointestinal imaging or subspecialty gastrointestinal radiologists.

Image analysis was undertaken using the commercial software platform used locally, provided by the manufacturer of their CT scanners (GE, Siemens, Phillips or Toshiba). Commercial software platforms were based on different kinetic analysis models, depending on the CT manufacturer and included the deconvolution, distributed parameter, slope or Patlak models.

Data sets were also transferred for review centrally, performed by three radiologists with 5–18 years' experience of perfusion CT analysis. Radiologists undertaking central review were blinded to standard imaging investigations and outcomes for individual patients. Image analysis was performed using the same commercial software that had been used locally (GE, Siemens, Phillips or Toshiba).

Image analysis was undertaken as described previously to generate the following tumour vascular parameters: regional blood flow, regional blood volume, mean transit time and permeability surface area product.

Statistical analysis

Scatter plots for regional blood flow, blood volume, mean transit time and permeability surface measurements from local and central review, respectively, were generated. Agreement between paired local and central review data was assessed using Bland–Altman statistics for each vascular parameter: regional blood flow, regional blood volume, mean transit time and permeability surface area product. The role of the size of ROI and machine type was also investigated as a factor for differences in vascular parameters for local and central review.

Results

Of the 303 participants who had undergone successful perfusion CT imaging, 291 (96%) participants were included in this analysis. Central review was not possible in 14/303 (4%) participants due to technical issues related to corrupt data storage and data transfer from the local site.

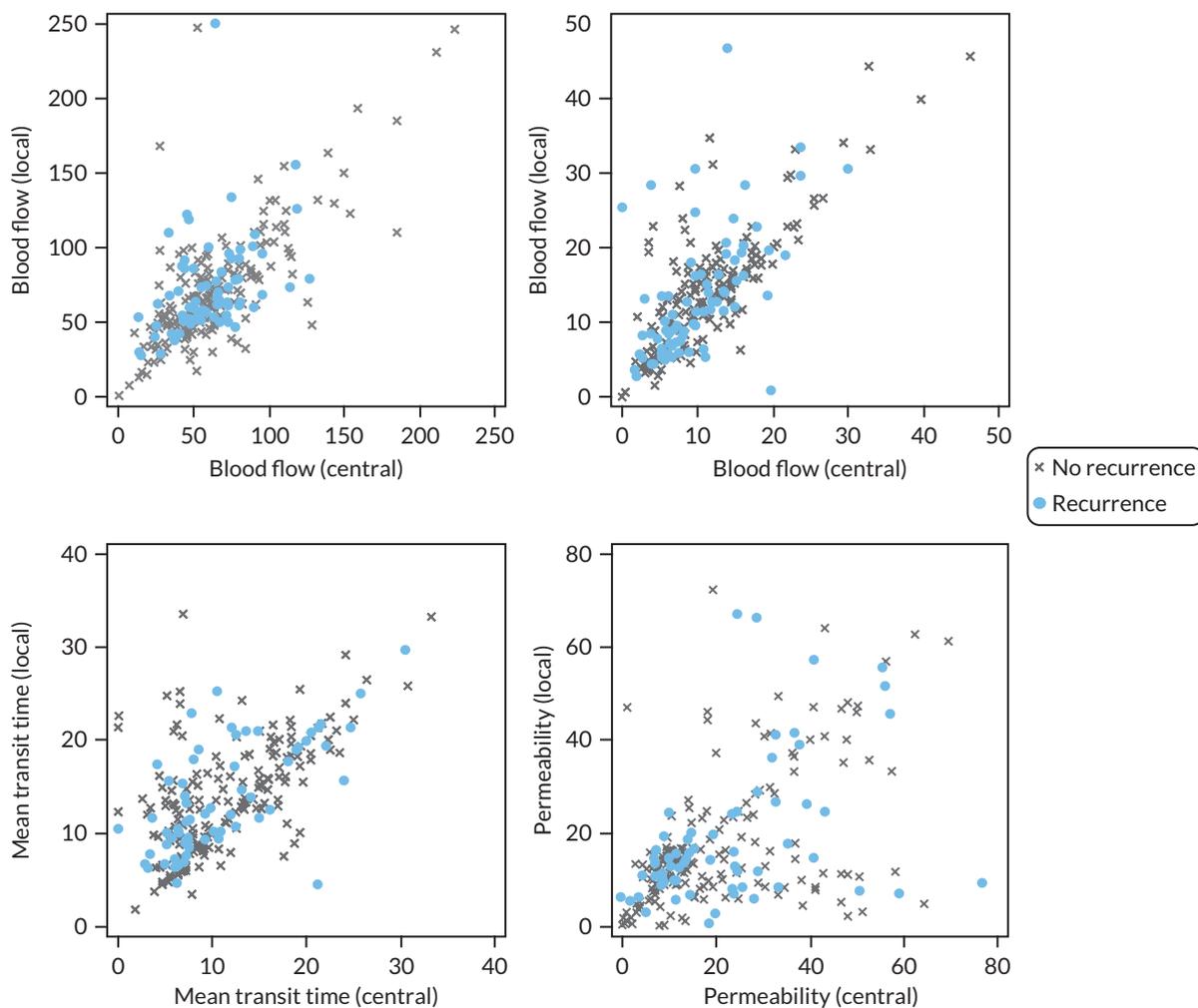


FIGURE 17 Scatter plots showing the values from local and central review on the same plot. Participants with and without recurrence are also highlighted differently.

Image analysis was undertaken on four different manufacturer platforms (GE, Siemens, Phillips, Toshiba) and across different manufacturer software versions (GE, three versions; Siemens, three versions; Phillips, one version; Toshiba, one version) depending on corresponding availability at the local sites.

Scatter plots for blood flow, blood volume, mean transit time and permeability surface measurements from local and central measurements are shown in [Figure 17](#). These plots highlight variability between local and central perfusion CT measurements, particularly for permeability surface area product.

The mean difference and 95% limits of agreement are summarised in [Appendix 1, Table 58](#).

Bland–Altman plots for blood flow, blood volume, mean transit time and permeability surface measurements for local and central reviews are shown in [Figure 18](#). Again, these indicate that local and central perfusion CT measurements varied, particularly for permeability surface area product measurements, where greater variation was noted at higher values, demonstrated by a greater range of measurements outside of the 95% limits of agreement line compared with other measurements.

Differences in the size of the ROI placed and in each vascular parameter between local and central measurements are plotted in [Figure 19](#). Most scans were performed on CT scanners manufactured by GE, followed by Siemens. Data were acquired using Toshiba and Phillips scanners from only one site each. The plotted data indicate that ROIs differed between local and central estimates, reflecting observer variation, but that this was not a major contributor to differences in vascular parameters.

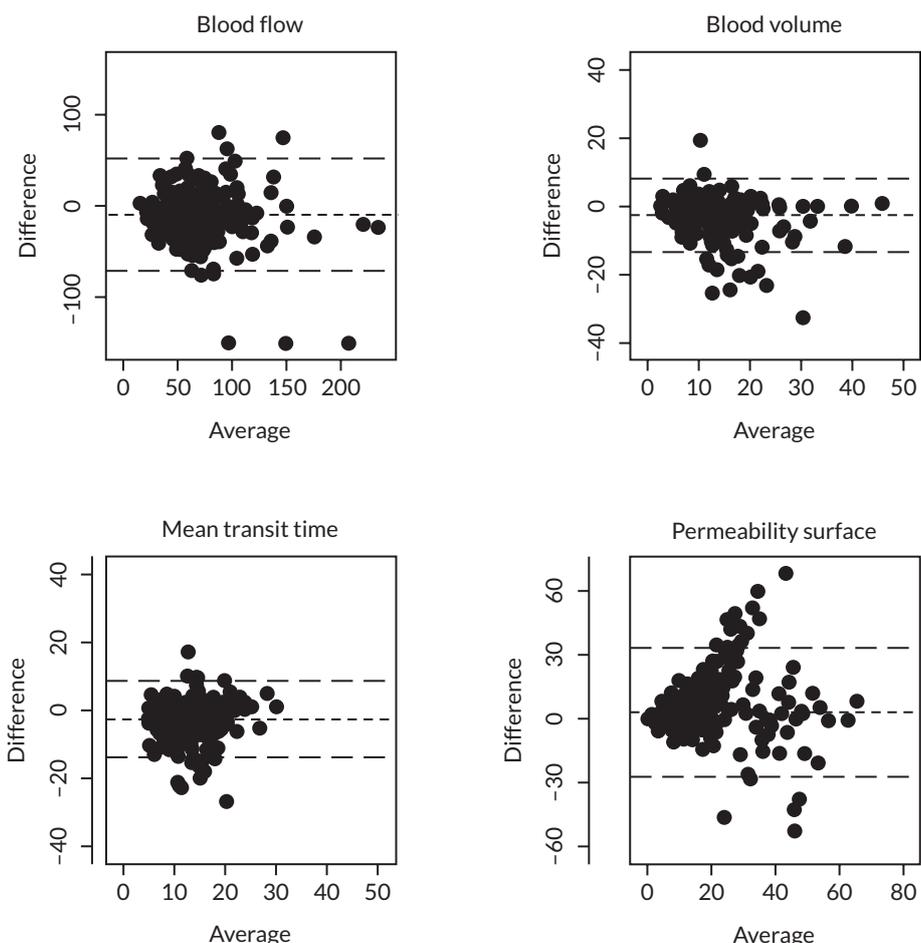


FIGURE 18 Bland–Altman plots showing the difference plotted against the average measurement for blood flow, blood volume, mean transit time and permeability surface area product, respectively; central dashed line represents the mean difference, the outer dashed lines represent the 95% limits of agreement. Three blood flow values > 150 ml/minute/ml were plotted at -150 for illustrative purposes only.

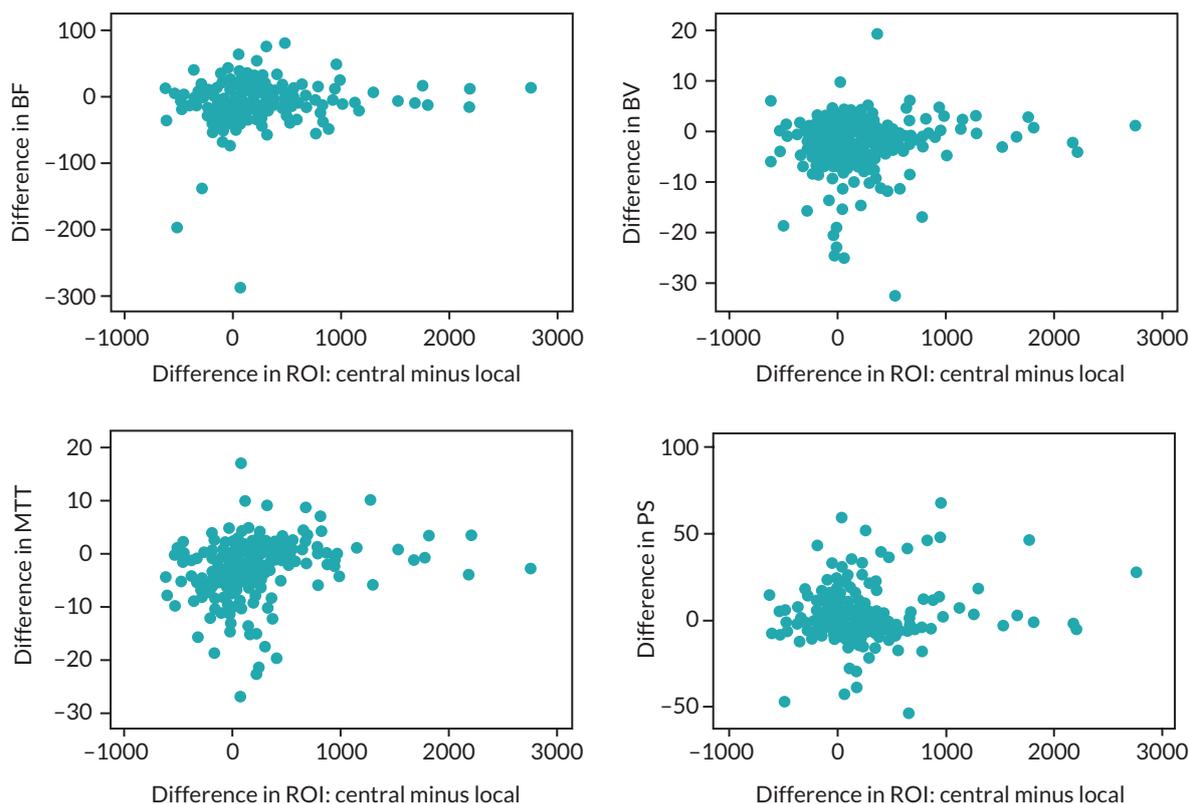


FIGURE 19 Difference in size of ROI placed plotted against the difference in each vascular parameter between local and central review.

Similarly, the scanner manufacturer did not appear to impact on the variance with the two most common manufacturers (GE and Siemens) displaying large differences in vascular parameters (see [Appendix 2, Figure 33](#)). However, there may be a greater data spread (i.e. less precision) with respect to differences in ROIs with GE scanners, based on visual inspection of the spread of the majority of measurements.

Chapter 6 Associations between perfusion computed tomography and pathological variables

Introduction

Hypoxia and angiogenesis are believed to be important drivers of tumour growth and progression.⁸³⁻⁸⁵ Poor vascularisation and hypoxia are also associated with reduced efficacy of radiotherapy, chemotherapy and immunotherapy.⁸⁶ Tumours acquire a blood supply via a number of mechanisms including vasculogenesis, co-option and intussusception.⁸⁷ Limited oxygen availability within tumours arises from increasing distance from the vasculature.

Hypoxia contributes to angiogenesis early in the adenoma–carcinoma sequence through HIF-1 mediated overproduction of proangiogenic growth factors, such as VEGF, which induces formation of new blood vessels.⁸⁸ VEGF expression has been shown to increase across the adenoma phase.⁸⁸

Angiogenesis in colorectal cancer is characterised structurally by fragile tortuous blood vessels that are hyperpermeable owing to an incomplete endothelium, with a relative absence of smooth muscle and pericyte coverage. Unlike the distribution of normal tissue vessels, the distribution of tumour vessels is chaotic, with areas of low vessel density mixed with regions of high angiogenic activity and an overall centripetal decline in vessel density of both colorectal adenomas and carcinomas.⁸⁹ These tumour vessels also function differently from normal vessels, with arteriovenous shunting and intermittent, or even reversed, flow, which may be due to a combination of high haematocrit level and vessel compression from raised interstitial pressure.

Hypoxia-inducible factor also induces expression of glucose transporters and glycolytic enzymes, including GLUT-1 and hexokinase-2.⁹⁰ Thus, GLUT-1 may serve as an additional downstream indicator for hypoxia.

In this chapter, we explore potential relationships between perfusion CT and pathological biomarker variables, looking at biological plausibility. If perfusion CT variables genuinely reflect vascular parameters closely, we would expect an association with known/proven pathological markers of the same.

First, we explored the relationship between perfusion CT and immunohistochemistry, specifically CD105 (reflecting neovascularisation), VEGF (reflecting angiogenesis), HIF-1 and GLUT-1 (both reflecting hypoxia). Here, we hypothesised that increasing perfusion CT vascular blood volume and permeability reflect increasing neovascularisation, respectively, while lower blood flow is associated with hypoxia. Second, we explored whether there were relationships between perfusion CT parameters and pathological venous invasion. Here, we hypothesised that there was an association between poorly perfused tumours and the presence of venous invasion. Third, we assessed whether there were relationships between perfusion CT parameters and tumour regression following neoadjuvant therapy in rectal cancers. Here, we hypothesised that there was an association between poorly perfused tumours and poor response to neoadjuvant therapy.

Methods

Participants

All participants consenting for the main trial gave permission for their data to be used for this substudy.

Perfusion computed tomography imaging and analysis

As described previously in [Chapter 2](#), perfusion CT imaging was performed across 13 local hospital sites, using a dynamic acquisition, centred on the tumour, undertaken every 1.5 seconds for a total of 45 seconds, then every 15 seconds for an additional 75 seconds, to capture contrast agent inflow, outflow and recirculation. Image analysis was undertaken by 25 radiologists across the local hospital sites, using their respective commercial software platforms (GE,

Siemens, Phillips, Toshiba). The following vascular parameters were recorded: blood flow, blood volume, mean transit time and permeability surface area product.

Pathological analysis

Immunohistochemistry

As described previously, additional immunohistochemical staining was performed centrally for vascularisation, angiogenesis and hypoxia using the following concentrations of antibodies: CD105 (Novocastra; 1/200 dilution, discontinued); VEGF (Dako; concentration: 0.45 µg/ml); HIF-1α (Abcam; concentration: 1.94 µg/ml) and Glut-1 (Millipore; concentration: 2.5 µg/ml). For MMR status the following was used: MLH1 (Novocastra; concentration: 19.5 µg/ml), MSH2 (Novocastra; concentration: 1.94 µg/ml), MSH6 (Dako; ready-to-use antibody in 0.015 mol/l sodium azide) and PMS2 (BD Pharmingen; concentration: 1.6 µg/ml).

Sections were stained in batches on the fully automated BOND-Max system (Leica Biosystems), which was used in conjunction with the BOND Polymer Refine detection system (Leica Biosystems). All slides were scanned at ×20 magnification using a Hamamatsu Nanozoomer 2.0 RS, and images were exhibited in a liquid crystal display monitor under contrast, focus, saturation and white balance standardisation.

CD105-stained vessels with a clearly defined lumen or well-defined linear vessel shape were considered for microvessel assessment. The invasive front of each sample was selected and the two areas of highest vascularisation (hot spots) averaged and expressed as a count per mm².

Scores for VEGF, GLUT-1 and HIF-1α were based on staining intensity and the proportion of positively stained cells, according to previously published systems. VEGF and GLUT-1 expression was calculated by combining the intensity of stained cells (0–3) with the percentage of positive cells (0–4) and HIF-1α expression on the combined cytoplasmic and nuclear staining (range, 0–6).

Mismatch repair protein expression was assessed by determining retained expression or lack of staining in tumour areas of the sample of MLH1, MSH2, MSH6 and PMS2.

The image analysis software Visiopharm was used to evaluate CD105 staining. For the remaining markers, semiquantitative analysis of immunoreactivity was performed by a histopathologist with more than 15 years of experience in gastrointestinal pathology.

Venous invasion

The presence or absence of venous invasion was recorded. Venous invasion was defined as tumour present within an endothelium-lined space, either surrounded by a rim of muscle or containing red blood cells.⁹¹ Venous invasion was also recorded when a rounded or elongated tumour profile that was not in direct continuity with the advancing tumour margin was identified adjacent to an artery, especially when no accompanying vein was seen.⁹¹

Tumour regression grade

In patients with rectal cancer, TRG following neoadjuvant therapy was assessed using a three-point system as follows:⁹²

TRG1: no cells, single cells or scattered small groups of cancer cells

TRG 2: residual cancer outgrown by fibrosis

TRG 3: extensive residual cancer, or fibrosis outgrown by cancer

Statistical analysis

Potential relationships between perfusion CT variables and pathology characteristics in the tumours were assessed by descriptive analyses.

Results

CD105 microvessel density

Scatter plots of both local and central measurements of perfusion CT variables plotted against CD105 microvessel density are shown in [Figure 20](#). There was no clear association between perfusion CT variables and CD105 expression.

Vascular endothelial growth factor expression

Scatter plots of both local and central measurements of perfusion CT variables plotted against VEGF score are shown in [Appendix 2, Figure 34](#). Most tumours were VEGF negative. There was no clear relationship between perfusion CT variables and VEGF expression.

Hypoxia-inducible factor-1 expression

Scatter plots of both local and central measurements of perfusion CT variables plotted against HIF-1 score are shown in [Appendix 2, Figure 35](#). Most tumours were HIF-1 negative. There was no clear relationship between perfusion CT variables and HIF-1 expression. Left-sided graphs are local review and right-sided graphs are central review.

Glucose transporter protein-1 expression

Scatter plots of both local and central measurements of perfusion CT variables plotted against GLUT-1 score are shown in [Appendix 2, Figure 36](#). There was no clear relationship between perfusion CT variables and GLUT-1 expression.

Mismatch repair status

Scatter plots of both local and central measurements of perfusion CT variables plotted against MMR status are shown in [Figure 21](#). Most tumours were MMR proficient. There was no difference between perfusion CT variables and MMR-deficient/MMR-proficient tumours.

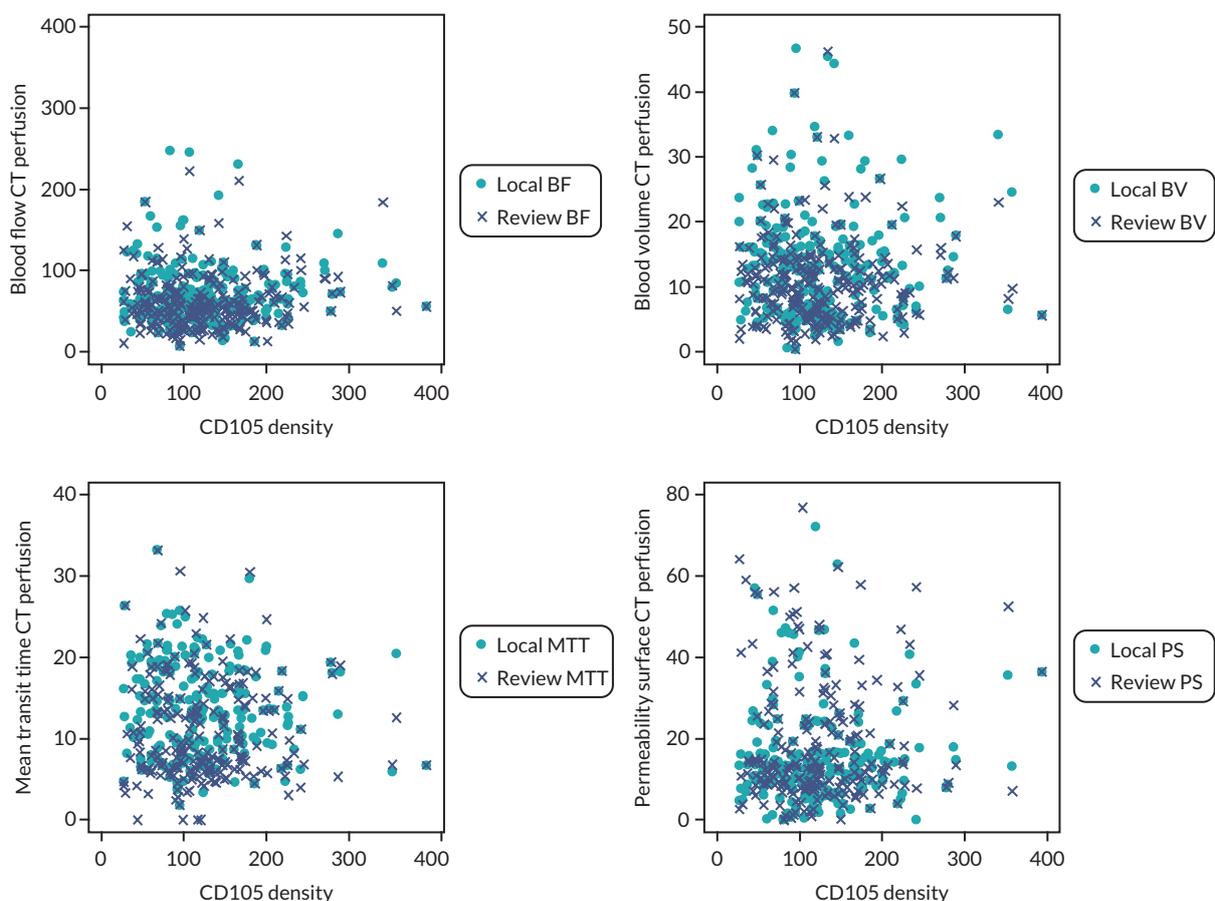


FIGURE 20 Scatter plots of blood flow, blood volume, mean transit time and permeability surface area product against CD105 microvessel density.

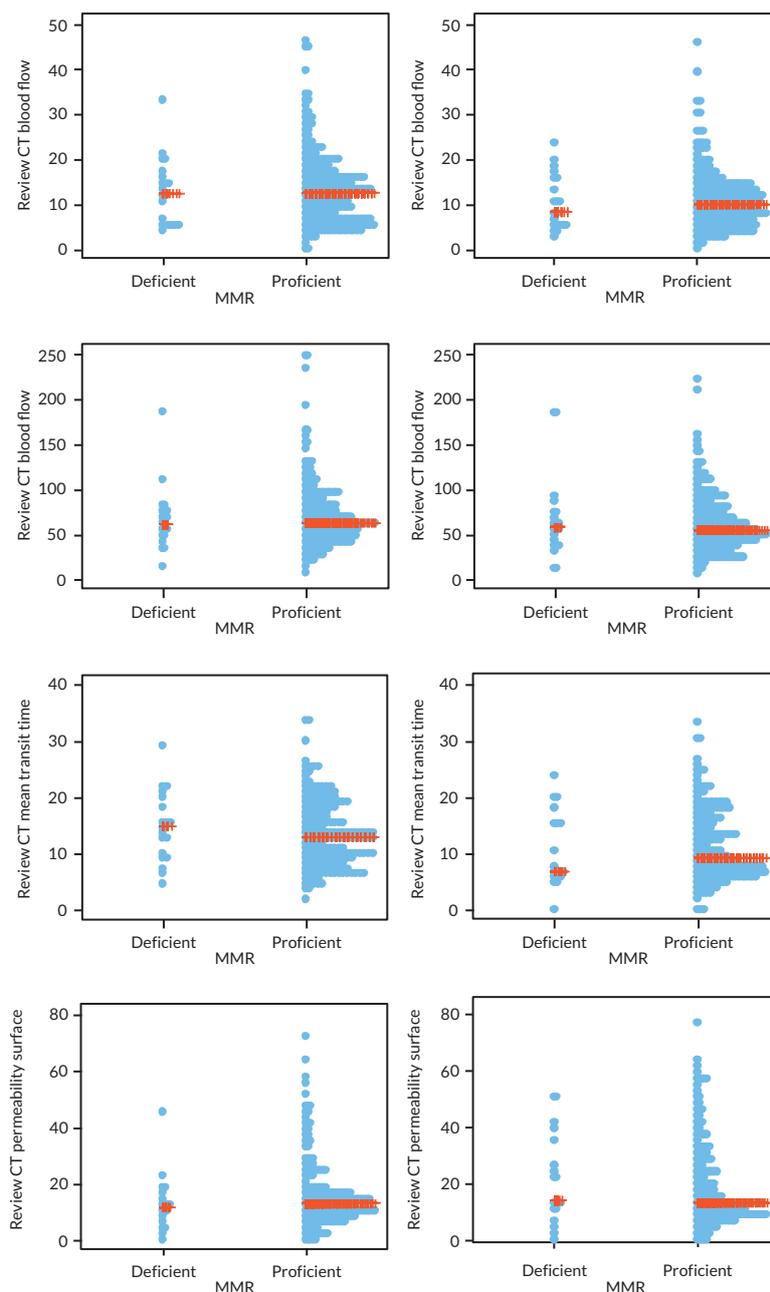


FIGURE 21 Scatter plots of blood flow, blood volume, mean transit time and permeability surface area product against MMR status. Left-sided graphs are local review and right-sided graphs are central review.

Venous invasion

Scatter plots of both local and central measurements of perfusion CT variables plotted against the presence or absence of venous invasion are shown in [Appendix 2, Figure 37](#). There was no clear difference between perfusion CT variables and venous invasion status.

Tumour regression grade

Scatter plots of both local and central measurements of perfusion CT variables plotted against TRG for the subset of patients with rectal cancer treated with neoadjuvant therapy are shown in [Appendix 2, Figure 38](#). There was no clear relationship between baseline perfusion CT variables and TRG.

Chapter 7 Discussion

Colorectal cancer is a leading cause of cancer-related death. Around half of patients treated with curative intent will develop recurrent disease, usually metastasis, and die from this. Earlier identification of these at-risk patients remains an unmet need so that chemotherapy can be offered on a more personalised basis than presently. Prognostic models combine multiple datapoints relating to an individual patient and weight them to estimate the risk of recurrence (expressed usually as overall survival and disease-free survival). However, although models for colorectal cancer exist, they are not used widely, perhaps because they omit 'cutting-edge' predictive biomarkers believed to be useful. Accordingly, in day-to-day practice, risk is most often estimated using the UICC/AJCC TNM system, yet this has its own limitations. In patients without metastasis at baseline staging, prognosis is derived from only two variables – 'T' (depth of tumour penetration into or through the gut wall) and 'N' (the extent to which local lymph nodes are involved). While simplicity is undoubtedly attractive, restricting to just two prognostic variables suggests that valuable information is being ignored. Supporting this, recent data have shown that patients with stage IIIA disease may experience better outcomes than those with stage IIB/IIC disease.^{21,22} Within stage II, there is also outcome variation related to tumour extent.

It follows that the major aim of PROSPECT was to improve prognostication for colorectal cancer by developing a multivariable model that incorporated not only standard clinicopathological variables, but also more novel imaging, immunohistochemical and genetic biomarker variables. We used robust statistical modelling techniques to counter emerging criticisms of that models are overfitted. Overfitting is especially problematic where modelling methods are not prespecified for data-driven approaches, such as exploratory multivariable modelling and machine learning. For example, a recent viewpoint concluded that publication of clinically useless prediction models is exponential, stating that while researchers can procure necessary data and computational power, they lack sufficient methodological expertise.⁹³ A 1994 article (28 years ago!) criticised researchers for focusing on 'pet' biomarkers, while discounting those already known to be useful and ignoring correct study designs.⁹⁴

Statisticians advocate building a 'baseline' model from predictors already known or believed to be clinically useful. The benefit of novel biomarkers is then determined by whether their addition to the baseline model improves prediction significantly. Accordingly, unlike the large majority of multivariable models developed in the literature, we did not use univariable significance within the study data set to select the predictor variables included in our models (or to exclude others), because this greatly increases the risk of overfitting. Instead, we prespecified all variables that would be included; that is, we selected variables upfront and all of these were then studied, irrespective of any univariable significance. We then obtained these variables via a prospective, multicentre trial to minimise bias and overfitting during model development. While, by definition, this results in an expensive and prolonged trial, the results are likely to have enhanced generalisability. Participants were enrolled from 13 hospital sites (district general and teaching hospitals) representing real-world practice. The trial was powered around novel imaging variables, specifically perfusion CT. Compared with therapeutic trials of colorectal cancer, participants were older and, in the main, underwent standard therapies.

The results of our modelling exercise are described in [Chapter 4](#). In summary, we found that the baseline model comprising standard clinicopathological variables demonstrated superior prognostic accuracy to TN staging. Sensitivity and specificity of TN staging for predicting recurrence within 3 years were 0.56 (95% CI 0.44 to 0.67) and 0.58 (95% CI 0.51 to 0.64), respectively. We deployed the standard model at two operating points. When used to distinguish 'high-' from 'medium-/low-risk' patients, the model improved specificity above standard TN prediction to 0.74 (95% CI 0.68 to 0.79), while maintaining equivalent sensitivity of 0.57 (95% CI 0.45 to 0.68). When used to distinguish high-/medium-risk patients from 'low'-risk patients, sensitivity improved to 0.89 (95% CI 0.80 to 0.95), but with diminished specificity of 0.40 (95% CI 0.31 to 0.47). While the model improves on standard clinical practice, at this stage, it is not clear to us which would be more desirable: improved specificity or improved sensitivity at the cost of diminished specificity. Work around patient and clinician preferences will help clarify this issue.

In addition to TN stage (used already for TN prognosis), the standard model also incorporated age, sex, tumour size, tumour location, treatment and the presence or absence of venous invasion. All these predictors are collected already in

routine clinical practice, which will greatly facilitate model adoption. While the equation comprises complex weightings, it is relatively simple to overcome this barrier by using an online calculator or spreadsheet to facilitate adoption. Furthermore, T and N stage and venous invasion can be assessed preoperatively via CT/MRI scanning, so it would be feasible to use the model for clinical decision-making in a wholly neoadjuvant setting. The same comments also apply when there is a complete response to neoadjuvant therapy and the primary (usually rectal cancer) is left unresected as a result.

A major objective of the PROSPECT trial was to counter criticisms that models do not incorporate the latest ‘cutting edge’ predictors. The results of our modelling exercise are described in [Chapter 4](#). Ultimately, we found that no novel imaging, immunohistopathological or genetic biomarker improved model performance significantly. Naturally, we were very disappointed by this finding but perhaps we should not be surprised: A recent article stated that ‘omic’ research often ignores clinical data and/or fails to develop models appropriately.⁹⁵ The authors pleaded that models must include all relevant clinical data before adding omics. As proof, they developed a model for breast cancer survival that included stage, age, receptors and grade. The addition of gene expression failed to improve prediction. Indeed, gene expression only became useful when basic clinical data were excluded altogether. Ultimately, the authors argued that omics ‘may not be much more than surrogates for clinical data.’⁹⁵ Similarly, other researchers found that novel biomarkers of cardiovascular disease contributed little over-and-above basic clinical measurements.⁹⁶

[Chapter 5](#) describes comparison of local and central measurements of perfusion CT. Perfusion CT was read by 26 different radiologists locally, and we chose local measurements for our primary outcome model, since that approach would be most generalisable when needed for model deployment. However, we recognise fully that rapid improvements in digital data transfer, particularly Digital Imaging and Communications in Medicine, means that central review could be a realistic proposition in the near future, so we decided to investigate this central reading paradigm as well, particularly as perfusion CT was a relatively new measurement for many local radiologists even after training. Bland–Altman plots did indicate important levels of variation between local and central measurements, especially for permeability surface area product measurements, where greater variation was noted at higher values, and with a wider range of values lying outside of the 95% limits of agreement line compared with other measurements. However, any variation (and the desire to reduce it) is rendered irrelevant by the fact that CT perfusion measurements, from whatever source, had no prognostic benefit when added to our standard baseline model.

A priori, we hypothesised that for CT perfusion measurements to be valid, then they must reflect biological processes around angiogenesis and tumour perfusion. We therefore reasoned that CT perfusion measurements would display significant relationships with pathological measurements known to reflect angiogenesis, etc. The results of these analyses are described in [Chapter 6](#). In summary, we could not identify direct relationships between imaging and pathological variables, suggesting that CT perfusion measurements, while providing physiological (vascular) measurements, are not biologically valid. However, we did not register levels and areas of vascular measurement by CT to pathological assessment, which may influence our findings. Precise registration between imaging and pathology specimens poses considerable technical difficulties and would have been unrealistic in a multicentre setting.

The PROSPECT trial does have limitations. In particular, at this stage, the final best model has not undergone external evaluation, notably in an NHS setting, so its generalisability is unknown. We would prefer that any external evaluation is led by independent researchers, to diminish potential bias. While our data suggest that the model will improve specificity for prediction of subsequent metastasis over current TN staging (without diminished sensitivity), we did not assess whether this is clinically useful. Ultimately, clinical utility is the most telling measure of whether clinicians and patients will find the model useful for clinical decision-making. We assessed our model against TN staging since this is standard current practice. We did not assess the model against any others available and any external evaluation should determine how our model compares against others that are promising. Novel commercial models are also increasingly available. For example, Immunoscore® (HaliuDx, Marseille, France), which determines the individual immune response to colorectal cancer, is used alongside TN staging to help predict relapse. Our model should be compared against these alternatives in an appropriate setting to generate data that are generalisable to the NHS.

DISCUSSION

Our findings are also limited by the fact that the number of patients developing metastasis was lower than expected from historical data and the number of exclusions/withdrawals was higher than anticipated, mostly due to a higher prevalence of metastasis at baseline. The number of patients undergoing additional histopathological analysis was relatively small because PROSPECT was not powered specifically to detect an effect for these variables. However, if a beneficial effect on prediction exists, it is likely to be small.

Chapter 8 Overall conclusions and implications for practice

In conclusion, we found that a prognostic model based on prospectively derived clinicopathological variables, including age, sex, treatment, tumour size, tumour location, TN stage and presence/absence of venous invasion, had the best performance for predicting recurrent disease. Performance was superior to TN staging at two operating points: one that improved specificity and one that improved sensitivity at the cost of diminished specificity. This is a promising step forward for clinical practice towards more personalised prognostication to guide management. We also found that none of the novel imaging, immunohistochemical or genetic biomarkers selected by us a priori was able to improve prognostication significantly. We must conclude that these models offer nothing over and above standard clinicopathological data.

Future research could consider the following issues:

- Prospective testing of the clinicopathological prognostic model in trial and clinical settings will be important to assess its added value to standard staging and potential impact on treatment decisions and outcomes.
- It is not clear to us which would be more desirable: improved specificity or improved sensitivity at the cost of diminished specificity. Prospective testing in clinical practice should also incorporate work around patient and clinician preferences regarding these issues.
- Venous invasion on pathological evaluation of the resected specimen was a strong prognostic factor in the baseline model: Further research regarding preoperative imaging assessment of venous invasion on CT for colon cancer, and MRI for rectal cancer, is warranted to facilitate model deployment for prognostication in the wholly neoadjuvant setting. If preoperative imaging agrees with postoperative pathology sufficiently, then all model predictors could be obtained preoperatively without diminished prognostication.
- Data on the prognostic value of KRAS and BRAF in the primary tumour setting remain limited; further large scale prospective studies may still be required.

Current interest around artificial intelligence/deep learning is considerable; although based on methodological literature, it is unlikely that there will be an improvement in the model performance using these methods. Many have been developed without prespecified variables/analyses and with insufficient events for the number of variables studied. Ill-specified, underpowered research using our data set could undermine our findings by generating false-positive results. Accordingly, requests to use PROSPECT data will require submission of a full protocol and analysis plans to the PROSPECT Trial Management Group.

Additional information

Contributions of authors

Vicky Goh (<https://orcid.org/0000-0002-2321-8091>) (Professor of Clinical Cancer Imaging) helped conceive and design the trial and achieve funding. She was chief investigator and responsible for the overall conduct and day-to-day management of the trial. She led interpretation of imaging biomarkers and helped with data interpretation.

Susan Mallett (<https://orcid.org/0000-0002-0596-8200>) (Professor of Diagnostic and Prognostic Medical Statistics) helped conceive and design the trial and achieve funding. She developed the trial design, statistical analysis plan, conducted the statistical analysis, helped with data interpretation and prepared analyses for publication.

Manuel Rodriguez-Justo (<https://orcid.org/0000-0001-5007-1761>) (Consultant Pathologist) helped conceive and design the trial and achieve funding. He led those aspects of the trial dealing with histopathological, genetic and immunohistochemical biomarkers.

Victor Boulter (Patient Representative) helped with trial conduct and data interpretation from a patient and lay perspective.

Rob Glynne-Jones (<https://orcid.org/0000-0002-6742-222X>) (Consultant Clinical Oncologist) helped conceive and design the trial and achieve funding. He helped with trial conduct and data interpretation.

Saif Khan (<https://orcid.org/0000-0002-5957-328X>) (Research Technician) helped with trial conduct and assisted with those aspects of the trial dealing with histopathological, genetic and immunohistochemical biomarkers.

Sarah Lessels (<https://orcid.org/0000-0002-7373-3067>) helped with trial conduct and day-to-day management.

Dominic Patel (<https://orcid.org/0000-0001-9223-6632>) (Research Manager) helped with trial conduct and assisted with those aspects of the trial dealing with histopathological, genetic and immunohistochemical biomarkers.

Davide Prezzi (<https://orcid.org/0000-0002-0878-2196>) (Consultant Radiologist) helped with trial conduct and assisted with interpretation of imaging biomarkers.

Stuart Taylor (<https://orcid.org/0000-0002-6765-8806>) (Professor of Medical Imaging) helped conceive and design the trial and achieve funding. He helped with trial conduct and data interpretation.

Steve Halligan (<https://orcid.org/0000-0003-0632-5108>) (Professor of Radiology) helped conceive and design the trial and achieve funding. He helped with trial conduct and data interpretation.

Vicky Goh, Susan Mallett and Steve Halligan produced the first draft of this monograph. All authors then helped with revision prior to submission.

Acknowledgements

The authors of this monograph would like to thank all those patients and their carers who took the time to consider participating at what will have been a very difficult time of their lives.

We would like to thank the PROSPECT Investigators (in alphabetical order): radiologists (with UK address at the time the data were acquired): Richard Beable, Department of Radiology, Queen Alexandra Hospital, Portsmouth. Margaret Betts, Department of Radiology, Oxford Radcliffe University Hospitals, Oxford. David Breen, Department of Radiology, University Hospital Southampton, Southampton. Ingrid Britton, Department of Radiology, North

Staffordshire University Hospital, Stoke-on-Trent. John Brush, Department of Radiology, Western General Hospital, Edinburgh. Peter Correa, Department of Oncology, University Hospitals and Coventry and Warwickshire NHS Trust, Coventry. Nick Dodds, Clinical Imaging, Royal Cornwall Hospitals NHS Trust, Truro. Catherine Grierson, Department of Radiology, University Hospital Southampton, Southampton. Nyree Griffin, Clinical Imaging, Guy's and St Thomas' NHS Foundation Trust, London. Sofia Gourtsoyianni, Clinical Imaging, Guy's and St Thomas' NHS Foundation Trust, London. James Hampton, Department of Radiology, Sheffield Teaching Hospitals NHS Trust, Sheffield. Anthony Higginson, Department of Radiology, Queen Alexandra Hospital, Portsmouth. Andrew Lowe, Department of Radiology, Bradford Teaching Hospitals NHS Trust, Bradford. Richard Mannion, Department of Radiology, York Teaching Hospital NHS Foundation Trust, York. Amjad Mohammed, Department of Radiology, Bradford Teaching Hospitals NHS Foundation Trust, Bradford. Colin Oliver, Department of Radiology, University Hospitals and Coventry and Warwickshire NHS Trust, Coventry. Andrew Slater, Department of Radiology, Oxford Radcliffe University Hospitals, Oxford. Madeline Strugnell, Clinical Imaging, Royal Cornwall Hospitals NHS Trust, Truro. Biju Thomas, Department of Radiology, University Hospital Southampton, Southampton. Damian Tolan, Department of Radiology, Leeds Teaching Hospitals NHS Trust, Leeds. Raghu Vinayagam, Department of Radiology, Sheffield Teaching Hospitals NHS Trust, Sheffield. Ian Zealley, Department of Radiology, Western General Hospital, Edinburgh and Ninewells Hospital, Dundee.

Local site radiographers and research nurses and staff: D Barrie, G Cattini, P Clewer, N Cowan, M Dodds, S Dundas, N Gibbons, A Green, M Harris, G Haley, L Houston, K Jones, L Jones, M Lewis, C Marsh, E Ngandwe, B Pharoah, V Ritchie, C Roe, R Smith, D Tew, K Wallace, A Williams, P Woodburn, N Wragg.

Scottish Clinical Trials Research Unit: Joanna Dunlop, Scottish Clinical Trials Research Unit (SCTRU), NHS National Services Scotland, Edinburgh, UK.

We would also like to thank those independent members of the Trial Steering Committee, data monitoring committee and ad hoc advisors who are not authors of this monograph: D G Altman (deceased at the time of writing), CI Bartram, A Groves, A Hackshaw, P Hoskin, D Miles, K Miles, A Rockall, C Streete, J Wilson.

Data-sharing statement

Requests to use PROSPECT data should be submitted to the corresponding author.

Ethics statement

Ethical permission was granted by Bloomsbury Research Ethics Committee (reference number 10/H0713/84) in 2011 and the trial was conducted in accordance with the principles of good clinical practice.

Information governance statement

King's College London is committed to handling all personal information in line with the UK Data Protection Act (2018) and the General Data Protection Regulation (EU GDPR) 2016/679. Under the Data Protection legislation, King's College London is the Data Controller, and you can find out more about how we handle personal data, including how to exercise your individual rights and the contact details for our Data Protection Officer here (www.kcl.ac.uk/professional-services/business-assurance/data-protection-introduction-data-losses-and-reporting).

Disclosure of interests

Full disclosure of interests: Completed ICMJE forms for all authors, including all related interests, are available in the toolkit on the NIHR Journals Library report publication page at <https://doi.org/10.3310/BTMT7049>.

Primary conflicts of interest: None declared.

References

1. UICC. *Global Cancer Data: GLOBOCAN 2018*. URL: www.uicc.org/news/new-global-cancer-data-globocan-2018 (accessed 19 July 2019).
2. Cancer Research UK. *Cancer Statistics*. URL: www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/incidence?_ga=2.99056855.855681031.1563547912-1431267947.1556168137 (accessed 19 July 2019).
3. Heald RJ, Husband EM, Ryall RD. The mesorectum in rectal cancer surgery: the clue to pelvic recurrence? *Br J Surg* 1982;**69**:613–6.
4. Birbeck KF, Macklin CP, Tiffin NJ, Parsons W, Dixon MF, Mapstone NP, *et al.* Rates of circumferential resection margin involvement vary between surgeons and predict outcomes in rectal cancer surgery. *Ann Surg* 2002;**235**:449–57.
5. Cedermark B, Dahlberg M, Glimelius B, Pahlman L, Rutqvist LE, Wilking N; Swedish Rectal Cancer Trial. Improved survival with preoperative radiotherapy in resectable rectal cancer. *N Engl J Med* 1997;**336**:980–7.
6. Kapiteijn E, Marijnen CA, Nagtegaal ID, Putter H, Steup WH, Wiggers T, *et al.*; Dutch Colorectal Cancer Group. Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer. *N Engl J Med* 2001;**345**:638–46.
7. Braendengen M, Tveit KM, Berglund A, Birkemeyer E, Frykholm G, Pahlman L, *et al.* Randomized phase III study comparing preoperative radiotherapy with chemoradiotherapy in nonresectable rectal cancer. *J Clin Oncol* 2008;**26**:3687–94.
8. Wolmark N, Fisher B, Rockette H, Redmond C, Wickerham DL, Fisher ER, *et al.* Postoperative adjuvant chemotherapy or BCG for colon cancer: results from NSABP protocol C-01. *J Natl Cancer Inst* 1988;**80**:30–6.
9. Moertel CG, Fleming TR, Macdonald JS, Haller DG, Laurie JA, Goodman PJ, *et al.* Levamisole and fluorouracil for adjuvant therapy of resected colon carcinoma. *N Engl J Med* 1990;**322**:352–8.
10. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, *et al.* *Globocan 2012: Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2012 v1.0*. IARC CancerBase No. 11. Lyon, France: International Agency for Research on Cancer; 2012.
11. Gill S, Loprinzi CL, Sargent DJ, Thomé SD, Alberts SR, Haller DG, *et al.* Pooled analysis of fluorouracil-based adjuvant therapy for stage II and III colon cancer: who benefits and by how much? *J Clin Oncol* 2004;**22**:1797–806.
12. International Multicentre Pooled Analysis of Colon Cancer Trials (IMPACT) investigators. Efficacy of adjuvant fluorouracil and folinic acid in colon cancer. International Multicentre Pooled Analysis of Colon Cancer Trials (IMPACT) investigators. *Lancet* 1995;**345**:939–44.
13. Andre T, Sargent D, Tabernero J, O'Connell M, Buyse M, Sobrero A, *et al.* Current issues in adjuvant treatment of stage II colon cancer. *Ann Surg Oncol* 2006;**13**:887–98.
14. Andre T, Boni C, Mounedji-Boudiaf L, Navarro M, Tabernero J, Hickish T, *et al.*; Multicenter International Study of Oxaliplatin/5-Fluorouracil/Leucovorin in the Adjuvant Treatment of Colon Cancer (MOSAIC) Investigators. Oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment for colon cancer. *N Engl J Med* 2004;**350**:2343–51.
15. Mamounas E, Wieand S, Wolmark N, Bear HD, Atkins JN, Song K, *et al.* Comparative efficacy of adjuvant chemotherapy in patients with Dukes' B versus Dukes' C colon cancer: results from four National Surgical Adjuvant Breast and Bowel Project adjuvant studies (C-01, C-02, C-03, and C-04). *J Clin Oncol* 1999;**17**:1349–55.

16. Wolmark N, Rockette H, Mamounas E, Jones J, Wieand S, Wickerham DL, *et al.* Clinical trial to assess the relative efficacy of fluorouracil and leucovorin, fluorouracil and levamisole, and fluorouracil, leucovorin, and levamisole in patients with Dukes' B and C carcinoma of the colon: results from National Surgical Adjuvant Breast and Bowel Project C-04. *J Clin Oncol* 1999;**17**:3553–9.
17. Figueredo A, Coombes ME, Mukherjee S. Adjuvant therapy for completely resected stage II colon cancer. *Cochrane Database Syst Rev* 2008;**2008**:CD005390.
18. 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.
19. Howlander N, Noone AM, Krapcho M, Miller D, Bishop K, Kosary CL, *et al.* editors. *SEER Cancer Statistics Review, 1975–2014*. Bethesda, MD: National Cancer Institute; 2018.
20. Brierley JD, Gospodarowicz MK, Wittekind C. *TNM Classification of Malignant Tumors*. 8th edn. Chichester: Wiley; 2022.
21. Kim MJ, Jeong SY, Choi SJ, Ryoo SB, Park JW, Park KJ, *et al.* Survival paradox between stage IIB/C (T4N0) and stage IIIA (T1-2N1) colon cancer. *Ann Surg Oncol* 2015;**22**:505–12.
22. Huang B, Mo S, Zhu L, Xu T, Cai G. The survival and clinicopathological differences between patients with stage IIIA and stage II rectal cancer: an analysis of 12,036 patients in the SEER database. *Oncotarget* 2016;**7**:79787–96.
23. Mahar AL, Compton C, Halabi S, Hess KR, Weiser MR, Groome PA. Personalizing prognosis in colorectal cancer: a systematic review of the quality and nature of clinical prognostic tools for survival outcomes. *J Surg Oncol* 2017;**116**:969–82.
24. Renfro LA, Grothey A, Xue Y, Saltz LB, Andre T, Twelves C, *et al.*; Adjuvant Colon Cancer Endpoints (ACCENT) Group. ACCENT-based web calculators to predict recurrence and overall survival in stage III colon cancer. *J Natl Cancer Inst* 2014;**106**:dju333.
25. Hurwitz H, Fehrenbacher L, Novotny W, Cartwright T, Hainsworth J, Heim W, *et al.* Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. *N Engl J Med* 2004;**350**:2335–42.
26. Cunningham D, Humblet Y, Siena S, Khayat D, Bleiberg H, Santoro A, *et al.* Cetuximab monotherapy and cetuximab plus irinotecan in irinotecan-refractory metastatic colorectal cancer. *N Engl J Med* 2004;**351**:337–45.
27. Jonker DJ, O'Callaghan CJ, Karapetis CS, Zalberg JR, Tu D, Au HJ, *et al.* Cetuximab for the treatment of colorectal cancer. *N Engl J Med* 2007;**357**:2040–8.
28. Van Cutsem E, Peeters M, Siena S, Humblet Y, Hendlisz A, Neyns B, *et al.* Open-label phase III trial of panitumumab plus best supportive care compared with best supportive care alone in patients with chemotherapy-refractory metastatic colorectal cancer. *J Clin Oncol* 2007;**25**:1658–64.
29. Siena S, Peeters M, Van Cutsem E, Humblet Y, Conte P, Bajetta E, *et al.* Association of progression-free survival with patient-reported outcomes and survival: results from a randomised phase 3 trial of panitumumab. *Br J Cancer* 2007;**97**:1469–74.
30. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, *et al.* PD-1 blockade in tumors with mismatch-repair deficiency. *N Engl J Med* 2015;**372**:2509–20.
31. Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK, *et al.* Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 2017;**357**:409–13.
32. Sepulveda AR, Hamilton SR, Allegra CJ, Grody W, Cushman-Vokoun AM, Funkhouser WK, *et al.* Molecular biomarkers for the evaluation of colorectal cancer: guideline from the American Society for Clinical Pathology, College of American Pathologists, Association for Molecular Pathology, and the American Society of Clinical Oncology. *J Clin Oncol* 2017;**35**:1453–86.

33. National Institute for Health and Care Excellence. *Colorectal Cancer (Update) [B1] Use of Molecular Biomarkers to Guide Systemic Therapy*. NICE Guideline NG151 Evidence Reviews. London: NICE; 2020. URL: www.nice.org.uk/guidance/ng151/evidence/b1-use-of-molecular-biomarkers-to-guide-systemic-therapy-pdf-7029391215 (accessed 1 July 2020).
34. Spano JP, Lagorce C, Atlan D, Milano G, Domont J, Benamouzig R, *et al*. Impact of EGFR expression on colorectal cancer patient prognosis and survival. *Ann Oncol* 2005;**16**:102–8.
35. Amado RG, Wolf M, Peeters M, Van Cutsem E, Siena S, Freeman DJ, *et al*. Wild-type KRAS is required for panitumumab efficacy in patients with metastatic colorectal cancer. *J Clin Oncol* 2008;**26**:1626–34.
36. Karapetis CS, Khambata-Ford S, Jonker DJ, O’Callaghan CJ, Tu D, Tebbutt NC, *et al*. K-ras mutations and benefit from cetuximab in advanced colorectal cancer. *N Engl J Med* 2008;**359**:1757–65.
37. Dahabreh IJ, Terasawa T, Castaldi PJ, Trikalinos TA. Systematic review: anti-epidermal growth factor receptor treatment effect modification by KRAS mutations in advanced colorectal cancer. *Ann Intern Med* 2011;**154**:37–49.
38. De Roock W, Claes B, Bernasconi D, De Schutter J, Biesmans B, Fountzilias G, *et al*. Effects of KRAS, BRAF, NRAS, and PIK3CA mutations on the efficacy of cetuximab plus chemotherapy in chemotherapy-refractory metastatic colorectal cancer: a retrospective consortium analysis. *Lancet Oncol* 2010;**11**:753–62.
39. National Institute for Health and Care Excellence. *Molecular Testing Strategies for Lynch Syndrome in People with Colorectal Cancer*. Diagnostics Guidance DG27. London: NICE; 2017. URL: www.nice.org.uk/guidance/dg27/chapter/2-Clinical-need-and-practice (accessed 1 July 2020).
40. Sauer R, Liersch T, Merkel S, Fietkau R, Hohenberger W, Hess C, *et al*. Preoperative versus postoperative chemoradiotherapy for locally advanced rectal cancer: results of the German CAO/ARO/AIO-94 randomized phase III trial after a median follow-up of 11 years. *J Clin Oncol* 2012;**30**:1926–33.
41. Dewdney A, Cunningham D, Tabernero J, Capdevila J, Glimelius B, Cervantes A, *et al*. Multicenter randomized phase II clinical trial comparing neoadjuvant oxaliplatin, capecitabine, and preoperative radiotherapy with or without cetuximab followed by total mesorectal excision in patients with high-risk rectal cancer (EXPERT-C). *J Clin Oncol* 2012;**30**:1620–7.
42. Glynne-Jones R, Hava N, Goh V, Bosompem S, Bridgewater J, Chau I, *et al*.; Bacchus Investigators. Bevacizumab and Combination Chemotherapy in rectal cancer Until Surgery (BACCHUS): a phase II, multicentre, open-label, randomised study of neoadjuvant chemotherapy alone in patients with high-risk cancer of the rectum. *BMC Cancer* 2015;**15**:764.
43. Dewdney A, Cunningham D, Chau I. Selecting patients with locally advanced rectal cancer for neoadjuvant treatment strategies. *Oncologist* 2013;**18**:833–42.
44. National Institute for Health and Care Excellence. *Colorectal Cancer: Diagnosis and Management*. Clinical Guideline CG131. London: NICE; 2014. URL: www.nice.org.uk/guidance/CG131 (accessed 19 July 2019).
45. National Institute for Health and Care Excellence. *Colorectal Cancer*, London: NICE 2020. URL: www.nice.org.uk/guidance/CG151 (accessed 12 February 2020).
46. Petersen RK, Hess S, Alavi A, Høilund-Carlsen PF. Clinical impact of FDG-PET/CT on colorectal cancer staging and treatment strategy. *Am J Nucl Med Mol Imaging* 2014;**4**:471–82.
47. Brush J, Boyd K, Chappell F, Crawford F, Dozier M, Fenwick E, *et al*. The value of FDG positron emission tomography/computerised tomography (PET/CT) in pre-operative staging of colorectal cancer: a systematic review and economic evaluation. *Health Technol Assess* 2011;**15**:1–192, iii.
48. Goh V, Halligan S, Daley F, Wellsted DM, Guenther T, Bartram CI. Colorectal tumor vascularity: quantitative assessment with multidetector CT – do tumor perfusion measurements reflect angiogenesis? *Radiology* 2008;**249**:510–7.

49. Goh V, Engledow A, Rodriguez-Justo M, Shastry M, Peck J, Blackman G, *et al.* The flow-metabolic phenotype of primary colorectal cancer: assessment by integrated 18F-FDG PET/perfusion CT with histopathologic correlation. *J Nucl Med* 2012;**53**:687–92.
50. Mandeville HC, Ng QS, Daley FM, Barber PR, Pierce G, Finch J, *et al.* Operable non-small cell lung cancer: correlation of volumetric helical dynamic contrast-enhanced CT parameters with immunohistochemical markers of tumor hypoxia. *Radiology* 2012;**264**:581–9.
51. Garcia-Figueiras R, Goh VJ, Padhani AR, Baleato-Gonzalez S, Garrido M, Leon L, Gómez-Caamaño A. CT perfusion in oncologic imaging: a useful tool? *AJR Am J Roentgenol* 2013;**200**:8–19.
52. Goh V, Glynne-Jones R. Perfusion CT imaging of colorectal cancer. *Br J Radiol* 2014;**87**:20130811.
53. Sahani DV, Kalva SP, Hamberg LM, Hahn PF, Willett CG, Saini S, *et al.* Assessing tumor perfusion and treatment response in rectal cancer with multisection CT: initial observations. *Radiology* 2005;**234**:785–92.
54. Goh V, Halligan S, Wellsted DM, Bartram CI. Can perfusion CT assessment of primary colorectal adenocarcinoma blood flow at staging predict for subsequent metastatic disease? A pilot study. *Eur Radiol* 2009;**19**:79–89.
55. Hayano K, Fujishiro T, Sahani DV, Satoh A, Aoyagi T, Ohira G, *et al.* Computed tomography perfusion imaging as a potential imaging biomarker of colorectal cancer. *World J Gastroenterol* 2014;**20**:17345–51.
56. Haider MA, Milosevic M, Fyles A, Sitartchouk I, Yeung I, Henderson E, *et al.* Assessment of the tumor microenvironment in cervix cancer using dynamic contrast enhanced CT, interstitial fluid pressure and oxygen measurements. *Int J Radiat Oncol Biol Phys* 2005;**62**:1100–7.
57. Hermans R, Meijerink M, Van den Bogaert W, Rijnders A, Weltens C, Lambin P. Tumor perfusion rate determined noninvasively by dynamic computed tomography predicts outcome in head-and-neck cancer after radiotherapy. *Int J Radiat Oncol Biol Phys* 2003;**57**:1351–6.
58. Hayano K, Desai GS, Kambadakone AR, Fuentes JM, Tanabe KK, Sahani DV. Quantitative characterization of hepatocellular carcinoma and metastatic liver tumor by CT perfusion. *Cancer Imaging* 2013;**13**:512–9.
59. Chen SH, Miles K, Taylor SA, Ganeshan B, Rodriguez M, Fraioli F, *et al.* FDG-PET/CT in colorectal cancer: potential for vascular-metabolic imaging to provide markers of prognosis. *Eur J Nucl Med Mol Imaging* 2021;**49**:371–84.
60. Sargent DJ, Patiyl S, Yothers G, Haller DG, Gray R, Benedetti J, *et al.*; ACCENT Group. End points for colon cancer adjuvant trials: observations and recommendations based on individual patient data from 20,898 patients enrolled onto 18 randomized trials from the ACCENT Group. *J Clin Oncol* 2007;**25**:4569–74.
61. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008;**27**:157–72; discussion 207.
62. Vergouwe Y, Royston P, Moons KG, Altman DG. Development and validation of a prediction model with missing predictor data: a practical approach. *J Clin Epidemiol* 2010;**63**:205–14.
63. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med* 2006;**25**:127–41.
64. Royston P, Lambert PC. *Flexible Parametric Survival Analysis Using Stata: Beyond the Cox Model*. College Station, TX: STATA Press; 2011.
65. Harrell FE. *Regression Modeling Strategies*. 2nd edn. Cham, Switzerland: Springer; 2006.
66. Blettner M, Sauerbrei W. Influence of model-building strategies on the results of a case-control study. *Stat Med* 1993;**12**:1325–38.
67. Royal College of Pathologists. *Standards and Datasets for Reporting Cancers: Dataset for Histopathological Reporting of Colorectal Cancer*. London: RCR; 2018.

68. Riley RD, Van de Windt D, Croft P, Moons KGM. *Prognosis Research in Health Care: Concepts, Methods and Impact*. Oxford: Oxford University Press; 2019.
69. Cianchi F, Messerini L, Comin CE, Boddi V, Perna F, Perigli G, Cortesini C. Pathologic determinants of survival after resection of T3N0 (Stage IIA) colorectal cancer: proposal for a new prognostic model. *Dis Colon Rectum* 2007;**50**:1332–41.
70. Lu HJ, Lin JK, Chen WS, Jiang JK, Yang SH, Lan YT, *et al*. The prognostic role of para-aortic lymph nodes in patients with colorectal cancer: is it regional or distant disease? *PLOS ONE* 2015;**10**:e0130345.
71. Peng J, Ding Y, Tu S, Shi D, Sun L, Li X, *et al*. Prognostic nomograms for predicting survival and distant metastases in locally advanced rectal cancers. *PLOS ONE* 2014;**9**:e106344.
72. Weiser MR, Gönen M, Chou JF, Kattan MW, Schrag D. Predicting survival after curative colectomy for cancer: individualizing colon cancer staging. *J Clin Oncol* 2011;**29**:4796–802.
73. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med* 2015;**162**:55–63.
74. Goh V, Bartram C, Halligan S. Effect of intravenous contrast agent volume on colorectal cancer vascular parameters as measured by perfusion computed tomography. *Clin Radiol* 2009;**64**(4):368–72.
75. Goh V, Halligan S, Bartram CI. Quantitative tumor perfusion assessment with multidetector CT: are measurements from two commercial software packages interchangeable? *Radiology* 2007;**242**(3):777–82.
76. Goh V, Halligan S, Gartner L, Bassett P, Bartram CI. Quantitative colorectal cancer perfusion measurement by multidetector-row CT: does greater tumour coverage improve measurement reproducibility? *Brit J Radiol* 2006;**79**(943):578–83.
77. Goh V, Halligan S, Hugill JA, Bassett P, Bartram CI. Quantitative assessment of colorectal cancer perfusion using MDCT: inter- and intraobserver agreement. *AJR Am J Roentgenol* 2005;**185**(1):225–31.
78. Goh V, Halligan S, Gharpuray A, Wellsted D, Sundin J, Bartram CI. Quantitative assessment of colorectal cancer tumor vascular parameters by using perfusion CT: influence of tumor region of interest. *Radiology* 2008;**247**(3):726–32.
79. Ng CS, Chandler AG, Wei W, Anderson EF, Herron DH, Kurzrock R, *et al*. Effect of duration of scan acquisition on CT perfusion parameter values in primary and metastatic tumors in the lung. *Eur J Radiol* 2013;**82**(10):1811–8.
80. Ng CS, Chandler AG, Wei W, Anderson EF, Herron DH, Kurzrock R, *et al*. Effect of sampling frequency on perfusion values in perfusion CT of lung tumors. *AJR Am J Roentgenol* 2013;**200**(2):W155–62.
81. Ng CS, Wei W, Ghosh P, Anderson E, Herron DH, Chandler AG. Observer Variability in CT Perfusion Parameters in Primary and Metastatic Tumors in the Lung. *Technol Cancer Res Treat* 2018;**17**:1533034618769767.
82. Ng CS, Chandler AG, Wei W, Herron DH, Anderson EF, Kurzrock R, *et al*. Reproducibility of CT perfusion parameters in liver tumors and normal liver. *Radiology* 2011;**260**(3):762–70.
83. Holleb AI, Folkman J. Tumor angiogenesis. *CA Cancer J Clin* 1972;**22**:226–9.
84. Folkman J. Tumor angiogenesis: therapeutic implications. *N Engl J Med* 1971;**285**:1182–6.
85. Höckel M, Vaupel P. Biological consequences of tumor hypoxia. *Semin Oncol* 2001;**28**:36–41.
86. Vaupel P. Tumor microenvironmental physiology and its implications for radiation oncology. *Semin Radiat Oncol* 2004;**14**:198–206.
87. Carmeliet P, Jain RK. Molecular mechanisms and clinical applications of angiogenesis. *Nature* 2011;**473**:298–307.
88. Staton CA, Chetwood AS, Cameron IC, Cross SS, Brown NJ, Reed MW. The angiogenic switch occurs at the adenoma stage of the adenoma carcinoma sequence in colorectal cancer. *Gut* 2007;**56**:1426–32.

89. Konerding MA, Fait E, Gaumann A. 3D microvascular architecture of pre-cancerous lesions and invasive carcinomas of the colon. *Br J Cancer* 2001;**84**:1354–62.
90. Denko NC. Hypoxia, HIF1 and glucose metabolism in the solid tumour. *Nat Rev Cancer* 2008;**8**:705–13.
91. Talbot IC, Ritchie S, Leighton M, Hughes AO, Bussey HJ, Morson BC. Invasion of veins by carcinoma of rectum: method of detection, histological features and significance. *Histopathology* 1981;**5**:141–63.
92. Ryan R, Gibbons D, Hyland JM, Treanor D, White A, Mulcahy HE, *et al.* Pathological response following long-course neoadjuvant chemoradiotherapy for locally advanced rectal cancer. *Histopathology* 2005;**47**:141–6.
93. Adibi A, Sadatsafavi M, Ioannidis JPA. Validation and utility testing of clinical prediction models: time to change the approach. *JAMA* 2020;**324**:235–6.
94. Simon R, Altman DG. Statistical aspects of prognostic factor studies in oncology. *Br J Cancer* 1994;**69**:979–85.
95. Volkmann A, De Bin R, Sauerbrei W, Boulesteix AL. A plea for taking all available clinical information into account when assessing the predictive value of omics data. *BMC Med Res Methodol* 2019;**19**:162.
96. Melander O, Newton-Cheh C, Almgren P, Hedblad B, Berglund G, Engstrom G, *et al.* Novel and conventional biomarkers for prediction of incident cardiovascular events in the community. *JAMA* 2009;**302**:49–57.

Appendix 1 Supplemental tables

TABLE 21 Union for International Cancer Control/AJCC staging classification for colorectal cancer (8th edition)

AJCC stage	Stage grouping	Description
I	T1 or T2	Cancer has grown through the muscularis mucosa into the submucosa (T1), or into the muscularis propria (T2). There is no spread to nearby lymph nodes (N0) or to distant sites (M0)
	N0	
	M0	
IIA	T3	Cancer has grown into the outermost layers of the colon or rectum but has not gone through them (T3). There is no spread to nearby lymph nodes (N0) or to distant sites (M0)
	N0	
	M0	
IIB	T4a	Cancer has grown through the wall of the colon or rectum but has not grown into other nearby tissues or organs (T4a). There is no spread to nearby lymph nodes (N0) or to distant sites (M0)
	N0	
	M0	
IIC	T4b	Cancer has grown through the wall of the colon or rectum and into other nearby tissues or organs (T4b). There is no spread to nearby lymph nodes (N0) or to distant sites (M0)
	N0	
	M0	
IIIA	T1 or T2	Cancer has grown through the muscularis mucosa into the submucosa (T1), or into the muscularis propria (T2). Cancer has spread to 1–3 nearby lymph nodes (N1) or into areas of fat near the lymph nodes but not the nodes themselves (N1c). It has not spread to distant sites (M0)
	N1/N1c	
	M0	
	<i>or</i>	
	T1	
	N2a	
M0		
IIIB	T3 or T4a	Cancer has grown into the outermost layers of the colon or rectum (T3) or through the visceral peritoneum (T4a) but has not reached nearby organs. Cancer has spread to 1–3 nearby lymph nodes (N1a or N1b) or into areas of fat near the lymph nodes but not the nodes themselves (N1c). It has not spread to distant sites (M0)
	N1/N1c	
	M0	
	<i>or</i>	
	T2 or T3	
	N2a	
	M0	
	<i>or</i>	
	T1 or T2	
N2b		
M0		

TABLE 21 Union for International Cancer Control/American Joint Committee on Cancer staging classification for colorectal cancer (8th edition) (continued)

AJCC stage	Stage grouping	Description	
IIIC	T4a	Cancer has grown through the wall of the colon or rectum (including the visceral peritoneum) but has not reached nearby organs (T4a). Cancer has spread to 4–6 nearby lymph nodes (N2a). It has not spread to distant sites (M0)	
	N2a		
	M0		
		or	
	T3 or T4a	Cancer has grown into the outermost layers of the colon or rectum (T3) or through the visceral peritoneum (T4a) but has not reached nearby organs. Cancer has spread to seven or more nearby lymph nodes (N2b). It has not spread to distant sites (M0)	
	N2b		
	M0		
		or	
	T4b	Cancer has grown through the wall of the colon or rectum and is attached to or has grown into other nearby tissues or organs (T4b). Cancer has spread to at least one nearby lymph node or into areas of fat near the lymph nodes (N1 or N2). It has not spread to distant sites (M0)	
N1 or N2			
M0			
IVA	Any T	Cancer has spread to one distant organ (such as the liver or lung) or distant set of lymph nodes, but not to distant parts of the peritoneum (the lining of the abdominal cavity) (M1a)	
	Any N		
	M1a		
IVB	Any T	Cancer has spread to more than one distant organ (such as the liver or lung) or distant set of lymph nodes, but not to distant parts of the peritoneum (the lining of the abdominal cavity) (M1b)	
	Any N		
	M1b		
IVC	Any T	Cancer has spread to distant parts of the peritoneum (the lining of the abdominal cavity) and may or may not have spread to distant organs or lymph nodes (M1c)	
	Any N		
	M1c		

TABLE 22 List of participating sites and site principal investigators

Site	Local investigator	Cancer network
Bradford Teaching Hospitals NHS Foundation Trust	A Lowe, A Mohammed	Yorkshire
Guys and St Thomas's NHS -Foundation Trust	N Griffin, S Gourtsoyiani	South East London
NHS Lothian	J Brush	Scottish
University Hospital of North Staffordshire NHS Trust	I Britton	Greater Midlands
Oxford University Hospitals NHS Foundation Trust	A Slater	Thames Valley
Portsmouth Hospitals University NHS Trust ^a	A Higginson	Central South Coast
Royal Cornwall Hospitals NHS Trust	M Strugnell	Peninsula
Sheffield Teaching Hospitals NHS Foundation Trust	R Vinayagam	North Trent

continued

TABLE 22 List of participating sites and site principal investigators (continued)

Site	Local investigator	Cancer network
University Hospital Southampton NHS FoundationTrust	D Breen, C Grierson	Central South Coast
NHS Tayside	I Zealley	Scottish
York Hospitals NHS Foundation Trust	R Mannion	Yorkshire
University Hospitals Coventry and Warwickshire NHS Trust	P Correa	Warwickshire
Leeds Teaching Hospitals NHS Trust	D Tolan	Yorkshire

a Pilot sites.

TABLE 23 Recruitment sites and multidetector CT capability

Recruitment site	CT vendor	Detector rows	Number of scanners
Bradford Teaching Hospitals NHS Foundation Trust	GE	16–64	2
Guys and St Thomas's NHS Foundation Trust	Phillips	16–256	3
Leeds Teaching Hospitals NHS Trust	Siemens	64	2
NHS Lothian	Toshiba	320	1
University Hospital of North Staffordshire NHS Trust	Siemens	16–128	3
Oxford University Hospitals NHS Foundation Trust	GE	16–64	> 3
Portsmouth Hospitals University NHS Trust	Siemens	128	3
Royal Cornwall Hospitals NHS Trust	GE	16–64	3
Sheffield Teaching Hospitals NHS Foundation Trust	GE	32–64	> 3
University Hospital Southampton NHS Foundation Trust	GE	64–128	2
NHS Tayside	Siemens	128	3
York Hospitals NHS Foundation Trust	Siemens	16–64	1
University Hospitals Coventry and Warwickshire NHS Trust	GE	64	2

TABLE 24 Perfusion CT models and vascular parameters

Kinetic model	Compartments	Assumptions	Quantitative parameter
Distributed parameter	Dual	Constrained input function	Blood flow, blood volume, permeability surface area product
Patlak	Dual	One-way transfer Well-mixed compartments	Extraction fraction (permeability surface area product), blood volume
Deconvolution	Single	Instantaneous input function	Blood flow, blood volume, mean transit time
Maximum slope	Single	No venous outflow	Blood flow

TABLE 25 Definitions for end of period for disease-free survival

Situation	Date	Outcome
Annual surveillance CT scan shows metastasis	Date of first scan showing recurrence	Recurrence
Chest X-ray abnormal, clinic visit, multiple imaging, some showing metastasis		
CT scan – scheduled or unscheduled	Date of CT scan showing recurrence	Recurrence
No CT scan	Date of subsequent clinic visit where decision made on basis of recurrence event	
Clinical suspicion (e.g. CEA raised) but no FU imaging and no CT scan		
Hospital visit	Date of subsequent clinic visit where decision made on basis of recurrence	Recurrence Recurrence censored at last scan ^a
GP visit only (likely only very elderly participants, who are unable to attend hospital)	GP – individual patient follow-up would be difficult and laborious	Death from ONS
Loss to follow-up or patient withdraws consent	Date of last CT scan or baseline	Censored recurrence Death from ONS+
Clinical suspicion but patient too ill to attend any tests; no CT scan		
Hospital visit	Date of clinic visit or inpatient admission when decision made that patient is too ill to attend for CT	Censored recurrence Death from ONS
GP visit only	Hospice admission or GP visit	
GP, general practitioner; ONS: Office for National Statistics. a This censoring may be informative: we hope numbers are low and would not affect trial results. A sensitivity analysis could be performed including these participants but would require follow-up with individual patient GP.		

TABLE 26 Definitions employed for high-risk (stage III) disease

Stage	Definition
Stage IIIA	T1, N1, M0 or T2, N1, M0: the cancer has grown through the mucosa into the submucosa (T1) or it may also have grown into the muscularis propria (T2). It has spread to 1–3 nearby lymph nodes, but not to distant sites
Stage IIIB	T3, N1, M0 or T4, N1, M0: the cancer has grown into the outermost layers of the colon or rectum but has not reached nearby organs (T3), or the cancer has grown through the wall of the colon or rectum and into other nearby tissues or organs (T4). It has spread to 1–3 nearby lymph nodes but not distant sites
Stage IIIC	Any T, N2, M0: the cancer may or may not have grown through the wall of the colon or rectum, but it has spread to 4 or more nearby lymph nodes. It has not spread to distant sites

TABLE 27 Summary of treatments that participants received

	Recurrence, N (%)	No recurrence, N (%)	Total, N (%)
Treatment^{a,b,c}			
No treatment	5 (6)	0 (0)	5 (2)
Treatment	76 (94)	245 (100)	321 (98)
Missing data	0 (0)	0 (0)	0 (0)
Total	81 (100)	245 (100)	326 (100)

continued

TABLE 27 Summary of treatments that participants received (continued)

	Recurrence, N (%)	No recurrence, N (%)	Total, N (%)
Neoadjuvant therapy ± surgery			
No neoadjuvant therapy	187 (76)	59 (74)	246 (76)
Radiotherapy and surgery	6 (7)	14 (6)	20 (6)
Radiotherapy only	2 (2)	0 (0)	2 (1)
Chemotherapy and surgery	0 (0)	9 (4)	9 (3)
Chemotherapy only	1 (1)	0 (0)	1 (0)
Chemoradiotherapy and surgery	8 (10)	29 (12)	37 (11)
Chemoradiotherapy only	5 (6)	4 (2)	9 (3)
Unspecified, no surgery	0 (0)	1 (0)	1 (0)
Missing data	0 (0)	1 (0)	1 (0)
Surgery			
No surgery	12 (15)	5 (2)	17 (5)
Surgery only	32 (39)	118 (48)	150 (46)
Surgery with both neoadjuvant and adjuvant therapy	3 (4)	26 (11)	29 (9)
Surgery with neoadjuvant therapy only	11 (14)	25 (10)	36 (11)
Surgery with adjuvant therapy only	22 (27)	70 (29)	92 (28)
Unspecified surgery		1 (0)	1(0)
Missing data	1 (0)		1 (0)
Adjuvant therapy			
No adjuvant therapy	53 (66)	146 (59)	199 (61)
Prior neoadjuvant therapy and surgery	3 (4)	26 (11)	29 (9)
Prior neoadjuvant therapy only	2 (2)	2 (1)	4 (1)
Prior surgery only	22 (27)	70 (29)	92 (28)
Missing data	1 (1)	1 (0)	2 (1)

a Participants can appear in more than one category.
b Four participants received no treatment.
c One patient missing all treatment.

TABLE 28 Local site perfusion CT measurements by tumour T stage

Local review	Tumour confined to bowel wall (T1, T2)				Tumour extending beyond bowel (T3, T4)			
	Participants with data (n)	Mean (SD)	Median (IQR)	Range	Participants with data (n)	Mean (SD)	Median (IQR)	Range
Blood flow (ml/minute/100 ml or 100 g)	94	67.2 (38.6)	59.8 (42.8–80.9)	0.0–231.1	209	71.5 (36.2)	64.3 (50.7–84.2)	23.0–350.8
Blood volume (ml/100 ml or 100 g)	92	11.9 (7.4)	11.1 (5.9–15.9)	0.0–44.4	204	13.5 (7.8)	12.6 (7.4, 16.8)	0.6, 46.7
Permeability surface area product (ml/minute/100 ml or 100 g)	88	15.8 (13.3)	12.6 (7.9–17.2)	0.0–72.1	181	17.5(14)	13.4(9.1–18.8)	0.0–66.8
Mean transit time (seconds)	79	13.9(6.4)	13(8.7–18.4)	(4.6–33.2)	175	13.5 (5.5)	13.1(9.3–17.7)	3.4–33.6

TABLE 29 Local site perfusion CT measurements for node-positive and node-negative cancers

Local review Variable	Node positive (N1, N2)				Node negative (N0)			
	Participants with data (n)	Mean (SD)	Median (IQR)	Range	Participants with data (n)	Mean (SD)	Median (IQR)	Range
Blood flow (ml/minute/100 ml or 100 g)	140	64.5 (25.2)	60.8 (48.0–80.1)	17.3–155.8	163	75 (44.1)	64.3 (49.6–86.9)	0–350.8
Blood volume (ml/100 ml or 100 g)	136	12.9 (7.4)	12.5 (7.0–16.1)	3.3–46.7	160	13 (7.9)	11.8 (6.6–16.8)	0–45.5
Permeability surface area product (ml/minute/100 ml or 100 g)	123	16.3 (12.2)	13.3 (9.1–16.9)	0.0–57.3	146	17.4 (15)	13.1 (8.8–19.4)	0.0–72.1
Mean transit time (seconds)	118	13.9 (5.5)	13.4 (9.7–18.0)	4.7–33.6	136	13.4 (6)	12.9 (8.4–18.2)	3.4–33.2

TABLE 30 Central review of perfusion CT measurements by tumour T stage

Central review Variable	Tumour confined to bowel wall (T1, T2)				Tumour extending beyond bowel (T3, T4)			
	Participants with data (n)	Mean (SD)	Median (IQR)	Range	Participants with data (n)	Mean (SD)	Median (IQR)	Range
Blood flow (ml/minute/100 ml or 100 g)	86	59.6 (35.5)	52.2 (33.6–75.1)	10.8–211.1	205	62.1 (30.4)	57.9 (43.7–73.4)	7.1–222.8
Blood volume (ml/100 ml or 100 g)	85	10.1 (6.0)	9 (5.5–13.5)	1.9–32.8	205	11.2 (6.6)	10.4 (6.7–14)	0–46.3
Permeability surface area product (ml/minute/100 ml or 100 g)	78	18.5 (14.1)	13.6 (9.6–23.5)	0–69.5	177	19.8 (15.7)	13.6 (8.6–28.9)	0–76.9
Mean transit time (seconds)	78	11 (6.4)	8.4 (6.2–16.1)	2.8–30.6	174	11 (6)	9.3 (6.5–15.0)	0–26.4

TABLE 31 Central review of perfusion CT measurements for node-positive and node-negative cancers

Central review Variable	Node positive (N1, N2)				Node negative (N0)			
	Participants with data (n)	Mean (SD)	Median (IQR)	Range	Participants with data (n)	Mean (SD)	Median (IQR)	Range
Blood flow (ml/minute/100 ml or 100 g)	134	57.4 (26.7)	53.7 (39.9–72.5)	7.1–150.1	157	64.7 (35.6)	58.2 (43.3–79.3)	13.2–222.8
Blood volume (ml/100 ml or 100 g)	133	10.5 (6)	9.9 (6.4–12.7)	0.5–39.8	157	11.1 (6.8)	9.9 (6.1–14.6)	0–46.3
Permeability surface area product (ml/minute/100 ml or 100 g)	118	18.9 (15.4)	12.9 (8.7–24.1)	0–76.9	137	19.9 (15.1)	13.9 (9–28.8)	0–69.5
Mean transit time (seconds)	117	11.3 (6.3)	9.7 (6.8–15.0)	0–30.6	135	10.7 (6.0)	8.5 (6.3–16.0)	0–30.5

TABLE 32 Central review of perfusion CT measurements for participants with and without recurrence

Central review Variable	Recurrence				No recurrence			
	Participants with data (n)	Mean (SD)	Median (IQR)	Range	Participants with data (n)	Mean (SD)	Median (IQR)	Range
Blood flow (ml/minute/100 ml or 100g)	72	60.3 (24.2)	58.5 (43.8–73.3)	13.2–127.0	219	61.7 (34.2)	54.9 (39.7–76.7)	7.1–222.8
Blood volume (ml/100 ml or 100g)	72	10.3 (5.9)	9.8 (6.2–13.8)	0–30.2	218	11 (6.6)	10.3 (6.4–13.9)	0.5–46.3
Permeability surface area product (ml/minute/100 ml or 100g)	63	21.8 (15.9)	14.8 (10–32)	0–76.9	192	18.7 (14.9)	12.8 (8.6–25.4)	0–69.5
Mean transit time (seconds)	63	10.8 (6.6)	7.7 (6.3–14.1)	0–30.5	189	11.1 (6)	9.3 (6.4–16)	0–30.6

TABLE 33 CD105 expression (number of CD105 stained vessels/mm² field) in participants with and without recurrence

Variable	Recurrence				No recurrence			
	Participants with data (n)	Mean (SD)	Median (IQR)	Range	Participants with data (n)	Mean (SD)	Median (IQR)	Range
CD105	58 ^a	126 (65)	113 (85–166)	35–357	199 ^b	125 (63)	116 (82–57)	27–393

a Data missing, n = 4.
b Data missing, n = 13.

TABLE 34 Immunohistochemical scores for HIF-1 α , VEGF and GLUT-1 in participants with and without recurrence

Variable score	Recurrence, N (%)	No recurrence, N (%)	Total, N (%)
HIF-1α			
0	18 (29)	58 (27)	76 (28)
1	5 (8)	24 (11)	29 (11)
2	11 (18)	35 (17)	46 (17)
3	1 (2)	4 (2)	5 (2)
4	11 (18)	40 (19)	51 (19)
6	15 (24)	46 (22)	61 (22)
Missing data	1 (2)	5 (2)	6 (2)
VEGF			
0	55 (89)	199 (94)	254 (93)
2	2 (3)	4 (2)	6 (2)
3	2 (3)	3 (1)	5 (2)
4	2 (3)	2 (1)	4 (1)
Missing data	1 (2)	4 (2)	5 (2)
GLUT-1			
0	2 (3)	10 (5)	12 (4)
1	0 (0)	1 (0)	1 (0)

TABLE 34 Immunohistochemical scores for HIF-1 α , VEGF and GLUT-1 in participants with and without recurrence (*continued*)

Variable score	Recurrence, N (%)	No recurrence, N (%)	Total, N (%)
2	0 (0)	5 (2)	5 (2)
3	2 (3)	8 (4)	10 (4)
4	4 (6)	13 (6)	17 (6)
5	5 (8)	18 (8)	23 (8)
6	18 (29)	56 (26)	74 (27)
7	10 (16)	46 (22)	56 (20)
8	20 (32)	51 (24)	71 (26)
Missing data	1 (2)	4 (2)	5 (2)

TABLE 35 Frequency of genetic mutations (MMR, KRAS, HRAS, NRAS and BRAF) for participants with and without recurrence

Gene mutation	Recurrence, N (%)	No recurrence, N (%)	Total, N (%)
MMR			
Deficient	5 (8)	16 (8)	21 (8)
Proficient	56 (90)	192 (91)	248 (91)
Missing data	1 (2)	4 (2)	5 (2)
KRAS			
Wild type	34 (55)	96 (45)	130 (47)
Mutation	28 (45)	112 (53)	140 (51)
Missing data	0 (0)	4 (2)	4 (1)
NRAS			
Wild type	55 (89)	188 (89)	243 (89)
Mutation	7 (11)	20 (9)	27 (10)
Missing data	0 (0)	4 (2)	4 (1)
HRAS			
Wild type	52 (84)	184 (87)	236 (86)
Mutation	10 (16)	24 (11)	34 (12)
Missing data	0 (0)	4 (2)	4 (1)
BRAF_V600E			
Wild type	59 (95)	192 (91)	251 (92)
Mutation	3 (5)	16 (8)	19 (7)
Missing data	0 (0)	4 (2)	4 (1)
BRAF_other			
Wild type	53 (85)	172 (81)	225 (82)
Mutation	9 (15)	36 (17)	45 (16)
Missing data	0 (0)	4 (2)	4 (1)

TABLE 36 Mean (SD) image noise obtained from the bladder ROI per participating centre and CT scanner model

Participating centre	Scanner model	Scans (n)	Noise	
			Mean (HU)	SD (HU)
01	GE Lightspeed VCT	11	24.1	1.7
03	Siemens Sensation 16	3	25.3	3.7
03	Siemens Definition AS+	9	17.3	3.5
04	GE Lightspeed VCT	16	20.4	2.6
05	Siemens Definition AS+	4	17.0	2.1
07	GE Lightspeed VCT	1	20.1	-
08	GE Discovery CT750 HD	15	20.3	1.5
09	Siemens Definition AS+	9	20.4	5.2
11	Siemens Sensation 64	13	24.0	5.9

TABLE 37 Summary of peak arterial iodine enhancement and full-width half maximum of the initial arterial enhancement peak

Participating centre	Mean	Minimum	Scanner model	Peak arterial enhancement			Full width half maximum
				Maximum	Mean	Minimum	Maximum
<i>GE scanners</i>							
01	GE Lightspeed VCT	395	314	458	9.6	7.0	12.0
04	GE Lightspeed VCT	405	316	546	13.8	8.0	29.0
07	GE Lightspeed VCT	490	490	490	9.0	9.0	9.0
08	GE Discovery CT750 HD	423	316	546	10.5	7.0	26.0
<i>Siemens scanners</i>							
03	Siemens Sensation 16	285	409	356	8.7	8.0	9.0
03	Siemens Definition AS+	292	189	420	11.1	9.0	15.0
05	Siemens Definition AS+	336	259	476	9.5	8.0	10.0
09	Siemens Definition AS+	279	210	330	11.1	8.0	15.0
11	Siemens Sensation 64	331	249	396	10.5	9.0	12.0

TABLE 38 Summary of timing of follow-up visits

Follow-up visit	Participants (n)	Median (IQR) (days)	Range (days)
Clinic visit 1	311	127 (97-172)	5-289
Clinic visit 2	305	321 (251-369)	147-658
Clinic visit 3	260	498 (456-561)	337-874
Clinic visit 4	269	729 (646-778)	279-995
Clinic visit 5	244	1072 (952-1137)	685-1455
CT imaging 1	296	391 (358-419)	115-668
CT imaging 2	253	755 (725-778)	321-975
CT imaging 3	225	1125 (1074-1151)	756-1478
Colonoscopy	35	1027 (826-1200)	494-1320

TABLE 39 Location of disease recurrence

Location of recurrence	N (%)
Distant metastases	
Liver	28 (53)
Lung	16 (31)
Peritoneal	2 (4)
Retroperitoneal	2 (4)
Soft tissue	2 (4)
Brain	1 (2)
Renal	1 (2)
Total	52 (100)
New primary	
Lung	2 (17)
Prostate	2 (17)
Breast	2 (17)
Colorectal	2 (17)
Bladder	1 (8)
Melanoma	1 (8)
Renal	1 (8)
Squamous cancer (skin)	1 (8)
Total	12 (100)

TABLE 40 Summary of serum CEA levels at baseline and subsequent follow-up clinic visits

Clinic visit	Recurrence		No recurrence		Total	
	Participants with data (n)	Median (IQR), range	Participants with data (n)	Median (IQR), range	Participants with data (n)	Median (IQR), range
Baseline	61	4.0 (1.7–6.9), 0.7–222.0	183	2.4 (1.5–5.9), 0.4–124.0	244	2.5 (1.6–6.0), 0.4–222.0
1	31	3.0 (1.5–8.0), 0.5–37.5	114	1.7 (1.0–2.4), 0.3–22.8	145	1.9 (1.0–3.0), 0.3–37.5
2	34	2.0 (1.4–4.8), 0.5–930.8	145	1.5 (1.0–2.2), 0.5–106.0	179	1.6 (1.0–2.6), 0.5–930.8
3	25	2.0 (1.0–4.7), 0.5–14.2	125	1.4 (1.0–2.0), 0.0–8.2	150	1.5 (1.0–2.3), 0.0–14.2
4	30	3.0 (2.0–6.0), 0.5–81.3	129	1.7 (1.0–2.3), 0.0–10.6	159	2.0 (1.0–2.8), 0.0–81.3
5	26	2.5 (2.0–5.1), 1.0–473.0	122	1.6 (1.0–2.3), 0.5–20.0	148	1.8 (1.0–2.5), 0.5–473.0

TABLE 41 Summary of candidate variables included in model B

Variable	Group	Participants with CT measurements (included in model B)	Participants without CT measurements (not in model B)	All patients
		N (%)	N (%)	N
T stage	1	92 (11)	8 (1)	12
	2	89 (77)	11 (10)	87
	3	84 (154)	16 (29)	183
	4	80 (35)	20 (9)	44
	Total	85 (277)	15 (49)	326
N stage	0	85 (149)	15 (26)	175
	1	85 (83)	15 (15)	98
	2	85 (45)	15 (8)	53
	Total	85 (277)	15 (49)	326
Venous invasion	Yes	83 (193)	17 (40)	233
	No	90 (84)	10 (9)	93
	Total	85 (277)	15 (49)	326
Treatment group	Surgery only	83 (124)	17 (26)	150
	Surgery and adjuvant therapy	88 (81)	12 (11)	92
	Neoadjuvant therapy and surgery	84 (56)	16 (11)	67
	No surgery	94 (16)	6 (1)	17
	Total	85 (277)	15 (49)	326
Primary tumour location	Colon	79 (113)	21 (30)	143
	Rectum	90 (164)	10 (19)	183
	Left colon	85 (228)	15 (40)	268
	Right colon	84 (49)	16 (9)	58
	Total	85 (277)	15 (49)	326
Primary tumour size (mm)	10–29	88 (59)	12 (8)	67
	30–39	82 (61)	18 (13)	74
	40–49	84 (61)	16 (12)	73
	50–150	86 (96)	14 (16)	112
	Total	85 (277)	15 (49)	326
Age at diagnosis (years)	< 50	81 (17)	19 (4)	21
	50–59	92 (47)	8 (4)	51
	60–69	80 (103)	20 (25)	128
	70–79	86 (86)	14 (14)	100

TABLE 41 Summary of candidate variables included in model B (continued)

Variable	Group	Participants with CT measurements (included in model B)	Participants without CT measurements (not in model B)	All patients
		N (%)	N (%)	N
	> 80	92 (24)	8 (2)	26
	Total	85 (277)	15 (49)	326
Sex	Female	85 (191)	15 (35)	226
	Male	86 (86)	14 (14)	100
	Total	85 (277)	15 (49)	326

TABLE 42 Correlation coefficients between local hospital perfusion CT variables

Variables	(1)	(2)	(3)	(4)
(1) Blood flow	1.000			
(2) Blood volume	0.411	1.000		
(3) Permeability	0.355	-0.001	1.000	
(4) Transit time	-0.285	0.471	-0.364	1.000

TABLE 43 Univariable hazard ratios for perfusion CT variables from central review

Variable	N	Hazard ratio	Hazard ratio, 95% CI lower	Hazard ratio, 95% CI upper	p-value
Blood volume	311	0.99	0.95	1.02	0.41
Blood flow	313	1.00	0.99	1.01	0.84
Permeability	277	1.01	1.00	1.03	0.12
Mean transit time	277	0.99	0.95	1.03	0.53
Central PCA variables^a					
PCA1	277	0.91	0.76	1.08	0.27
PCA2	277	1.04	0.87	1.25	0.67

a PCA1 and PCA2 represents first and second eigenvalue from PCA.

TABLE 44 Multivariable-adjusted hazard ratios for model E with model A as reference

	Model A				Model E			
	Hazard ratio	Hazard ratio, 95% CI lower	Hazard ratio, 95% CI upper	p-value	Hazard ratio	Hazard ratio, 95% CI lower	Hazard ratio, 95% CI upper	p-value
ModelA_324	1.000				1.000			
PCA1 score ^a					-0.048	-0.222	0.126	0.59
PCA2 score ^a					0.073	-0.111	0.256	0.44
Baseline survival (rcs1)	0.820	0.652	0.989	< 0.001	0.822	0.653	0.991	< 0.001
Constant	-4.577	-4.812	-4.341	< 0.001	-4.581	-4.818	-4.344	< 0.001
Log likelihood	-176.84				-176.42			
In model (n)	277				277			

a PCA1 and PCA2 represent first and second eigenvalues from PCA.

TABLE 45 Summary of model variables of participants in model F

Variable	Group	Participants with pathology measurements (included in model F, N = 212)	Participants without pathology measurements (not in model F, N = 112)	All patients (N = 324)
T stage	1	9 (4)	3 (3)	12 (4)
	2	57 (27)	30 (27)	87 (27)
	3	119 (56)	62 (55)	181 (56)
	4	27 (13)	17 (15)	44 (14)
N stage	0	116 (55)	59 (53)	175 (54)
	1	61 (29)	36 (32)	97 (30)
	2	35 (17)	17 (15)	52 (16)
Extramural venous invasion	Yes	143 (67)	89 (79)	232 (72)
	No	69 (33)	23 (21)	92 (28)
Treatment group	Surgery only	106 (50)	43 (38)	149 (46)
	Surgery and adjuvant therapy	72 (34)	20 (18)	92 (28)
	Neoadjuvant therapy and surgery	34 (16)	33 (29)	67 (21)
	No surgery	0 (0)	16 (14)	16 (5)
Primary tumour location	Colon	94 (44)	48 (43)	142 (44)
	Rectum	118 (56)	64 (57)	182 (56)
	Left colon	168 (79)	98 (88)	266 (82)
	Right colon	44 (21)	14 (13)	58 (18)

TABLE 45 Summary of model variables of participants in model F (continued)

Variable	Group	Participants with pathology measurements (included in model F, N = 212)	Participants without pathology measurements (not in model F, N = 112)	All patients (N = 324)
Primary tumour size (mm)	10–29	47 (22)	20 (18)	67 (21)
	30–39	49 (23)	23 (21)	72 (22)
	40–49	46 (22)	27 (24)	73 (23)
	50–150	70 (33)	42 (38)	112 (35)
Age at diagnosis (years)	< 50	13 (6)	8 (7)	21 (6)
	50–59	36 (17)	14 (13)	50 (15)
	60–69	81 (38)	46 (41)	127 (39)
	70–79	63 (30)	37 (33)	100 (31)
Sex	> 80	19 (9)	7 (6)	26 (8)
	Female	69 (33)	31 (28)	100 (31)
	Male	143 (67)	81 (72)	224 (69)

TABLE 46 Multivariable-adjusted hazard ratios for somatic mutation analysis variables

Variables	N	Hazard ratio (95% CI)	p-value
ModelA_324	212	2.718 constrained	
BRAF not 600	212	0.585 (0.244 to 1.403)	0.23
BRAF 600	212	0.371 (0.087 to 1.573)	0.18
KRAS	212	0.524 (0.287 to 0.955)	0.04
HRAS	212	1.806 (0.810 to 4.025)	0.15
NRAS	212	2.414 (0.998 to 5.836)	0.05
Baseline survival (rcs1)	212	2.177 (1.776 to 2.669)	< 0.001
Constant	212	0.013 (0.008 to 0.019)	< 0.001

TABLE 47 Difference in sensitivity and specificity for model B compared with rule C

Model	Sensitivity (95% CI)	Specificity (95% CI)	True positives ^a (of 330 with recurrence) (95% CI)	True negatives ^a (of 670 with no recurrence) (95% CI)
Model B, high vs. medium/low risk	0.59 (0.46 to 0.70)	0.75 (0.68 to 0.81)	193 (152 to 231)	503 (456 to 543)
Rule C ^b	0.54 (0.42 to 0.66)	0.57 (0.50 to 0.64)	178 (139 to 218)	382 (335 to 429)
Difference	0.05 (–0.09 to 0.18)	0.18 (0.09 to 0.27)	17 (–30 to 59)	121 (60 to 181)

a Individual model true-positive and true-negative numbers are rounded to reflect differences in true positive and true negative.

b Rule C in 277 patients has similar but slightly lower sensitivity than in 324 patients.

TABLE 48 Sensitivity and specificity analysis for model B (alternative threshold) compared with Rule C

Model	Risk group	Survival (95% CI)	Total	Recurrences	No recurrences	Sensitivity	Specificity
						(95% CI), n/N	(95% CI), n/N
Model B, sensitivity	High/medium risk (top 66%)	0.671	185	63	122	0.90	0.41
	High/med vs. low risk	(0.595 to 0.736)				(0.80 to 0.96)	(0.34 to 0.48)
	Low risk	0.941	92	7	85	(63/70)	(85/207)
		(0.861 to 0.975)					
	Total PCT		277	70	207		
	Missing		49	11	38		
	Total		326	81	245		
Rule C	High risk	0.713	127	38	89	0.54	0.57
	Group 2 : 2 risk groups	(0.623 to 0.786)				(0.42 to 0.66)	(0.50 to 0.64)
	Low risk	0.803	150	32	118	(38/70)	(118/207)
		(0.727 to 0.860)					
	Total model B		277	70	207		
	Missing PCT		49	11	36		
	Total		326	81	247		

PCT, perfusion computed tomography.

TABLE 49 Difference in sensitivity and specificity for model B (alternative threshold) compared with rule C

Model	Sensitivity (95% CI)	Specificity (95% CI)	True positives (of 330 with recurrence) (95% CI)	True negatives ^a (of 670 with no recurrence) (95% CI)
Model B, high/med vs. low risk	0.90 (0.80 to 0.96)	0.41 (0.34 to 0.48)	296 (264 to 317)	275 (228 to 322)
Rule C ^b	0.54 (0.42 to 0.66)	0.57 (0.50 to 0.64)	178 (139 to 218)	382 (335 to 429)
Difference	0.36 (0.21 to 0.50)	-0.16 (-0.26 to -0.06)	118 (69 to 165)	-107 (-40 to -174)

a Individual model true-positive and true-negative numbers are rounded to reflect differences in true positive and true negative.

b Rule C on 277 patients has similar but slightly lower sensitivity than in 324 patients.

TABLE 50 Sensitivity and specificity analysis for model A (alternative threshold) compared with rule C

Model	Risk group	Survival (95% CI survival)	Total	Recurrences	No recurrences	Sensitivity (95% CI), (n/N)	Specificity (95% CI), (n/N)
Model A sensitivity analysis	High/medium risk (top 66%)	0.677 ^a	217	72	145	0.89	0.40
	Group 2 : 2 risk groups	(0.608 to 0.736)				(0.80 to 0.95)	(0.31 to 0.47)
	Low risk	0.93	107	9	98	(72/81)	(98/243)
		(0.858 to 0.966)					
	Missing		2	0	2		
	Total		326	81	245		
Rule C	High risk	0.706 ^a	148	45	103	0.56	0.58
	Group 2 : 2 risk groups	(0.624 to 0.774)				(0.44 to 0.67)	(0.51 to 0.64)
	Low risk	0.809 ^a	176	36	140	(45/81)	(140/243)
		(0.740 to 0.861)					
	Missing		2	0	2		
	Total		326	81	245		

a Survival calculated from Nelson–Aalen cumulative hazard at 3 years, as these estimates allow for censored data due to loss to follow-up. Survival probability multiplied for both high- and low-risk groups to standardise to a common value of 81 recurrences multiplied by 0.985.

TABLE 51 Difference in sensitivity and specificity for model A (alternative threshold) compared with rule C

Model (n = 324)	Sensitivity (95% CI)	Specificity (95% CI)	True positives (of 330 with recurrence) (95% CI)	True negatives (of 670 with no recurrence) (95% CI)
Model A, high/medium vs. low	0.89 (0.80 to 0.95)	0.40 (0.31 to 0.47)	294 (264 to 314)	268 (208 to 315)
Rule C	0.56 (0.44 to 0.67)	0.58 (0.51 to 0.64)	185 (145 to 221)	389 (342 to 429)
Difference	0.33 (0.21 to 0.46)	-0.18 (-0.08 to -0.26)	109 (69 to 152)	-121 (-54 to -174)

TABLE 52 Sensitivity and specificity analysis for model A compared with model B

Model	Risk group	Survival (95% CI survival)	Total	Recurrences	No Recurrences	Sensitivity (95% CI) (n/N)	Specificity (95% CI) (n/N)
Model B, group 2: 2 risk groups	High risk (top 33%)	0.573 ^a (0.462 to 0.670)	93	41	52	0.58 (0.46 to 0.70), (41/70)	0.75 (0.68 to 0.81), (155/207)
	Med/low risk	0.857 ^a (0.795 to 0.901)	184	29	155		
	Total model B		277	70	207		
	Missing PCT		49	11	36		
	Total		326	81	247		
Model A, group 2: 2 risk groups	High risk (top 33%)	0.575 ^a (0.465 to 0.671)	93	41	52	0.58 (0.46 to 0.70), (41/70)	0.75 (0.68 to 0.81), (155/207)
	Med/low risk	0.857 ^a (0.795 to 0.901)	184	29	155		
	Total		277	70	207		
	Missing		49	11	36		
	Total		326	81	247		

a Survival calculated from Nelson–Aalen cumulative hazard at 3 years, as these estimates allow for censored data due to loss to follow-up. Survival probability multiplied for both high- and low-risk groups to standardise to a common value of 70 recurrences – multiplied by 0.980.

TABLE 53 Difference in sensitivity and specificity for model A compared with model B

Model, n = 277	Sensitivity (95% CI)	Specificity (95% CI)	True positives (of 330 with recurrence) (95% CI)	True negatives (of 670 with no recurrence) (95% CI)
Model A (high vs. medium/low)	0.58 (0.46 to 0.70)	0.75 (0.68 to 0.81)	191 (152 to 231)	503 (456 to 543)
Model B (high vs. medium/low)	0.58 (0.46 to 0.70)	0.75 (0.68 to 0.81)	191 (152 to 231)	503 (456 to 543)
Difference	0.00 (–0.01 to 0.01)	0.00 (–0.02 to 0.02)	0 (–3 to 3)	0 (–13 to 13)

TABLE 54 Sensitivity and specificity analysis for model D compared with model B

Model	Risk group	Survival (95% CI survival)	Total	Recurrences	No recurrences	Sensitivity (95% CI) (n/N)	Specificity (95% CI) (n/N)
Model D	High risk (top 33%)	0.575 ^a (0.465 to 0.671)	93	41	52	0.58 ^b (0.46 to 0.70)	0.75 ^b (0.68 to 0.81) (155/207)
	Med/low risk	0.857 ^a (0.795 to 0.901)	184	29	155		
	Total PCT		277				
	Missing		49				
	Total		326				
Model B	High risk (top 33%)	0.573 ^a (0.462 to 0.670)	93	41	52	0.58 ^b (0.46 to 0.70) (41/70)	0.75 ^b (0.68 to 0.81) (155/207)
	Med/low risk	0.857 ^a (0.795 to 0.901)	184	29	155		
	Total PCT		277	70	207		
	Missing		49	11	38		
	Total		326	81	245		

a Adjusted as total number of events is 70.

b Estimates of sensitivity and specificity at 3 years is based on 63 patients with event at 3 years with local perfusion CT measurements. This is calculated from the Kaplan–Meier estimates of the number of events (see events in Kaplan–Meier graph number at risk). The modelling is based on 70 patients with events during the study period; 7 patients had an event between 3 years and their last study follow-up.

Note

NA = Nelson–Aalen cumulative hazard estimates are used to estimate the number of events at 3 years, as these estimates allow for censored data due to loss to follow-up.

TABLE 55 Sensitivity and specificity analysis for model E compared with model A

Model	Risk group	Survival (95% CI survival)	Total	Recurrences	No recurrences	Sensitivity (95% CI) (n/N)	Specificity (95% CI) (n/N)
Model E, group 2 : 2 risk groups	High risk (top 33%)	0.545 ^a (0.434 to 0.642)	93	43	50	0.61 ^b (0.49 to 0.73), (43/70)	0.75 ^b (0.69 to 0.82), (157/207)
	Med/low risk	0.873 ^a (0.813 to 0.915)	184	27	157		
	Total model E		277	70	207		
	Missing		49	11	38		
	Total		326	81	245		
Model A, group 2: 2 risk groups	High risk (top 33%)	0.575 ^a (0.465 to 0.671)	93	41	52	0.58 ^b (0.46 to 0.70), (41/70)	0.75 ^b (0.68 to 0.81), (155/207)
	Med/low risk	0.857 ^a (0.795 to 0.901)	184	29	155		
	Total model A		277	70	207		
	Missing		49	11	38		
	Total		326	81	245		

a Adjusted as total number of events is 70 (multiplied by 0.98). Nelson–Aalen cumulative hazard estimates are used to estimate the number of events at 3 years, as these estimates allow for censored data due to loss to follow-up.

b The modelling is based on 70 patients with events during the study period; 7 patients had an event between 3 years and their last study follow-up.

TABLE 56 Difference in sensitivity and specificity for model E compared with model A

Model (N = 277)	Sensitivity (95% CI)	Specificity (95% CI)	True positives (of 330 with recurrence) (95% CI)	True negatives (of 670 with no recurrence) (95% CI)
Model E (high vs. medium/low)	0.61 (0.49 to 0.73)	0.75 (0.69 to 0.82)	210 (162 to 241)	503 (462 to 549)
Model A (high vs. medium/low)	0.58 (0.46 to 0.70)	0.75 (0.68 to 0.81)	191 (152 to 231)	503 (456 to 543)
Difference	0.03 (-0.04 to 0.09)	0.00 (-0.02 to 0.04)	9 (-13 to 30)	0 (-13 to 27)

TABLE 57 Sensitivity and specificity analysis for model F compared with model A

Model	Risk group	Survival (95% CI survival)	Total	Recurrences	No recurrences	Sensitivity (95% CI) (n/N)	Specificity (95% CI) (n/N)
Model F, high vs. med/low	High risk (top 33%)	0.560 ^a (0.432 to 0.670)	72	32	40	0.68 ^b (0.53 to 0.81) (32/47)	0.76 ^b (0.68 to 0.82) (125/165)
	Med/low risk	0.911 ^a (0.848 to 0.949)	140	15	125		
	Total PCT		212	47	165		
	Missing		114	34	80		
	Total		326	81	245		
Model A, high vs. med/low	High risk (top 33%)	0.618 ^a (0.490 to 0.723)	72	28	44	0.60 ^b (0.44 to 0.74) (28/47)	0.73 ^b (0.66 to 0.80) (121/165)
	Med/low risk	0.882 ^a (0.814 to 0.926)	140	19	121		
	Total PCT		212	47	165		
	Missing		114	34	80		
	Total		326	81	245		

PCT, perfusion computed tomography.

a Adjusted as total number of events is 47. Nelson–Aalen cumulative hazard estimates are used to estimate the number of events at 3 years, as these estimates allow for censored data due to loss to follow-up.

b Estimates of sensitivity and specificity at 3 years is based on 47 patients with event at 3 years with molecular measurements. This is calculated from the Kaplan–Meier estimates of the number of events (see events in Kaplan–Meier graph number at risk) which were calculated as 44 with events.

TABLE 58 Mean difference and 95% limits of agreement for perfusion CT variables between local hospital and central review

Variable	Mean difference	95% limits of agreement
Blood flow (ml/minute/100 ml or 100 g)	-9.5	-71.2, + 52.3
Blood volume (ml/100 ml or 100 g)	-2.6	-13.4, + 8.19
Permeability surface area product (ml/minute/100 ml or 100 g)	3.1	-27.0, + 33.1
Mean transit time (seconds)	-2.8	-14.1, + 8.4

Appendix 2 Supplemental figures

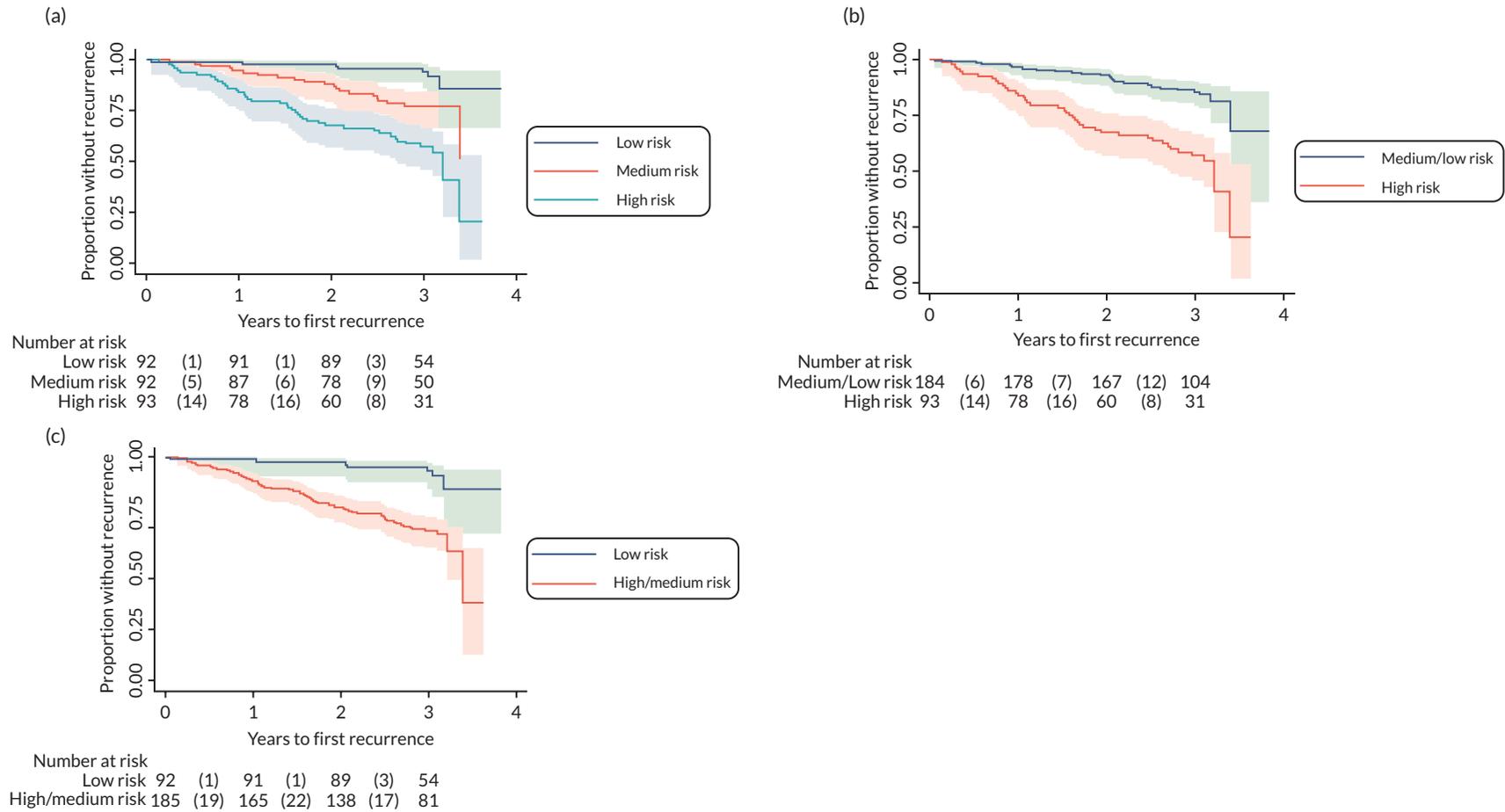


FIGURE 22 Kaplan–Meier plots for the different risk groupings: (a), all three risk groups; (b), two risk groups, low/medium vs. high; and (c), low vs. medium/high groups. Note: study end was 3 years; data beyond 3 years will be sparse and should not be overinterpreted.

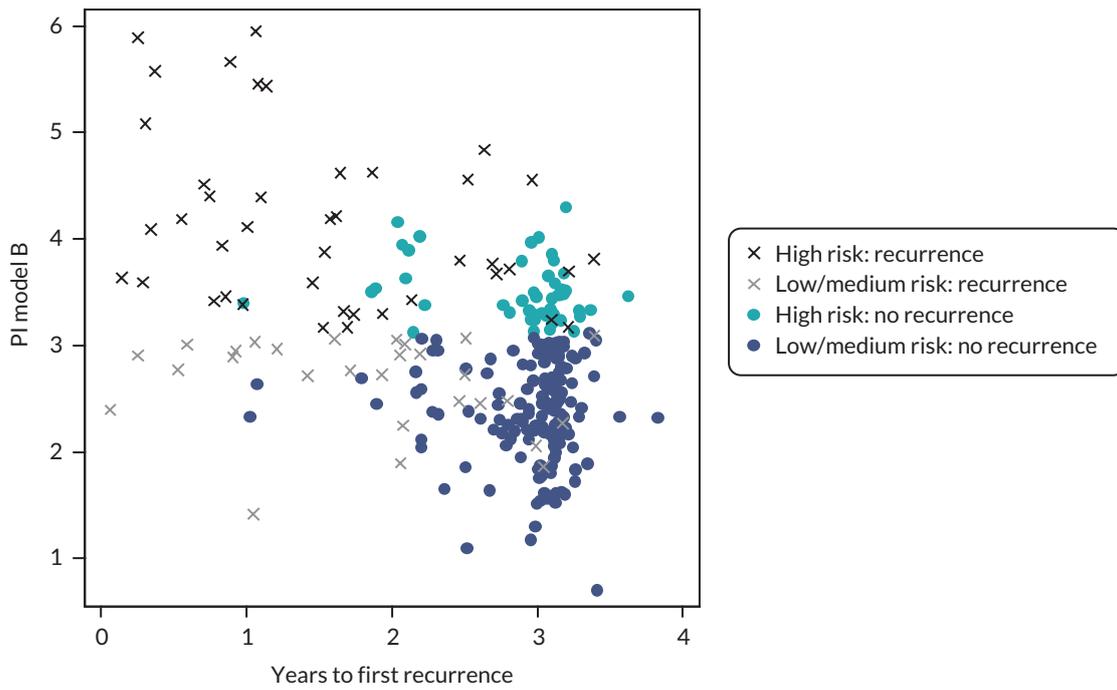


FIGURE 23 Scatter plot showing time to recurrence with respect to two risk groupings defined by model B prediction index (PI), low/medium risk vs. high risk.

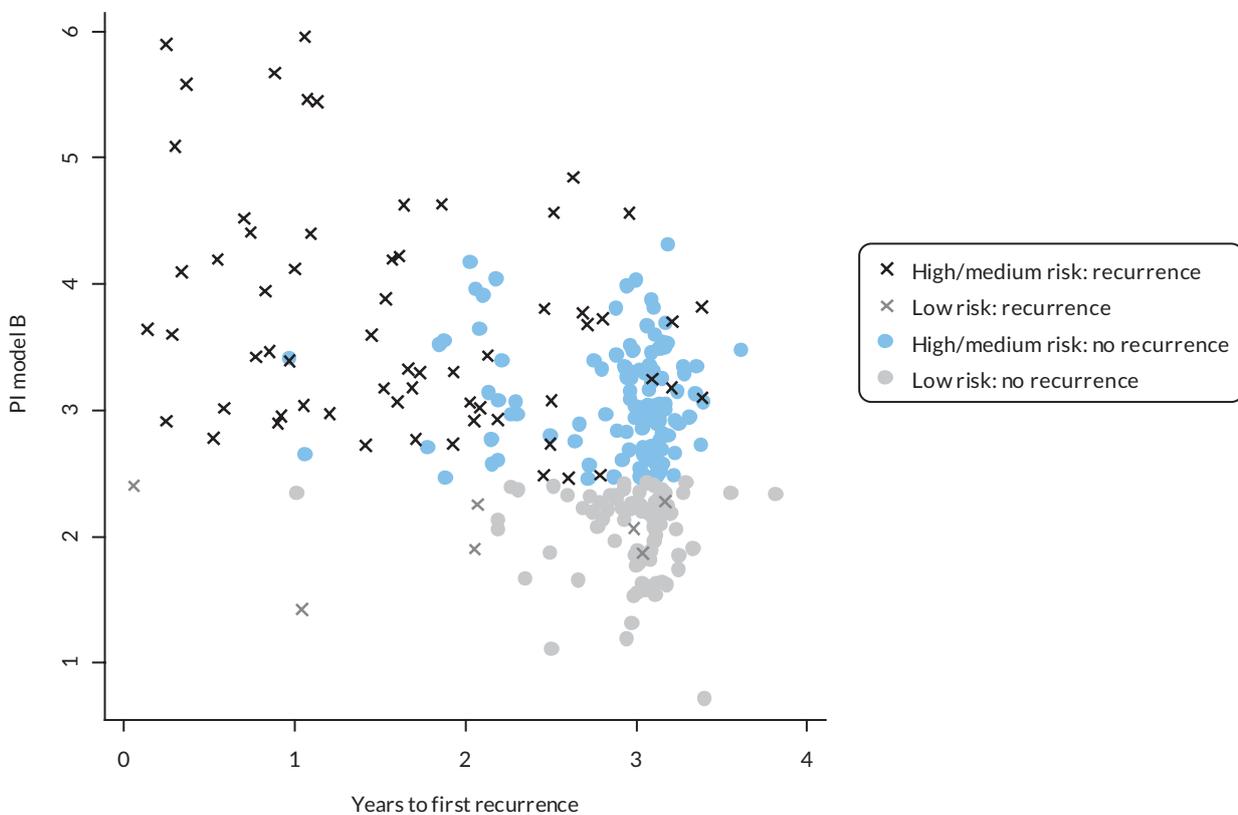


FIGURE 24 Scatter plot showing recurrences and time to recurrence with respect to two risk groupings defined by the model B prediction index (PI), high/medium vs. low risk.

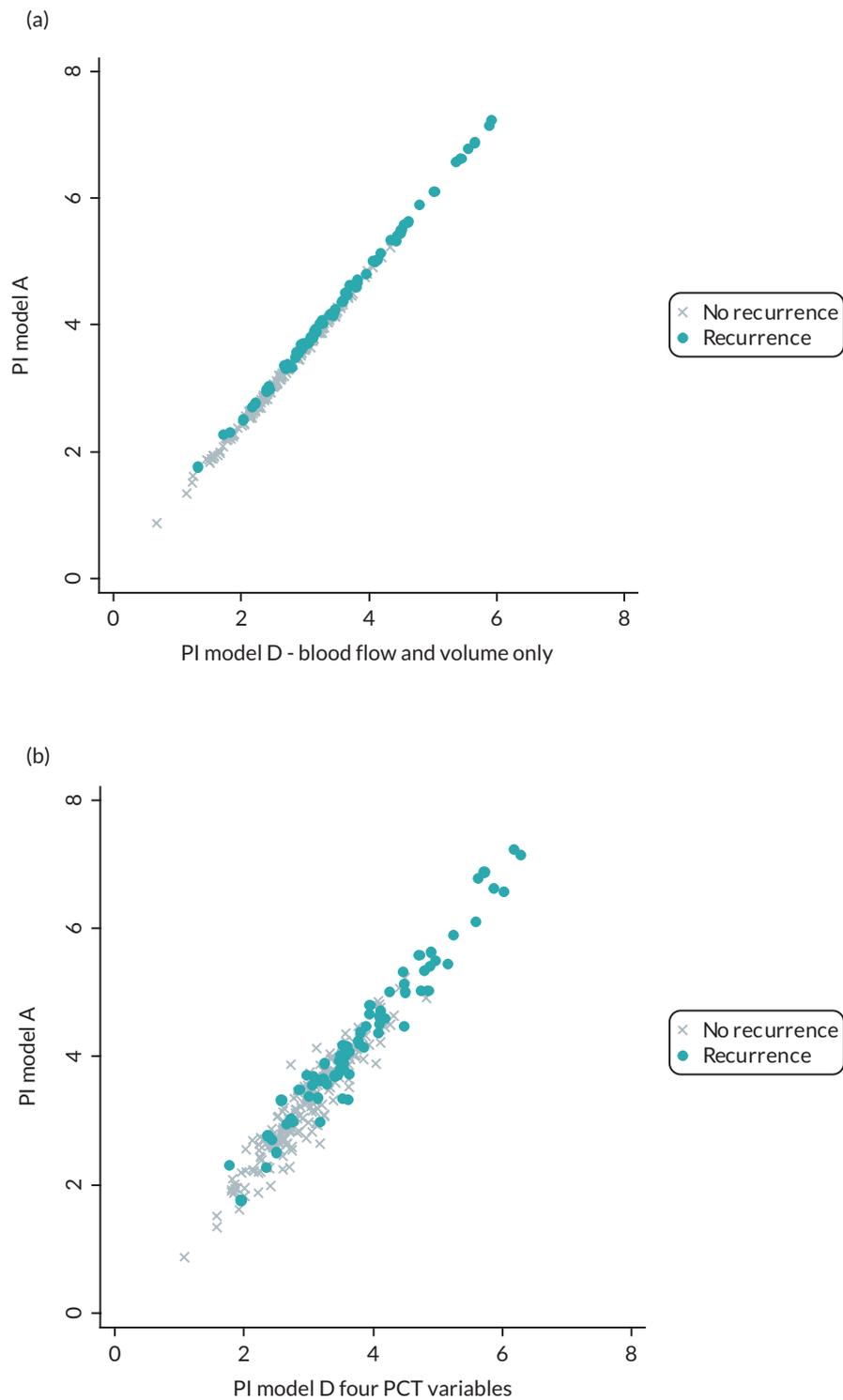


FIGURE 25 Scatter plot of prediction index (PI) for model A and model D. (a) model D with two perfusion variables and (b) model D with four perfusion variables. PCT, perfusion CT.

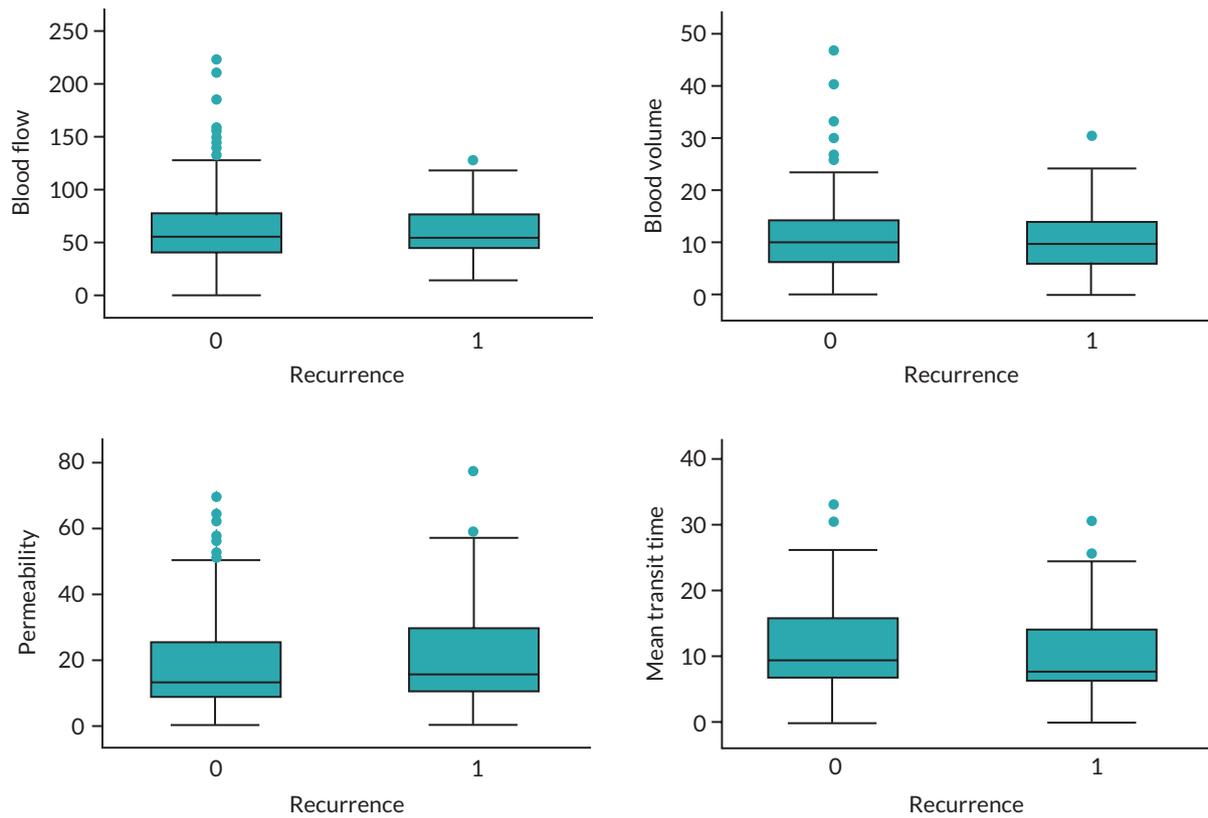


FIGURE 26 Box and whisker plots showing the distribution of perfusion CT variables by recurrence group. 0 = no recurrence.

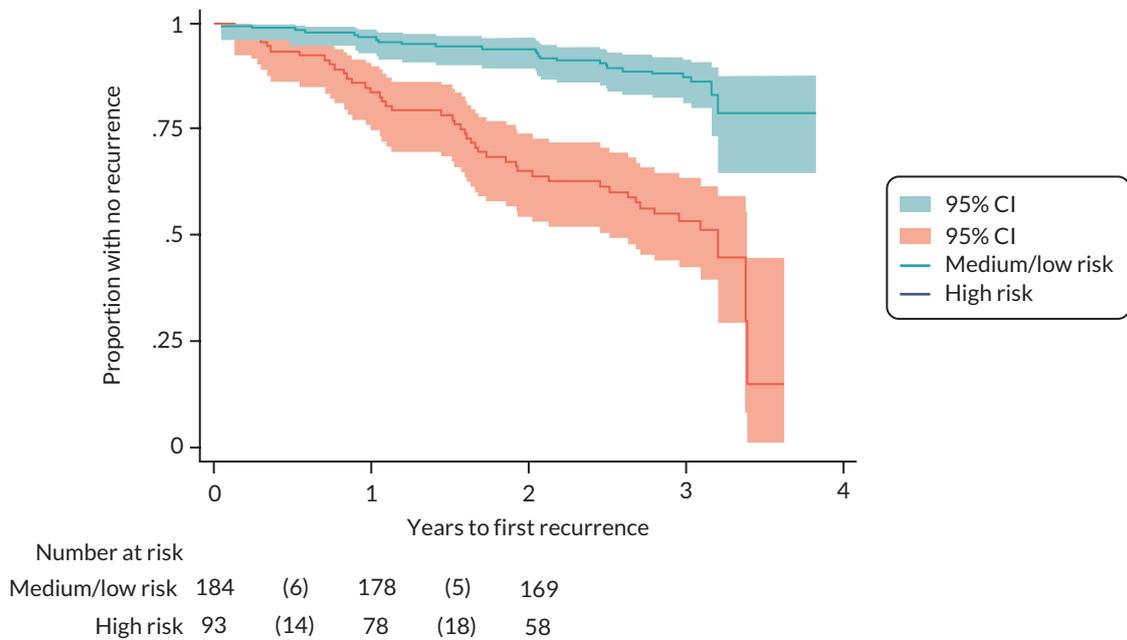


FIGURE 27 Kaplan-Meier plot for low/medium- vs. high-risk groups. Study end was 3 years; data beyond 3 years will be sparse and should not be overinterpreted.

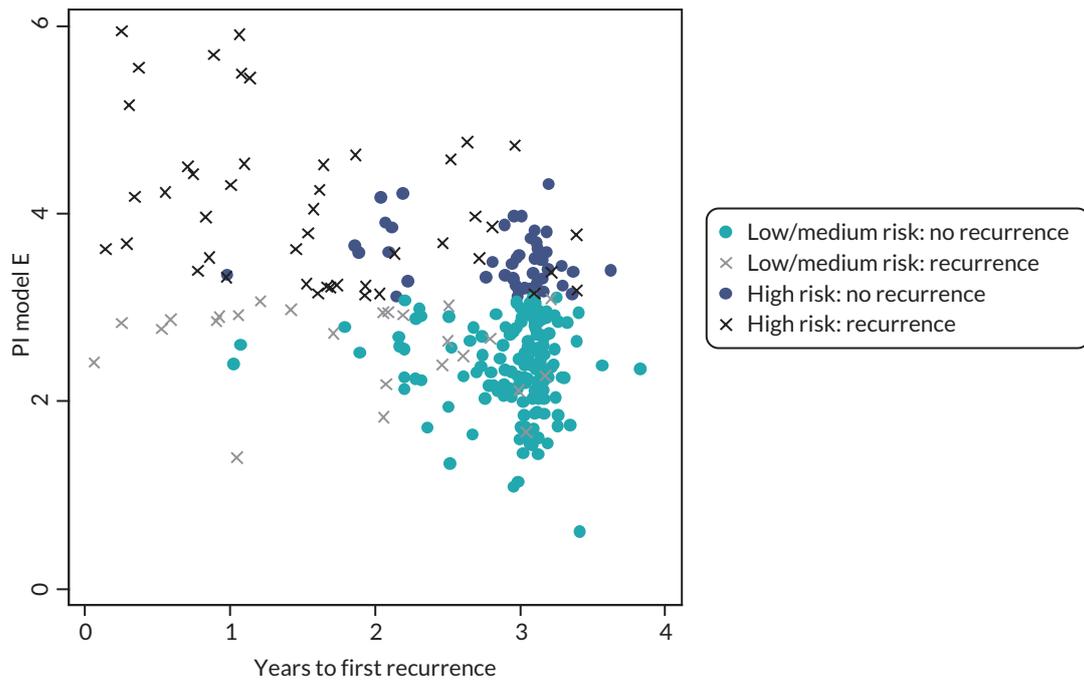


FIGURE 28 Scatter plot showing time to recurrence with respect to two risk groupings defined by model E prediction index (PI), low/medium risk vs. high risk.

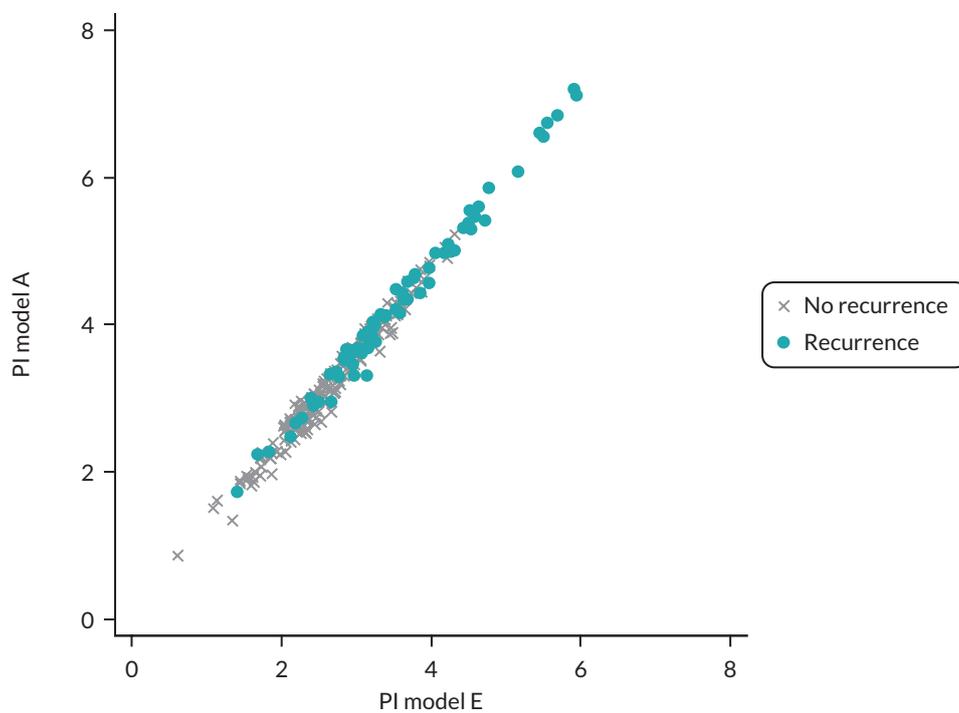


FIGURE 29 Scatter plot of prediction index (PI) for models A and E.

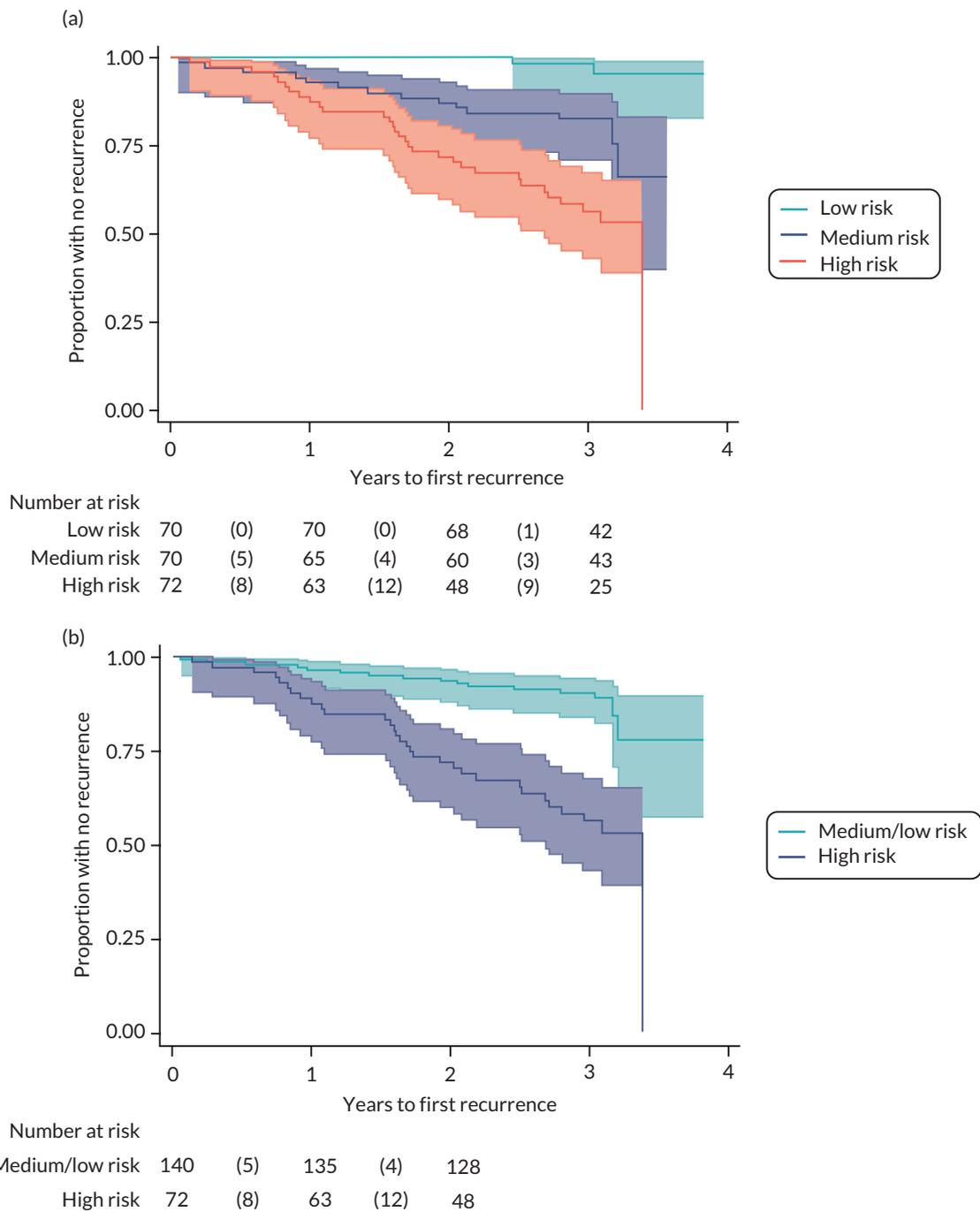


FIGURE 30 Kaplan–Meier plots (all variables, model F3) for the different risk groupings: (a) all three risk groups and (b) two risk groups, low-/medium- vs. high-risk groups.

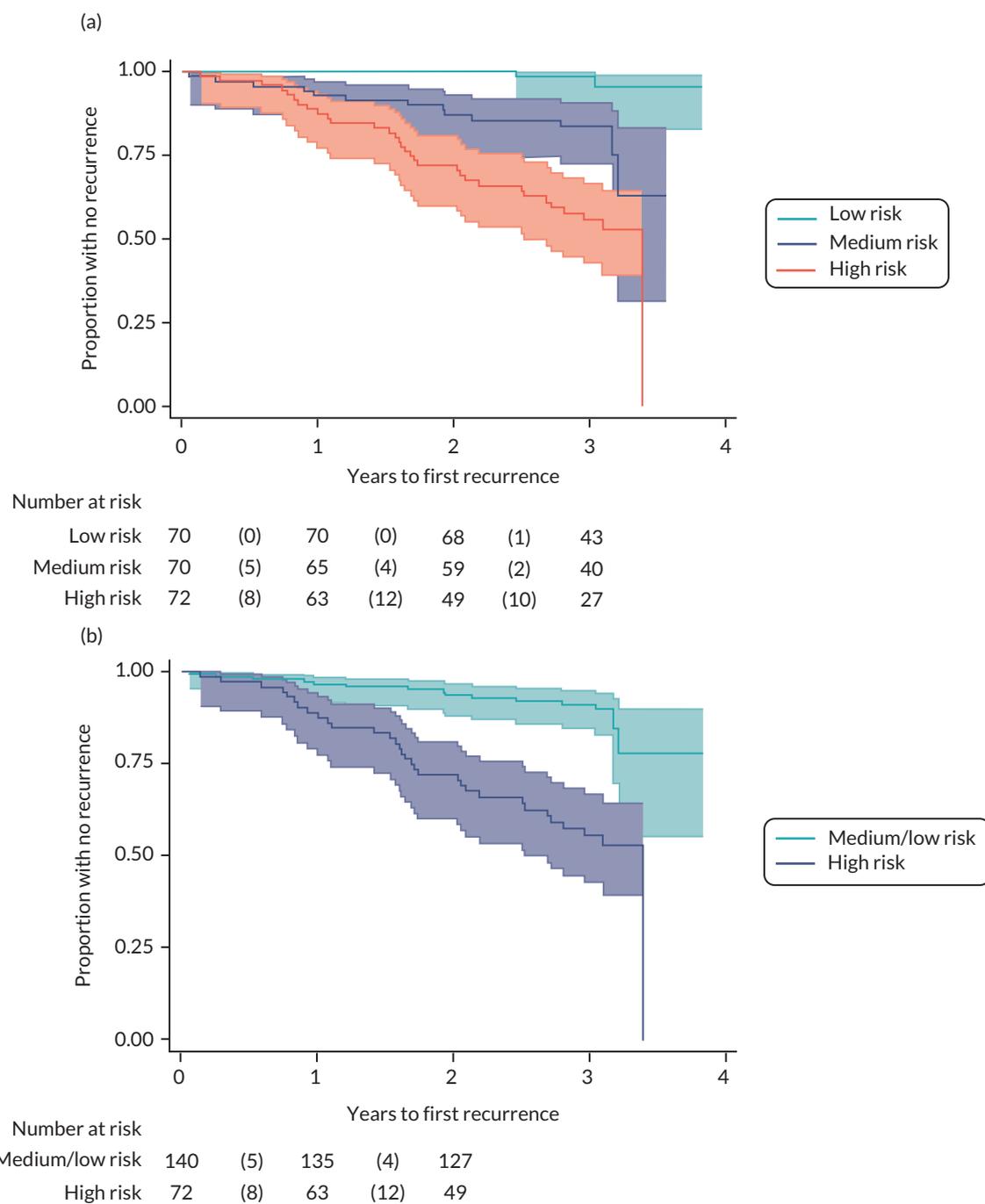


FIGURE 31 Kaplan-Meier plot (somatic mutation analysis variables, model F1) for the different risk groupings: (a) all three risk groups and (b) two risk groups, low-/medium- vs. high-risk groups. Study end was 3 years; data beyond 3 years will be sparse and should not be overinterpreted.

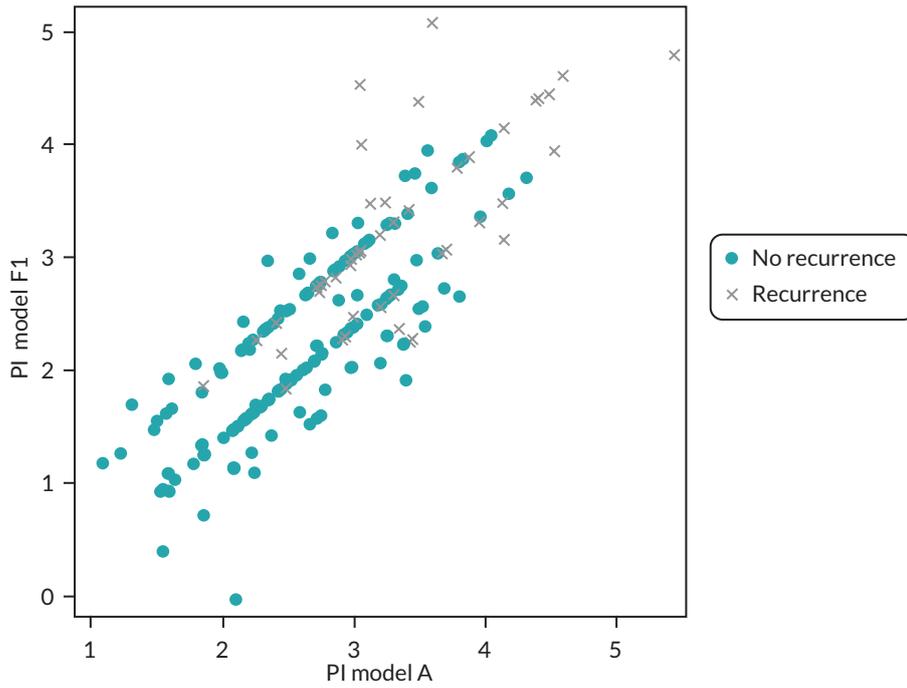


FIGURE 32 Scatter plot prediction index for model A and model F1 (somatic mutation analysis variables).

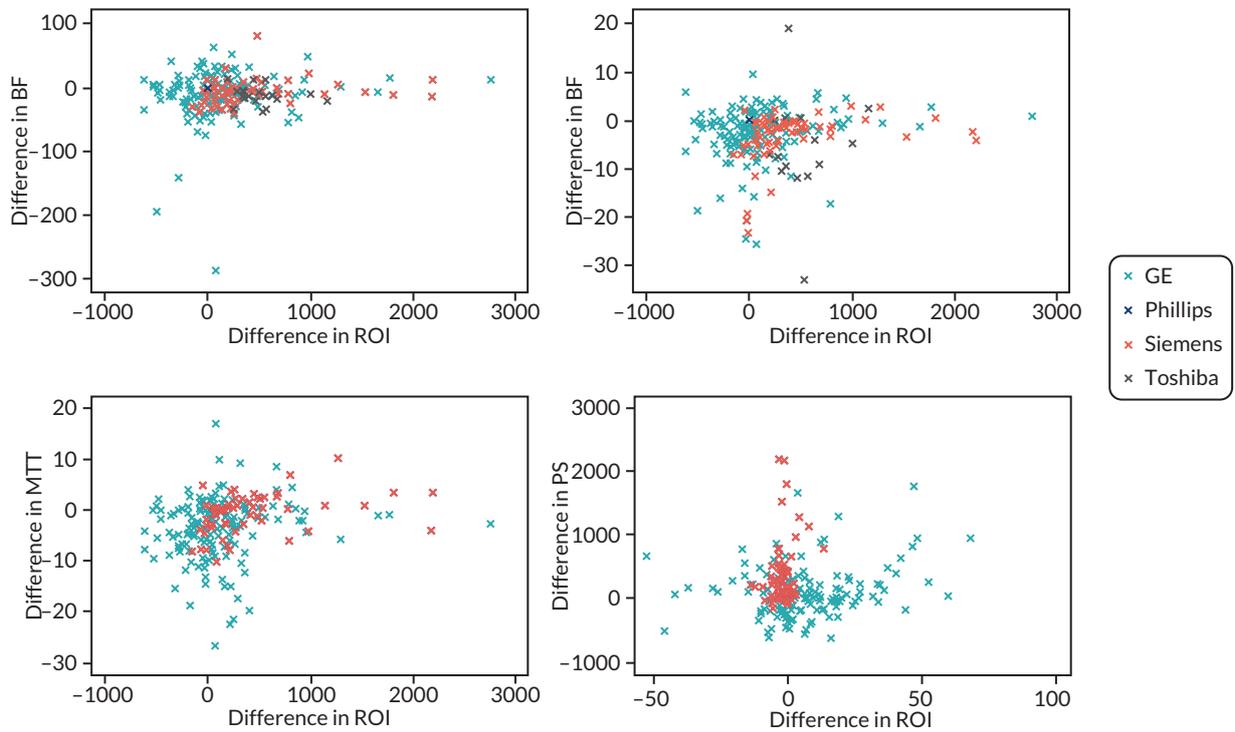


FIGURE 33 Difference in size of ROI placed plotted against the difference in each vascular parameter between local and central review with the scanner manufacturers highlighted by different colours. BF, blood flow; MTT, mean transit time; PS, permeability surface area product.

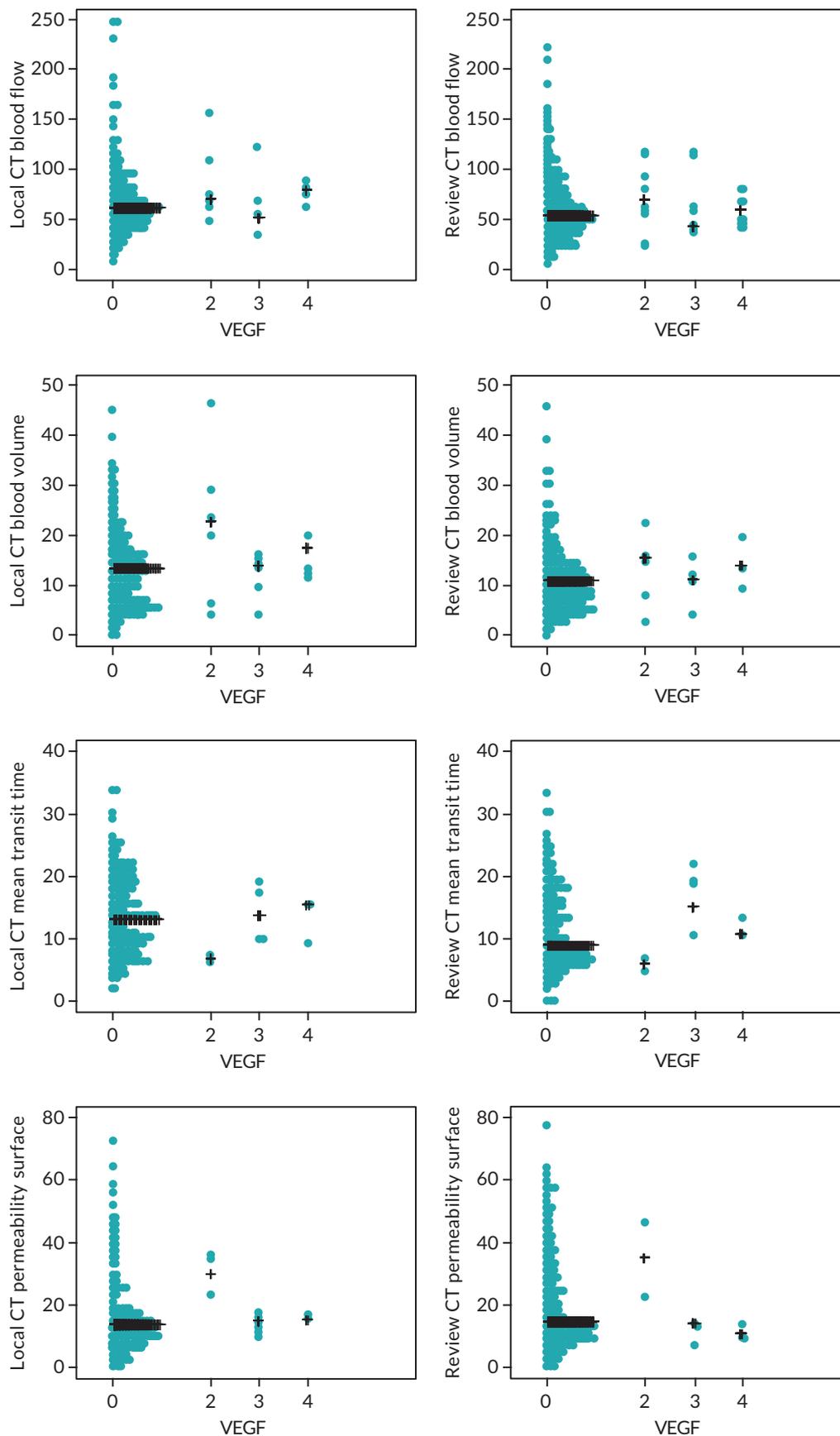


FIGURE 34 Scatter plots of CT blood flow, blood volume, mean transit time and permeability surface area product against VEGF score (0–4). Left-sided graphs are local review and right-sided graphs are central review.

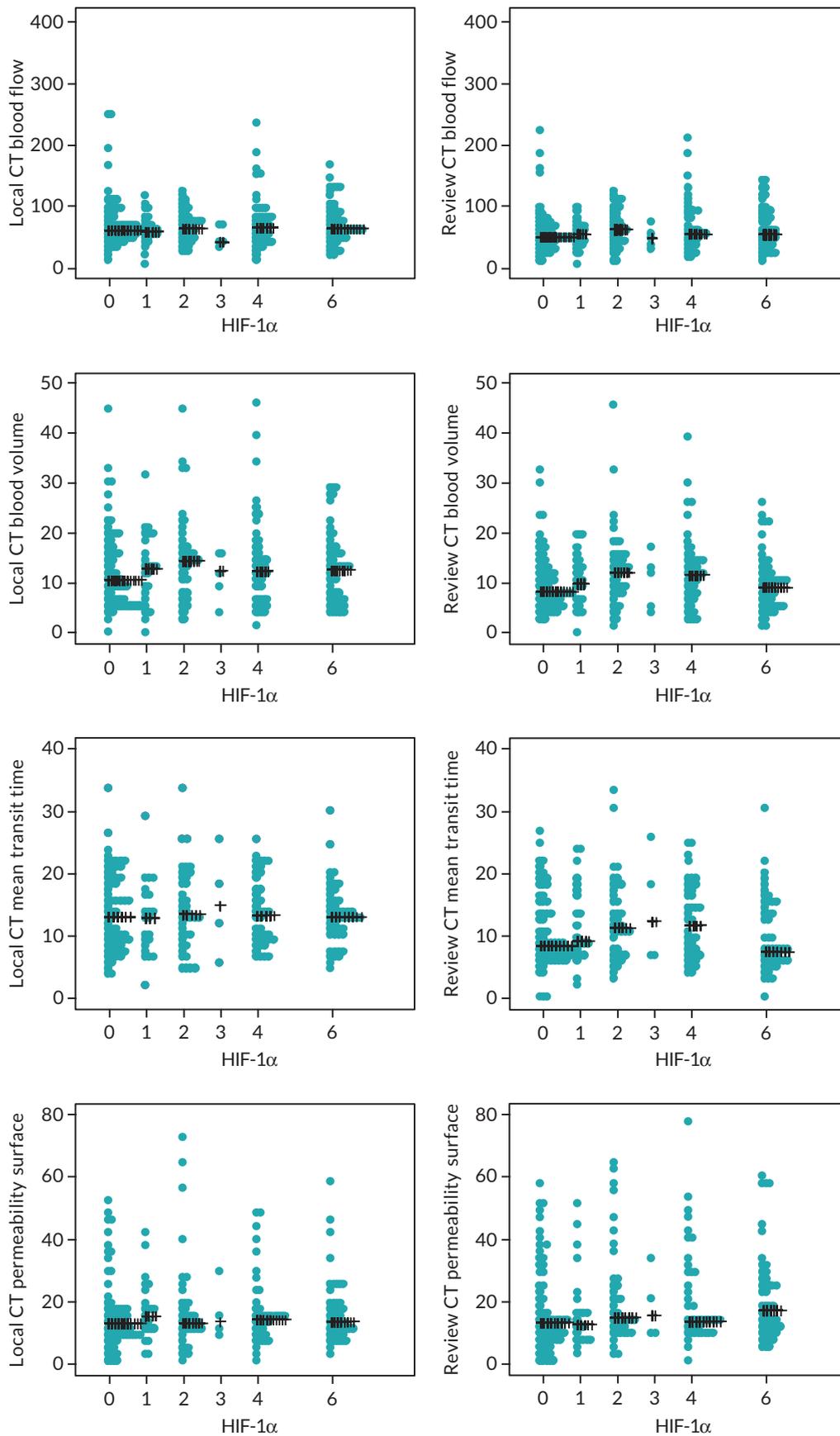


FIGURE 35 Scatter plots of CT blood flow, blood volume, mean transit time and permeability surface area product against HIF-1α score (0-6).

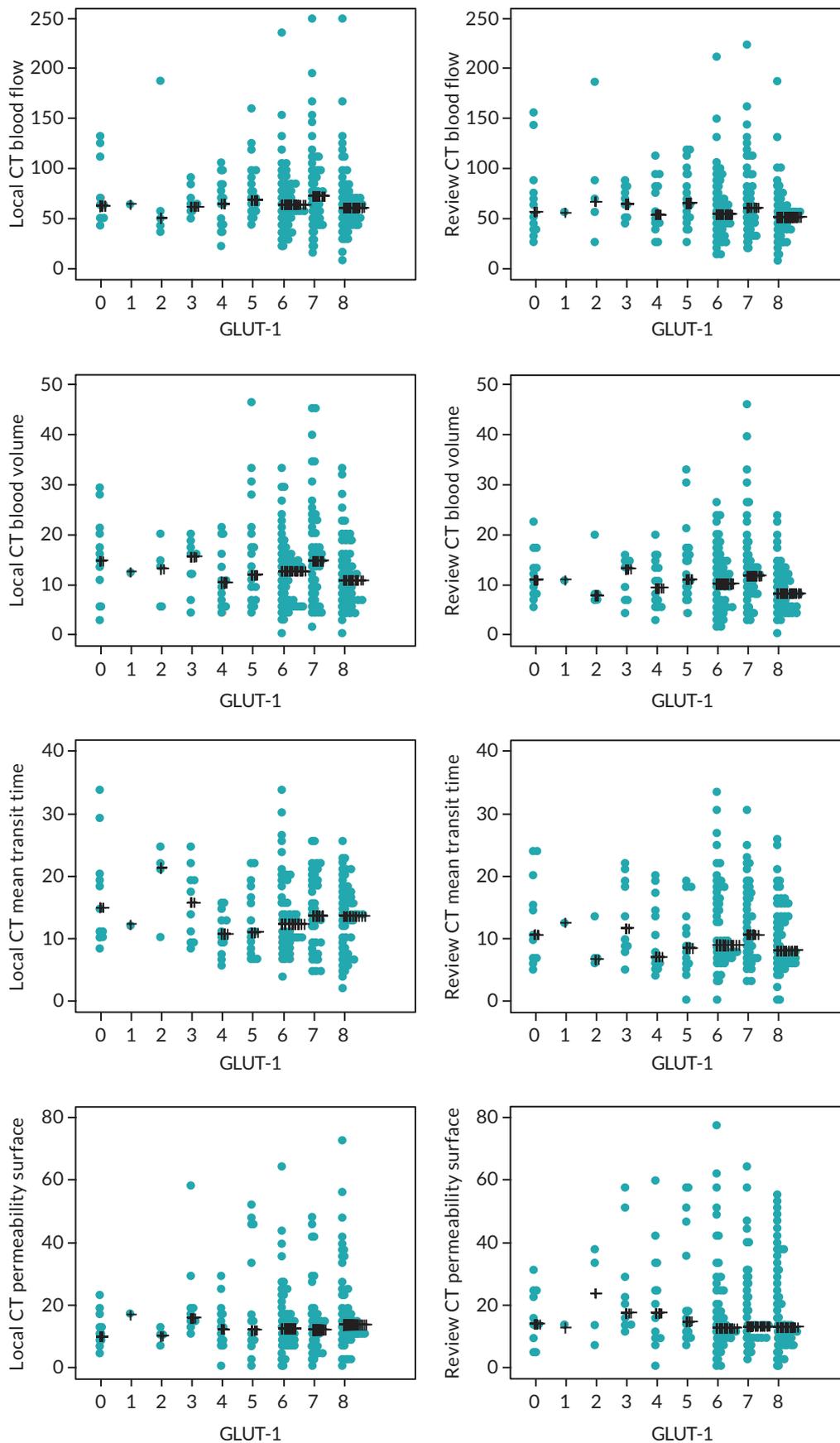


FIGURE 36 Scatter plots of CT blood flow, blood volume, mean transit time and permeability surface area product against GLUT-1 score (0-8). Left-sided graphs are local review and right-sided graphs are central review.

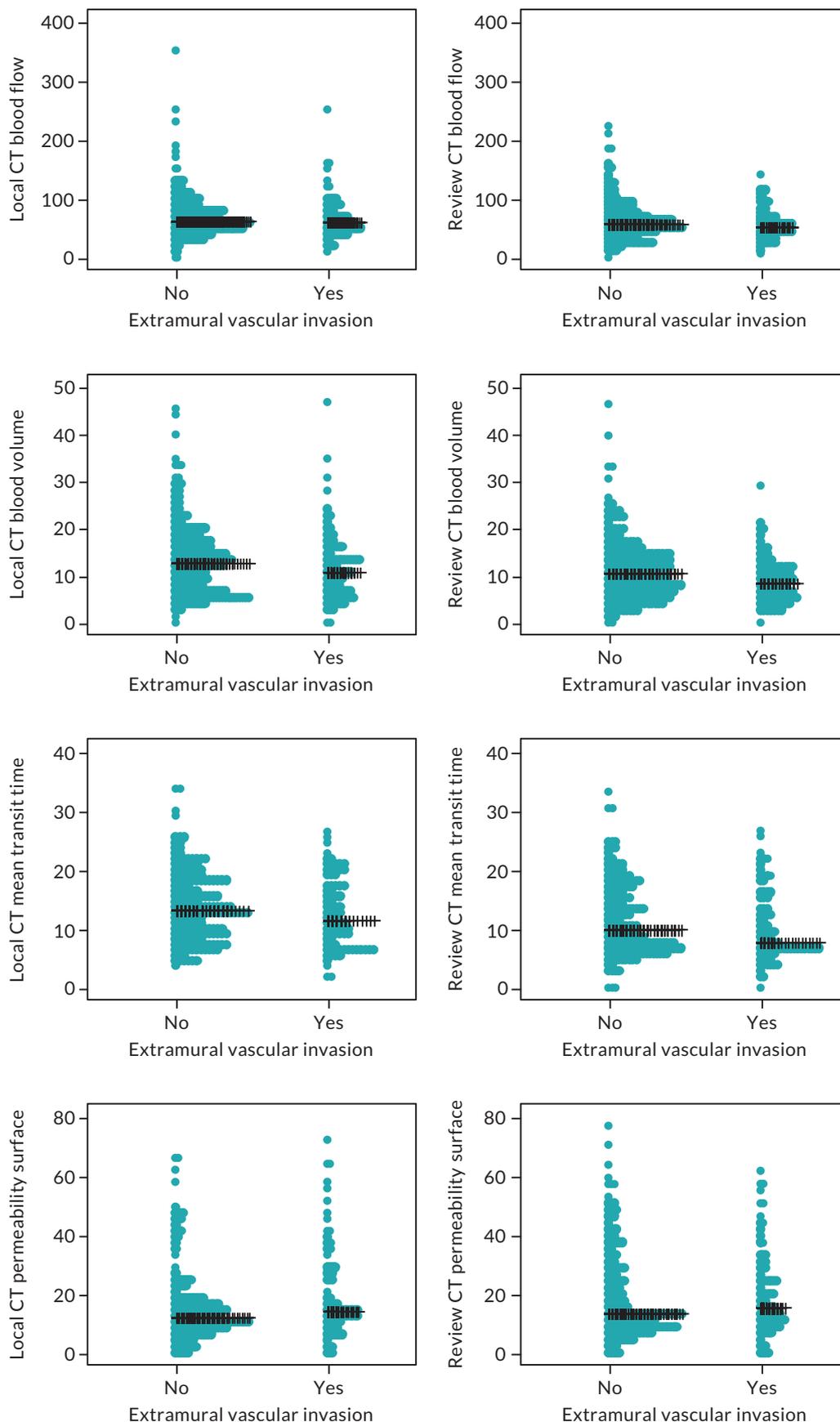


FIGURE 37 Scatter plots of CT blood flow, blood volume, mean transit time and permeability surface area product against the presence or absence of venous invasion. Left-sided graphs are local review and right-sided graphs are central review.

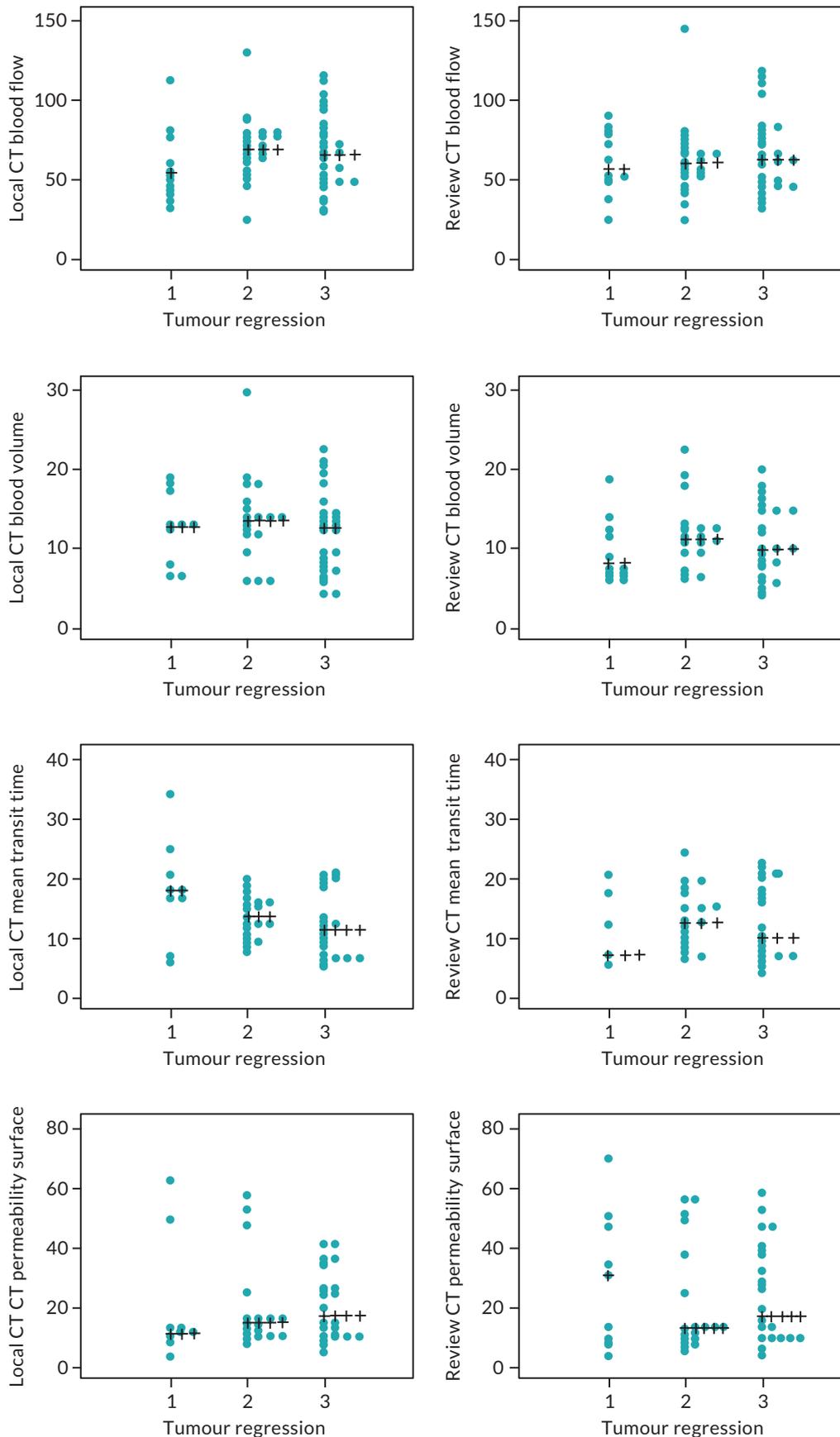


FIGURE 38 Scatter plots of CT blood flow, blood volume, mean transit time and permeability surface area product against the presence or absence of tumour regression. Left-sided graphs are local review and right-sided graphs are central review.

EME
HSDR
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 and Care 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