Factor influencing local regrowth after watch-and-wait for clinical complete response following chemoradiotherapy in rectal cancer: an individual participant data meta-analysis (InterCoRe consortium)

[1] Sami A Chadi MD MSc
[2,3] Lee Malcomson BSc
[4] Joie Ensor PhD
[4] Professor Richard D Riley PhD

Buenos Aires
[5] Carlos A Vaccaro MD PhD
[5] Gustavo L Rossi MD

Exeter, UK
[6] Ian R Daniels MD
[6] Neil J Smart PhD
[6] Melanie E Osborne MRCP FRCR

Maastricht, NL
[7,8] Professor Geerard L Beets PhD
[9] Monique Maas MD PhD

NYU, US
[10] Danielle S Bitterman MD
[11] Kevin Du MD PhD

OnCoRe
[12] Simon Gollins DPhil
[13] Professor Arthur Sun Myint FRCR
[14] Fraser M Smith MD
[15] Professor Mark P Saunders PhD
[16] Nigel Scott MD
[3] Professor Sarah T O'Dwyer MD

Rio de Janeiro, Brazil
[17] Rodrigo Otavio de Castro Araujo MD
[17] Marcus Valadao MD PhD
[17] Alberto Lopes MD

Taipei, Taiwan
[18] Cheng-Wen Hsiao MD
[18] Chien-Liang Lai MD

University of Pennsylvania, USA
[19] Radhika K Smith MD
[19] Emily Carter Poulson MD

Vejle, Denmark
[20,21] Ane Appelt PhD
[20] Professor Anders Jakobsen DMSc

[22] Professor Steven D Wexner MD PhD (Hon)
[23] Professor Angelita Habr-Gama MD
[24] Guilherme Sao Julião MD
[23,24] Rodiguo Perez MD PhD
[3,4] Professor Andrew G Renehan PhD

Affiliations

1 Division of Surgical Oncology and General Surgery, Princess Margaret Hospital and University Health Network, University of Toronto, Toronto, Ontario, Canada

2 Manchester Cancer Research Centre and NIHR Manchester Biomdeical Research Centre, Division of Cancer Sciences, School of Medical Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK

3 Colorectal and Peritoneal Oncology Centre, The Christie NHS Foundation Trust, Manchester, UK

4 Centre for Prognosis Research, Research Institute for Primary Care and Health Sciences, Keele University, Staffordshire, UK

5 Servicio Cirugia General, Sector de Coloproctologia, Hospital Italiano de Buenos Aires, Buenos Aires, Argentina

6 Exeter Colorectal Unit, and Exeter Surgical Health Sciences Research Unit (HESRU), Royal Devon & Exeter NHS Foundation Trust, Exeter, Devon, UK

7 Department of Surgery, Netherlands Cancer Institute, Amsterdam, The Netheralnds

8 GROW, School of Oncology and Developmental Biology, University of Maastricht, The Netherlands

9 Department of Radiology, Netherlands Cancer Institute, Amsterdam, the Netherlands

10 Harvard Radiation Oncology Program, Boston, MA, USA

11 Department of Radiation Oncology, New York University Langone Medical Center, New York, NY, USA

12 North Wales Cancer Treatment Centre, Rhyl, UK

13 Clatterbridge Cancer Centre, Liverpool, UK

14 Royal Liverpool Hospital NHS Foundation Trust, Liverpool, UK

15 Department of Clinical Oncology, The Christie NHS Foundation Trust, Manchester, UK

16 Royal Preston NHS Foundation Trust, Preston, UK

17 Department of Abdominal and Pelvic Surgery, Instituto Nacional de Câncer José Alencar Gomes da Silva - INCA, Rio de Janeiro, Brazil

18 Division of Colon and Rectal Surgery, Department of Surgery, Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan, Republic of China

19 Department of Surgery, Philadelphia VA Medical Center; Division of Colon and Rectal Surgery, Hospital of the University of Pennsylvania, Philadelphia, USA

20 Danish Colorectal Cancer Center South, Vejle Hospital, Vejle, Denmark

21 Leeds Cancer Centre, St. James's University Hospital; and Leeds Institute of Cancer and Pathology, University of Leeds, Leeds, UK

22 Department of Colorectal Surgery Cleveland Clinic Florida, Weston, Florida, USA.

23 Instituto Angelita & Joaquim Gama, São Paulo, Brazil

24 Ludwig Institute for Cancer Research, Molecular Biology and Genomics Lab, São Paulo Branch, Brazil

Running title: *Rectal cancer watch and wait IPD meta-analysis*

Keywords: rectal cancer; clinical complete response; watch and wait; individual participant data

Correspondence to:

Professor Andrew G Renehan
Division of Cancer Sciences, School of Medical Sciences,
Faculty of Biology, Medicine and Health,
University of Manchester,
The Christie NHS Foundation Trust,
Wilmslow Road
Manchester,
M20 4BX UK
Tel: +44 161 3060870
E-mail: andrew.renehan@manchester.ac.uk

Abstract: 393 words (max: 300); main text: 4758 words (max: 4000); x4 tables; x2 figures; 40 references (max: 40); supplemental material (19 pages); language: UK English.

ABSTRACT

Background: In patients with rectal cancer, 'watch-and-wait' (W&W) for clinical complete response (cCR) following neoadjuvant chemo-radiotherapy is a novel management strategy with potential to avoid major surgery. Study-level meta-analyses report wide variation in local regrowth rates. We performed an individual participant data (IPD) meta-analysis to evaluate factors influencing local regrowth occurrence as a potential explanation of this variation.

Methods: We updated a recent systematic review search (MEDLINE and Embase, from 01 Jan 2016 to 05 May 2017; plus expert knowledge) to identify published studies in patients with rectal cancer reporting local regrowth following W&W for cCR following neoadjuvant chemo-radiotherapy. We restricted studies to those that defined cCR using criteria equivalent to São Paulo benchmarks, and requested IPD. We assessed study quality using an 11-item checklist. The primary outcome was 2-year local regrowth cumulative incidence estimated using a two-stage random-effects (RE) IPD meta-analysis. We evaluated the impact of clinical and treatment factors using Cox frailty models, expressed as hazard ratios (HRs). From these models, we derived percentage differences in mean theta as an approximation of the impact of measured covariates on between-centre heterogeneity.

Results: We obtained IPD from 10 studies (11 datasets), totally 602 patients enrolled between 11 March 1990 and 13 February 2017, with a median follow-up of 37.6 (IQR: 25.0 – 58.7) months. Ten of the 11 studies were judged to be at low-risk of bias. There was wide between-centre variation in patient, tumour and treatment characteristics. The 2-year local regrowth cumulative incidence was 21.4% (RE 95% CIs: 15.3-27.6) with high levels of between-study heterogeneity (I^2 = 61%). There was some evidence that increasing cT stage was associated with increased risk of local regrowth (RE $HR_{per\ cT\ stage}$: 1.395, P_{trend} = 0.048). In a sub-cohort of patients managed post-2008 (after which high-resolution MR pre-treatment staging became standard), 2-year local regrowth cumulative incidences were 19% (95% CIs: 13-28) for cT1/cT2, 31% (95% CIs: 26-37) for cT3, and 37% (95% CIs: 30-60) for cT4 (RE $HR_{per\ cT\ stage}$: 1.482, P_{trend} = 0.033). We estimated that measured factors contributed 4.8% to 45.3% to the explanation of observed between-centre heterogeneity.

Interpretation: Among patients with rectal cancer and cCR managed by W&W, there was some evidence that increasing cT stage predicts for local regrowth. These data will inform clinician-patient decision-making in this setting. There is a research need to determine other predictors of a sustained clinical complete response.

Funder: None.

Registration: PROSPERO CRD42017070934

Research in context

Evidence before this study

In patients with rectal cancer who achieve a complete clinical response (cCR) after chemo-radiotherapy, the strategy of watch and wait (W&W) is new and offers an opportunity for patients to avoid major resection surgery. However, in the absence of randomised trials, this approach is not standard care. One recently published study-level meta-analysis of 23 studies (published and unpublished) including 871 patients, evaluated the outcome of patients managed by W&W and estimated a 2-year local regrowth rate of 15.7% but noted considerable between-study heterogeneity (I^2 = 55.9%), with rates ranging from 3.3% to 33.3%. A second updated study-level meta-analysis of 17 published-only studies (692 patients) reported a 3-year cumulative risk of local regrowth of 21.6% (I^2 = 66.5%). A register-based project, the International Watch and Wait Database (IWWD), reported on 880 patients with cCR managed by W&W, from 47 participating institutes (15 countries) and estimated a 2-year local regrowth cumulative incidence of 25.2%. Understanding factors that predict for local regrowth might explain the reported high levels of between-study heterogeneity. To-date, there is no large-scale study that has evaluated predictive factors for local regrowth because the study-level meta-analyses were unable to extract these data in an analysable form and there was considerable missingness in the IWWD registry-based report.

Added value of this study

This is the first reported individual participant data (IPD) meta-analysis in this field. By using the IPD methodology, there were two main advantages over study-level meta-analyses. First, we were able to test for predictive factors of local regrowth. And second, by incorporating Cox frailty models, we accounted for unmeasured factors at each study level. These factors might include centre-level protocols for staging, treatment, and follow-up. We obtained data from 10 studies (11 datasets) totally 602 patients, and with a median follow-up of 37.6 months, we estimated that the 2-year local regrowth cumulative incidence was 21.4%. There was some evidence that increasing cT stage was associated with increased risk of local regrowth, an association that remained after adjustments. We tested for other predictors including age, gender, cN stage, tumour distance to anal verge, serum CEA, radiotherapy dose, and time to W&W decision, and found no associations.

Implications of all the available evidence

The current literature notes wide variation in local regrowth rates after initial W&W and raised the concern that this strategy might not be generalisable to standard care. The

present analysis exploited this heterogeneity of outcomes and demonstrated that the latter is partly explained by differences in study baseline characteristics. For the first time at-scale, the present analysis shows that increasing cT stage is associated with increased risk of subsequent local regrowth. In a sub-cohort of patients managed after 2008 (reflecting current standard practice using high-resolution MR pre-treatment staging), 2-year local regrowth cumulative incidences were 19% for cT1/T2, 31% for cT3, and 37% for cT4. These estimates will inform clinician-patient decision making and future trials in the field of organ-preservation in patients with rectal cancer.

INTRODUCTION

Surgical resection is the mainstay of treatment for rectal cancer.[1] In patients who receive pre-operative neoadjuvant chemo-radiotherapy, up to a quarter have complete tumour regression, recognisable as a clinical complete response (cCR).[2] In these patients, 'watch-and-wait' (W&W) is a novel management strategy with potential to avoid major pelvic surgery.[3] This strategy originated from Habr-Gama and colleagues[4-6] in São Paulo, Brazil, over a decade ago, and extended, for example, to a large single institute series in the Netherlands[7, 8] and to a multi-centre network coordinated through Manchester in the North West of England and Wales (the OnCoRe project).[2] In a matched analysis of the OnCoRe data, survival rates were not inferior to those treated by standard surgical resection. Nonetheless, W&W has yet to reach universal acceptance in oncology and is not standard care.

In 2017, Dossa and colleagues[9] reported a study-level meta-analysis of 23 studies (15 published; 8 unpublished) including 871 patients, quantifying the risk of tumour local regrowth with W&W management in the setting of cCR. They estimated a 2-year local regrowth rate of 15.7% but noted considerable between-study heterogeneity (I^2 = 55.9%), with rates ranging from 3.3% to 33.3%.[9] A second updated study-level meta-analysis from Dattani et al.[10] identified 17 published-only studies (692 patients) and estimated a 3-year cumulative risk of local regrowth of 21.6%, again with high levels of heterogeneity (I^2 = 66.5%). Such between-study heterogeneity adds to concerns that W&W management, practiced as specialist centres, might not be generalisable to standard care. Alternatively, understanding factors that predict for local regrowth might explain the causes of between-study heterogeneity, ultimately better informing clinical pathways.

Here, we perform and report an individual participant data (IPD) meta-analysis, obtaining IPD from 10 published studies (11 datasets) within the International Complete Response (InterCoRe) consortium. The central aim was to evaluate for factors influencing local regrowth. The InterCoRe project parallels the International Watch and Wait Database (IWWD),[11] which recently reported on 880 patients with cCR managed by W&W, from 47

participating institutes (15 countries) and estimated a 2-year local regrowth cumulative incidence of 25·2%.

The IPD meta-analysis approach has several advantages over the study-level meta-analyses reported by Dossa et al.[9] and Dattani et al.[10], and over the registry-based IWWD reported by van der Valk and colleagues.[11] IPD afford the meta-analyst the opportunity to standardise inclusion criteria and analyses; obtain study results that had not been provided by the study publications; check modelling assumptions;[12] and importantly, for this study, model data as time-to-event cumulative incidence rather than crude rates. In the IPD meta-analysis framework, one models individual-level covariate-outcomes directly clustered within studies and minimises ecological bias compared with a meta-regression of aggregate data across studies.[13] To-date, there is no large-scale study that has evaluated predictive factors for local regrowth because the study-level meta-analyses[9, 10] were unable to extract these data in an analysable form and there was considerable missingness in the IWWD registry-based report.[11]

METHODS

Reporting was in accordance with PRIMA-IPD recommendations,[14] and the protocol was registered with PROSPERO (CRD42017070934).

Eligibility and study selection

The PICO (Population; Intervention; Comparator; Outcome) was as follows. We sought to identify studies of patients with locally advanced rectal cancer where the intervention was W&W after cCR following neoadjuvant chemo-radiotherapy, as the predominant treatment modality within each reported study, and followed-up to local regrowth, as defined by the 2014 Champalimaud conference.[15] We anticipated that the majority of studies would be treatment single-arm series, and accordingly, did not seek a comparator.

We used the systematic search published by Dossa and colleagues[9] (as our PICO was equivalent) and updated using MEDLINE and Embase databases. From the main

searches, we took a cut of identified studies from 01 Jan 2016 to 05 May 2017, and with studies identified through expert knowledge, added these to the studies identified by Dossa et al.[9] There was no language restriction. The search terms are detailed in webappendix p1.

As the central theme was the evaluation of predictive factors, we sought to have a baseline 'level playing field' and only included studies where the definition of cCR was judged to have used criteria equivalent to those of the São Paulo benchmarks, described by Habr-Gama et al. in 2004[5] and 2010[16] – namely, absence of residual ulceration, stenosis, or mass within the rectum using clinical and endoscopic examination. As abstracts did not allow this assessment, we excluded *a priori* unpublished studies. While, the Habr-Gama 'definition' papers[5, 16] restricted their cases to the distal rectum, subsequent large series,[8, 17] the two meta-analyses[9, 10] and the IWWD report[11] included proximal rectal tumours. Thus, we did not restrict by tumour distance from the anal verge.

Data collection and harmonisation

We approached chief investigators for identified studies and transferred fully anonymised data in encrypted files under centre-level governance arrangements. Data harmonisation is detailed in webappendix p2. To ensure homogeneity of patients entering into W&W management, from the received datasets, we excluded those who received short course radiotherapy as initial treatment; those treated by local excision or contact brachytherapy as part of the initial W&W management; and patients with distant metastases at baseline.

Risk of bias assessment in individual studies

We assessed study quality, modifying the Institute of Health Economics Quality Appraisal (IHEQA) Checklist for Case Series Studies.[18] This checklist comprises 18 'yes/no' items, with explanatory dictionaries. Only the first 11 items were relevant as subsequent items relate to reporting qualities, which did not apply to the IPD meta-analysis framework. Studies were considered to have a low-risk of bias if at least 80% of criteria were met, moderate-risk if 60% to 79% of criteria were met, and high-risk if less than 60% of criteria were met.

Outcome measures

The primary outcome was 2-year local regrowth cumulative incidence from date of W&W decision (we took this as equivalent to the date at which cCR was achieved). This allowed direct comparability with the aggregate-level meta-analysis from Dossa et al.[9] Secondary outcomes were: local regrowth cumulative incidence at 1-, 3-, 4- and 5-years; proportion of patients with local regrowth undergoing salvage surgery and proportion R0 (negative resection margin); 5-year overall survival (OS); 5-year non-regrowth disease-free survival (nrDFS), as detailed in our previous work;[2, 17] and 3-year distant metastasis rate, the latter three outcomes from date of first treatment. Post-protocol registration, we added 3-year post-salvage surgery survival, from date of salvage surgery.

Statistical Analysis

We used STATA version 14.0 (College Station, TX) in our analyses. For tables of study characteristics, we summarised proportions and medians (with inter-quartile ranges, IQRs) and compared with chi-squared and Kruskal- Wallis tests across studies.

To derive summary estimates of local regrowth cumulative incidences, we took two approaches. In our main model, we used a two-stage IPD approach; first undertaking time-to-event analyses per dataset to determine 2-year local regrowth cumulative incidence with 95% confidence intervals (95% CIs) using 1 – Kaplan-Meier (KM) analyses, and then combined the outputs using a random-effects methods with the `admetan` command. We assessed between-study heterogeneity with the I^2 statistic and assigned adjective low, moderate and high for values close to 25%, 50%, and 75%, respectively.[19] We repeated this for 1-, 3-, 4- and 5-year local regrowth cumulative incidences. For yearly summary estimates, we additionally derived prediction intervals. Second, we pooled data from all datasets and reported 1- through 5-year local regrowth cumulative incidence as 1 – KM and 95% CIs, without accounting for within centre correlations. We denoted our main (preferred) analysis as ‘RE’ (random-effects); and our second analysis as ‘pooled’ analysis.

We evaluated the impact of clinical and treatment covariates on local regrowth. Initially, we reported univariable pooled analysis, and compared as required using log-rank tests. For multivariable modelling, we used Cox frailty models, with results expressed as hazard ratios (HRs) and their 95% CIs. These models introduce a random-effects approach to account for associations and unobserved heterogeneity due to participation of different centres.[20] In the context of the present study, this approach takes account of unmeasured factors, sometimes called 'noise', at each study level such as centre-level protocols for staging, treatment, and follow-up. Frailty models are increasingly reported in multi-centre trial analyses to account for centre-level variations in clinical practice outside the trial protocol.[21] A limitation of the Cox frailty model occurs where one attempts to evaluate a predictor where certain values of that covariate exist only in specific centres. This is similar to the 'co-linearity' problem in regression models. From Cox frailty models, we derived theta (θ) values and their standard errors, and tested for $\theta = 0$ using the likelihood ratio test to quantify between-centre variability. P value < 0.01 was taken to mean that the correlation between participants within centres could not be ignored. To approximate the impact of measured factors on between-centre heterogeneity, we performed frailty models with and without covariates, and derived percentage mean differences in theta values. We tested assumptions of proportionality using Schoenfeld residuals and visualising predicted versus observed survival plots.

There were 20 core variables. Missingness was generally low. Data were complete for age and gender, and missing in 4.3% for cN stage; 7.6% for cT stage (none from one small study[22]); and 7.6% for tumour distance to anal verge (AV), which formed the basis for multivariable model A (10 datasets). Time to decision for W&W was not calculable for the two São Paulo datasets – thus, model B was model A plus time to decision for W&W based on 8 datasets. Serum CEA values were missing in 45.3% - thus, model C was model A plus serum CEA. Radiotherapy dose was missing in only 6.5% - but was near totally coincident with centre status (the co-linearity problem mentioned above), this was reported only in

univariable models. In multivariable models, the continuous variable, time to decision for W&W and serum CEA were modelling using fractional polynomials.[23]

For reporting proportions among patients undergoing salvage surgery, we used a two-stage IPD approach, first estimating proportions (using the `metaprop` command) with 95% CIs, and then combined using a random-effects methods. For the outcomes of OS, nrDFS and distant metastases, we used similar two-stage meta-analysis approaches as those for local-regrowth cumulative incidence.

For interpretation of statistical significance, we used the language recommended by Pocock and Ware,[24] namely: 'weak evidence' for $0.05 < p < 0.10$; 'some evidence' for $0.01 < p < 0.05$; and 'strong evidence' for $p < 0.001$.

Post-protocol stratified analysis

After full data collection, it became clear that enrolment dates ranged from 11 March 1990 to 13 February 2017; older than anticipated in the initial protocol. We posseted at there was risk of misclassification in pre-treatment staging across such a long period, and thus, we performed a post-protocol stratified analysis limited to patients enrolled into studies after 01 January 2008. We judged this to reflect contemporary clinical practice where pre-treatment staging is generally by high-resolution MR evaluation using the MERCURY study[25] principles.

Publication bias, data availability bias and reviewer selection bias

We assessed for *publication bias* using contour enhanced funnel plots and the asymmetry test in accordance with recommendations from Sterne et al.[26] As per principles set out by Ahmed et al.,[27] we assessed for *data availability bias* (IPD not available - e.g. unpublished but available as summary estimates in abstract form) by adding summary estimates from abstracts (from the Dossa et al.[9] meta-analysis) and comparing with our summary estimates for the IPD data. Similarly, we assessed for *reviewer selection bias* (IPD only sought from a subset of known studies) by adding summary estimates of other known published studies

(taken mainly from the Dossa et al.[9] meta-analysis as 2-year local regrowth was also primary outcome) and comparing with our summary estimates for the IPD data.

Role of the funding source

There was no funder of this study. Five members (SC, LM, JE, RR, AGR) of the writing sub-group had access to all the data. Senior members (SC, RR, GB, RP, AGR) of the writing sub-group shared the responsibility for the final decision to submit the report for publication.

RESULTS

Included studies

The flow diagram of the search, study identifications, and reasons for not including studies are detailed in webappendix p3-5. We initially received data from 11 studies, but excluded one study[28] where all patients received contact Papillon brachytherapy. For the large São Paulo series, we judged that there were two distinct cohorts – patients in the early series (denoted as São Paulo I), which were referred from two centres (University of São Paulo; Angelita & Joaquim Gama Institute, AJGI) and received neoadjuvant chemo-radiotherapy as 50.4 Gy and 2 cycles of 5-fluorouracil;[6] whereas the later series (denoted as São Paulo II) was treated from the outset through the AJGI, with an extended regimen of 54 Gy and 6 cycles of 5-fluorouracil.[6]

Our final analysis was from 10 studies (11 datasets).[2, 4, 6, 8, 22, 29-34] We judged that the definitions for cCR, across all datasets were equivalent to São Paulo benchmarks[5, 16] (evidenced in webappendix p6-7). The total number for analysis was 602 patients – 108 were not reported in previous publications. We noted two clinical indications among the studies: those termed standard practice neo-adjuvant chemoradiotherapy where cCR rates ranged from 12% to 49%, and two studies where there was an intentional enhanced cCR ranging from 68%[29] to 73%[4] (webappendix p8).

Study characteristics

Patient, tumour and treatment characteristics, by dataset, are summarised in **Table 1**. There was wide variation in characteristics and pathways: for example, median ages ranged from 59 to 75 years (p = 0.0001); proportion of men ranged from 40% to 91% (p = 0.001)); median tumour distance to AV from 3 to 6 cm (p = 0.0001); proportion of combined cT3/ cT4 stage from 43% to 83% (p = 0.007); and proportion of cN+ stage from 13% to 76% (p < 0.0001); and median time to W&W from 6 to 17 weeks (p = 0.0001). There were differences in radiotherapy treatment protocols – for example, for larger series, the radiotherapy dose regimen was predominantly 45 Gy in OnCoRe;[2] predominantly 50.4 Gy in Maastricht;[8] mainly 45 Gy and 50.4 Gy in São Paulo I;[16] mainly 54 Gy in São Paulo II;[6] and exclusively 60 Gy in Vejle.[29] Concurrent chemotherapy (5-fluorouracil-based in 518 out of 570 or 91%) was used in all series, and was used at least 95% of patients in seven datasets.

Assessment of Study Methodological Quality

Using the IHEQA Checklist,[18] ten of the 11 studies were judged to be at low-risk; one study[8] was judged to be moderate-risk of bias (webappendix p9).

Local regrowth

Overall, median follow-up was 37.6 (IQR: 25.0 to 58.7) months, but between studies, median follow-up ranged from 12.4 to 60 months. There were 166 local re-growths (crude proportion: 27.6%). The summary 2-year local regrowth cumulative incidence was 21.4% (RE 95% CIs: 15.3-27.6). There was a were high level of between-study heterogeneity (I^2 = 61%) **(Figure 1)**.

In the pooled analysis, the 1-, 2-, 3-, 4- and 5-year local regrowth rates were: 17.6% (95% CIs: 14.8-20.9), 24.7% (95% CIs: 21.4-28.5), 28.1% (95% CIs: 24.5-32.1), 31.1% (95% CIs: 27.2-35.5), and 31.6% (95% CIs: 27.6-36.0), respectively **(Figure 2A)**. By contrast, for 2-stage random-effects meta-analysis, summary point estimates for years 1 to 5 were more conservative at 15.6% (95% CIs: 9.9-21.4), 21.4% (95% CIs: 15.3-27.6), 24.9% (95% CIs: 18.5-31.3), 27.3% (95% CIs: 19.8-34.8), and 28.0 (95% CIs: 20.3-35.8), but with wider 95%

CIs **(Figure 2B)**. Local regrowth occurred almost exclusively in the first three years (155 out of 166 or 93.4%). We assessed visually for proportionality of local regrowth curves with time across the 11 datasets, and found similar patterns in all datasets (webappendix p10).

Cox frailty models

We tested for factors predicting local regrowth, initially for the total cohort, and then as a post-2008 sub-cohort analysis **(Table 2)**. For the total cohort, there was some evidence that increasing cT stage was associated with increased risk of local regrowth. By univariable analysis, 2-year cumulative incidences were 18% (95% CIs: 13-25) for cT1/T2, 29% (95% CIs: 24-34) for cT3, and 31% (95% CIs: 17-52) for cT4. In the multivariable frailty model, including age, gender, CT stage, N stage and distance to AV (model A), the HR per cT stage increase was 1.395 (RE 95% CI: 1.002, 1.941, P_{trend} = 0.048). There were no associations among other factors in model A (10 studies), model B (8 studies; incorporating time to W&W decision) or model C (8 studies; incorporating serum CEA).

For the sub-cohort of patients managed after 2008, 2-year local regrowth cumulative incidence increased in a stepwise manner from 19% (95% CIs: 13-28) for cT1/cT2, 31% (95% CIs: 26-37) for cT3, to 37% (95% CIs: 30-60) for cT4. In model A, the HR was per cT stage increase was 1.496 (RE 95% CI: 1.032, 2.168, P_{trend} = 0.033).

We tested (likelihood ratio test) for $\theta = 0$ and found statistical significance in all models, indicating that correlation within centres could not be ignored **(Table 3)**. We compared theta values in each model (A to C) with and without added factors, and noted that the likelihood ratio test remained statistically significant and that the addition of the measured factors only modestly influenced theta. We estimated that this contribution ranged from 4.8% to 45.3%.

Salvage surgery

Of the 166 patients with local regrowth, salvage surgery was performed in 137 (RE estimate: 89%, 95% CIs: 80-98), of which R0 status was achieved in 131 (RE: 98%, 95% CIs: 95-100)

(Table 4). After histopathological examination, only four patients were pT4; the majority (59 patients) were pT3 (RE: 44%, 95% CIs: 30-58). Node positivity was noted in 18 resections (RE: 16%, 95% CIs: 5-27).

The 137 patients with local regrowth undergoing salvage surgery were younger than the 29 patients treated by non-surgical strategies [median (IQR) age: 65.2 (57.4-71.2) versus 70.3 (60.9-76.0) years, p = 0.037). The commonest reason for no salvage surgery was synchronous distant metastases (12 patients) or unfit, mainly associated with older age (10 patients aged 75 years or older). The 3-year post-salvage survival rate was 80.1% (95% CIs: 70.3-87.0); the 3-year survival in patients not undergoing salvage surgery was 55.3% (95% CIs: 30.0-74.8) (webappendix p11). Accounting for age at local regrowth and between-centre variation, this was not statistical different (p = 0.153).

Survival and distant metastases rates

There were 68 deaths. The 5-year OS rate was 87.0 (RE 95% CIs: 81.5-92.4); and the 5-year nrDFS rate was 81.3% (RE 95% CIs: 74.9-87.6) (webappendix p12). Distant metastases were reported in 60 patients. The 3-year distant metastasis rate was 9.1% (RE 95% CIs: 8.7-9.5). The commonest sites of distant metastases were lung (31 of 60 patients) and liver (23 of 60 patients) (webappendix p13). Approximately half patients (31 of 60 patients) with distant metastases had local regrowth – these were identified synchronous with local regrowth in 12 patients; after local regrowth in 14 patients; and before local regrowth in only four.

Publication, data availability and reviewer selection biases

We visually inspected for asymmetry in the funnel plot for the 11 included datasets and found no evidence indicating publication bias (webappendix p14). For the primary outcome of 2-year local regrowth cumulative incidence, we found no evidence for data availability bias [RE: 21.4% (95% CIs: 15.1-27.7) versus 13.9% (95 CIs: 7.9-19.8), $p_{interaction}$ = 0.111]

(webappendix p15) and weak evidence for reviewer selection bias [RE: 21.4% (95% CIs: 15.1-27.7) versus 11.5% (95 CIs: 5.3-17.7), $p_{interaction}$ = 0.089] (webappendix p16).

DISCUSSION

Summary of main findings

We report five main findings. First, among studies of patients with rectal cancer and cCR managed by W&W, there was wide variation in baseline patient, tumour and treatment characteristics, but overall, the study quality was at low risk of bias. Second, the 2-year local regrowth cumulative incidence was approximately a fifth but there was wide variation across studies. Third, there was some evidence that increasing cT stage was associated with increased risk of local regrowth, particularly in sub-cohort of patients managed post-2008, but there was no clear signal of associations for other factors evaluated. Fourth, the observed between-study heterogeneity in local regrowth may partly be explained by study differences in measured factors, such as cT stage, but other unmeasured predictors might be relevant, and seeking these, should be a future research direction. Finally, we described several secondary outcomes, which will inform clinician-patient decision-making. These include that after tumour local regrowth, salvage rates were high, almost all achieved R0 status, and 3-year post-salvage survival was favourable; distant metastasis rates were low; and overall survival rates were favourable.

Context of other literature

There have been two published study-level meta-analyses[9, 10] and one large registry-based review[11] estimating local regrowth rates, and one meta-analysis[35] focusing on salvage in patients with local regrowth. Dossa et al.[9] performed a meta-analysis of 23 studies (published and unpublished) in 867 patients, and like our study, identified wide variation in baseline characteristics, but the authors were unable to directly test for differences. By contrast, our analysis directly reported these - for instance, median ages varied across the studies by as much as 16 years; and proportion of cT3/cT4 tumours varied from 43%[29] to

82%.[32] Our findings concur with Dossa and colleagues[9] that there was a wide variation on 2-year local regrowth rates across studies. They reported a summary 2-year local regrowth rate of 15.7%, lower than our summary estimate of 21.4%. Our assessment of data availability bias suggests that this difference was mainly driven by the inclusion of eight unpublished abstracts in the Dossa review,[9] but this difference was not statistically significant.

Dattani et al.[10] recently reported a study-level meta-analysis of 17 published-only studies in 692 patients. They reported a 3-year cumulative risk of local regrowth of 21.6%. This study did not have individual-level time to event data, but the investigators used a variety of methods to estimate numbers at risk at 3 years, thereby accounting for censoring. Thus, their estimate is broadly equivalent to our 2-year local regrowth cumulative incidence of 21.4%,

The recent IWWD report[11] was a registry-based pooled analysis of 880 participants from 47 centres (15 countries). There were data from five centres (AJGI; OnCoRe; Maastricht; Hospital Italiano, Buenos Ares; Vejle) from our IPD meta-analysis that contributed 552 participants to IWWD. Not unexpectedly, there were similar estimates for several outcomes, but not all. For IWWD[11] versus InterCoRe: 2-year local regrowth cumulative incidence was 25.2 (95% CIs: 22.2-28.5) versus 21.4% (RE 95% CIs: 15.3-27.6); 5-year OS was 84.7% (95% CIs: 80.9-87.7) versus 87.0% (RE 95% CIs: 81.5-92.4); and 3-year distant metastasis rate was 8.1% (95% CIs: 6.2-10.5) versus 9.1% (RE 95% CIs: 8.7-9.5). However, for patients with local regrowth, in the IWWD paper,[11] against a background of missing data, the salvage surgery rate was estimated to be 69%; that for InterCoRe was 89% (RE 95% CIs: 80-98). R0 status was attained in 88% for IWWD; and almost all salvage operations in InterCoRe (98% RE 95% CIs: 95-100). We added the new finding that 3-year post-salvage OS was 80.1%. We additionally reported the new finding that 5-year nrDFS was 81.3% (RE 95% CIs: 74.9-87.6), previously arguing that this is an informative outcome of disease control.[17]

Although, there were individual-level data in IWWD,[11] the data were pooled without taking account of between-study differences, and with high proportions of missing data for key confounder like cT stage (18%), the IWWD analysis was unable to evaluate for predictive factors of tumour local regrowth. From our analyses, we observed some evidence that increasing cT stage was associated with increased risk of local regrowth, and observation that had been noted at smaller scale from the São Paulo series.[36]

The systematic review of Kong et al.[35] focused on the rate of salvage surgery among studies where patients were managed by W&W. They included nine studies (370 patients) of which 256 (69.2%) had sustained cCR. In their analysis, the salvage surgery rate was 83.8%; the equivalent rate in our analysis was 89% (RE 95% CIs: 80-98).

Limitations and strengths

Our study has limitations. First, we did not collect data on surveillance protocols. The IWWD study[11] reported wide variation in frequency and assessment tools, and in theory, this might contribute to the observed between-study heterogeneity in key outcomes. We broadly controlled for this using frailty models, which takes account of centre-level heterogeneity, such as follow-up protocols. Second, the IPD meta-analysis approach does not resolve that included studies might be susceptible to bias. We formally assessed for this and found the great majority of studies were low risk. Third, we only sought data from a subset of published studies. We assessed for reviewer selection bias and found only weak evidence. Fourth, we only approached investigators of published studies, and thus data availability bias might occur. Again, we assessed for this, and found no strong evidence.

At first glance, a study weakness might be lack of a comparator group. There is debate what this comparator might be – from patients with rectal cancer undergoing resection surgery and found to have a pathological CR, to patients with a cCR and treated by surgery.[9] We previously argued that choice of comparator group depends on the question.[2] If the question is oncological safety, for example survival outcomes, the comparison group should be matched for key prognostic factors such as age, performance

status, and tumour stage to minimise selection bias. By contrast, the study aim here was to evaluate predictive factors for local regrowth, as these will inform clinical protocols.

Our analyses has several strengths. First, in contrast to study-level aggregate data meta-analyses,[9, 10] we assessed for predictors of local regrowth. To minimise the concern of baseline misclassification of cCR and facilitate interpretation of our predictions, we restricted studies to those that defined cCR using criteria equivalent to São Paulo benchmarks. Second, in common with IPD meta-analyses, in general, our platform allowed us to update and extend study-level information (for example, data on a sixth of participants were previously unreported); identify published studies which contained overlapping sets of participants; incorporated results from under-reported outcomes (for example, nrDFS[17]); verify results presented in the original study publications; standardised the strategy for statistical analysis; and assess model assumptions in each study. Specifically, we ran identical time-to-event analyses for each study, thus by-passing numbers at risk assumptions used in other meta-analyses. Third, we purposefully strengthened our analytical design seeking homogeneity of treatment – for example, some series[8, 16] historically included local excisions as part of the initial W&W management from an era when it was thought that this additional step was necessary. Similarly, we excluded patients with a 'near complete' clinical response,[37] some of whom were treated by Papillon brachytherapy.[38]

Clinical implications

The first clinical question is whether our findings have identified a patient sub-group unsuitable for W&W. The answer is no. For example, although in the post-2008 sub-analysis, cT4 tumours were associated with 2-year local regrowth cumulative incidence approximating 40%, there were still over half patients potentially benefiting from a sustained complete response. Going forward, there is a need to validate the associations between cT stage and local regrowth based on standardised MR-driven pre-treatment staging protocols.

The second clinical question is whether there should be a stratified approach to follow-up? Conceivably, one might argue that cT3 and cT4 tumours are at high-risk of local regrowth, but given the high salvage rates and attained R0 rates, it is questionable whether high-intensity surveillance in this patient sub-group would materially influence long-term outcomes. Similarly, the rate of distant metastases in all these patients is low, arguing that more regular CT surveillance is unlikely to make a major clinical impact.

Finally, what are the implications for future trials? There are now several ongoing and in-development trials where rectal organ preservation is the primary motivation. Our study included one such trial;[29] and the selection of patients in São Paulo II cohort[6] fulfil the same motivation. We showed that these sub-populations had similar local-regrowth rates as those achieving cCR through routine care.

Unanswered questions and future research

There are three key areas for research. First, there is a need to establish an internationally accepted definition of cCR, and in particular, establish the role of MR imaging in this definition. Second, there is a research need to determine other predictors of a sustained clinical complete response. There are several approaches including imaging, blood biomarkers, and tumour molecular phenotyping. Third, research is required to engage the options and preferences of patients. There is evidence that W&W is associated with substantially better quality of life and functional outcomes compared with the standard surgical resection.[39] But, there is a major caveat that chemo-radiotherapy itself might be associated with long-term morbidity. In studies to-date, no study included MR-tailored approaches by surgery alone as a comparator. All three pathways (chemo-radiotherapy plus resection versus chemo-radiotherapy plus W&W versus tailored resection alone) need to be evaluated. Only then, can we truly appraise the role of W&W in the overall standard care management of locally advanced rectal cancer.

Conflict of Interest

AGR reports personal fees from Merck Serona, personal fees from Janssen-Cilag, grants and personal fees from Sanofi Pasteur MSD, outside the submitted work. MPS reports personal fees from Merck, personal fees from Amgen, personal fees from Servier, personal fees from Eisai, personal fees from Roche, outside the submitted work. ID reports personal fees and other from Medtronic UK, personal fees and other from Gore UK, personal fees and other from Bard, personal fees, non-financial support and other from Molyncke, outside the submitted work. NJS reports personal fees from Medtronic, personal fees from WL Gore, outside the submitted work. The remaining authors declare no conflicts.

Contributions

SC performed literature searches, data extraction, and contributed to analyses, data interpretation and writing of the manuscript. SC, AGR, JE, RR, AH-G, SW contributed to the design of the study, data analysis and interpretation, and writing. LM assisted with the literature screening and data extraction and harmonisation. JE and RR contributed to statistical interpretation. SC, RP, AGR conceptualized the paper and contributed to all sections of the manuscript. All authors contributed to the final manuscript draft.

Acknowledgements

This research was supported by the NIHR Manchester Biomedical Research Centre.

Table 1 Characteristics of 11 datasets of 602 patient with rectal cancer and cCR initially managed by watch and wait in the InterCoRe consortium

	Totals	Buenos Aires, Arg[34]	Exeter, UK[32]	Maas-tricht, NL[8]	NYU, US[31]	OnCoRe, UK[2]	Rio de Janeiro, Brazil[30]	Sao Paulo I, Brazil[4]	Sao Paulo II, Brazil[6]	Taipei, Taiwan, China[22]	Universit y Penn, US[40]	Vejle, DK[29]	P values
Number of patients	602	23	11	84	8	162	42	131	66	18	17	40	
Study period		2005-14	2006-12	2005-14	2005-15	2005-17	2002-14	1990-2016	2001-16	2008-11	2001-14	2010-14	
Median age (range) years	64 (30-89)	75 (31-89)	64 (47-81)	63 (33-84)	63 (52-82)	67 (41-88)	64 (43-81)	62 (30-86)	59 (31-82)	64 (35-86)	63 (43-81)	68 (46-86)	0.0001*
Men (%)	401 (67)	11 (48)	10 (91)	55 (66)	6 (75)	114 (70)	17 (40)	85 (65)	42 (64)	15 (83)	14 (82)	32 (80)	0.001†
Median time to W&W (range) weeks	11 (8-15)	11 (8-16)	12 (11-16)	12 (8-20)	8 (6-19)	11 (10-14)	17 (10-26)	Not available	Not available	8 (7-9)	12 (6-19)	6 (6-6)	0.0001*
≥ 2 ECOG performance status (%)		Not available	Not available	Not available	0 (0)	9 (6)	0 (0)	Not available	Not available	Not available	Not available	Not available	
Median distance to AV (range) cm	5 (4-7)	5 (5-7)	4 (3-6)	5 (2-7)	5 (2-9)	5 (4-8)	3 (2-5)	5 (4-7)	6 (5-7)	6 (5-6)	5 (2-6)	6 (5-6)	0.0001*
Median serum CEA (range) ng/ml	2.5 (1.5-3.8)	2.9 (1.5-7.1)	Not available	2.1 (1.2-3.6)	3.0 (1.6-3.0)	2.9 (2.6-4.0)	2.4 (1.6-4.5)	2.0 (1.4-2.9)	2.2 (1.4-4.8)	1.6 (1.0-2.2)	5.6 (3.2-7.4)	Not available	Not applicable
cT stage													
cT1 & cT2 (%)	163 (29)	6 (30)	2 (18)	22 (26)	2 (25)	38 (23)	8 (29)	34 (28)	25 (38)	Not available	3 (18)	23 (58)	
cT3 & cT4 (%)	393 (71)	14 (70)	9 (82)	62 (74)	6 (75)	124 (77)	20 (71)	86 (72)	41 (62)	Not available	14 (83)	17 (43)	0.007‡
Missing		3	0	00	0	0	14	11	0	18	0	0	
cN stage													
cN0 (%)	228 (50)	9 (45)	4 (36)	20 (24)	3 (38)	51 (31)	26 (87)	89 (74)	39 (59)	13 (72)	11 (65)	23 (58)	
cN+(%)	228 (50)	11 (55)	7 (64)	64 (76)	5 (63)	111 (59)	4 (13)	31 (26)	27 (41)	5 (28)	6 (35)	17 (43)	< 0.0001‡
Missing		3	0	0	0	0	12	11	0	0	0	0	
Radiotherapy dose regimens													
45 cGy	212 (38)	5	3	1	1	153	5	29	0	14	1	0	
50.4 cGy	228 (41)	18	1	83	6	6	37	68	1	0	8	0	
54 cGy	79 (14)	0	0	0	1	2	0	7	64	4	1	0	

60 to 65 cGy	44 (8)	0	0	0	0	1	0	2	1	0	0	40	
Missing	39	0	7	0	0	0	0	25	0	0	7	0	
Concurrent chemotherapy (%)	570 (95)	23 (100)	8 (73)	84 (100)	7 (88)	143 (88)	40 (95)	126 (96)	66 (100)	18 (100)	15 (88)	40 (100)	NA
Chemotherapy regimens													
5FU/ LV	66 (12)	0	0	0	0	0	0	0	66	0	0	0	
Capecitabine	250 (44)	4	8	82	5	135	2	11	0	0	3	0	
Infusional 5-FU	202 (35)	19	0	0	2	5	38	115	0	18	5	0	
Oxaliplatin	9 (2)	0	0	2	0	0	0	0	0	0	7	0	
Tegafurur	40 (7)	0	0	0	0	0	0	0	0	0	0	40	
Others	3 (<1)	0	0	0	0	3	0	0	0	0	0	0	
Adjuvant chemotherapy (%)	51 (8)	0	0	35 (42)	0	13 (8)	1 (2)	0	0	0	2 (12)	0	NA
Median follow-up in months (IQR)	37.6 (25.0-58.7)	36.2 (36.2-36.2)	60 (38-81)	38.4 (24.7-57.6)	12.4 (10.4-52)	36.9 (22.8-53.1)	50.4 (32.7-63.8)	49 (18-86)	41 (25-58)	33.7 (25.4-52.6)	60 (35.4-91.8)	35.5 (25.6-42.2)	

Arg: Argentina. UK: United Kingdom. NL: the Netherlands. US: United States. NYU: New York University. Uni Penn: University of Pennsylvania. DK: Denmark. W&W: watch and wait.
AV: Anal verge. CEA: carcinoembryonic antigen. 5-FU: 5-fluoruracil. 5-FU/ LV: Concomitant chemotherapy (5-FU - 450 mg/m^2 and Leucovorin 50 mg fixed dose) delivered in a total of 6 cycles.
NA: not applicable. IQR: inter-quartile range
* Kruskal-Wallis test.
† Chi-squared test.
‡ Chi-squared test excluding missing data.

Table 2 Factors predicting local regrowth in patients initially managed by W&W in the InterCoRe consortium, accounting for centre effect in frailty models for the total cohort and post-2008 sub-cohort

	Total cohort (n: 602)				Post-2008 sub-cohort (n: 459)			
		IPD pooled analysis	IPD frailty models			IPD pooled analysis	IPD frailty models	
			Univariable	Multivariable*			Univariable	Multivariable*
	No. of patients	2-year local growth rate (95% CIs)	Hazard ratio (95% CIs)	Hazard ratio (95% CIs)	No. of patients	2-year local growth rate (95% CIs)	Hazard ratio (95% CIs)	Hazard ratio (95% CIs)
All patients	602	25 (21-28)			459	27 (23-31)		
Age group								
Per 10 years	602		1.007 (0.876, 1.157)	0.952 (0.820, 1.106)	459		0.924 (0.786, 1.088)	0.904 (0.762, 1.072)
Gender								
Women	201	23 (18-30)	1.000	1.000	155	22 (16-30)	1.000	1.000
Men	401	25 (21-30)	1.165 (0.835, 1.627)	1.193 (0.932, 1.056)	304	29 (24-31)	1.439 (0.972, 2.132)	1.534 (1.023, 2.298)
cT-stage								
cT1& cT2	163	18 (13-25)	1.000	1.000	125	19 (13-28)	1.000	1.000
cT3	367	29 (24-34)	1.400 (0.963, 2.029)	1.428 (0.954, 2.137)	282	31 (26-37)	1.553 (1.009, 2.392)	1.657 (1.065, 2.579)
cT4	26	31 (17-52)	1.527 (0.732, 3.185)	1.864 (0.840, 4.133)	22	37 (21-60)	1.710 (0.771, 3.794)	1.904 (0.849, 4.266)
per cT stage increase			1.348 (0.997, 1.822)	1.395 (1.002, 1.941)			1.454 (1.039, 2.035)	1.496 (1.032, 2.168)
cN-stage								
cN0	288	25 (21-31)	1.000	1.000	192	28 (22-35)	1.000	1.000
cN+	288	24 (19-30)	0.910 (0.652, 1.270)	0.869 (0.607, 1.242)	256	26 (21-32)	0.908 (0.629, 1.309)	0.751 (0.512, 1.100)
Distance to AV†								
< 6.0 cm	311	25 (20-30)	1.000	1.000	264	27 (22-33)	1.000	1.000
≥ 6.0 cm	246	23 (18-29)	0.937 (0.666, 1.317)	0.896 (0.630, 1.273)	160	23 (17-31)	0.810 (0.549, 1.196)	0.767 (0.511, 1.153)

Serum CEA categories†								
< 3.0 ng/ml	219	29 (23-35)	1.000	Not included‡	164	32 (25-40)	1.000	Not included‡
3.0 to 9.9 ng/ml	88	19 (12-29)	0.704 (0.422, 1.175)		71	20 (13-32)	0.704 (0.399, 1.243)	
≥ 10 ng/ml	22	36 (20-55)	1.544 (0.790, 3.017)		18	39 (30-65)	1.542 (0.754, 3.155)	
Radiotherapy dose group								
45 cGy	212	30 (24-37)	1.000	Not appropriate¶	187	33 (26-40)	1.000	Not appropriate¶
50.4 cGy	228	19 (14-25)	0.899 (0.563, 1.437)		161	19 (13-26)	0.568 (0.328, 0.985)	
54 cGy	79	30 (21-42)	1.537 (0.753, 3.140)		38	40 (26-60)	1.492 (0.740, 3.011)	
60 to 65 cGy	44	26 (15-41)	0.989 (0.409, 2.394)		43	26 (15-42)	0.812 (0.3622, 1.821)	
Intention to enhance cCR rate								
Yes (2 centres)	106	26 (19-36)	1.000	Not appropriate¶	67	28 (19-41)	1.000	Not appropriate¶
No (11 centres)	496	24 (21-29)	1.126 (0.573, 2.213)		392	26 (22-31)	1.105 (0.531, 2.296)	
Time to W&W¶¶								
< 13 wks	264	23 (18-29)	1.000	Not included‡	239	25 (20-33)	1.000	Not included‡
≥ 13 wks	141	25 (19-34)	1.211 (0.805, 1.824)		134	27 (20-36)	1.154 (0.770, 1.730)	

CEA: carcinoembryonic antigen. AV: distance to anal verge. cT and cN staging according to AJCC 7th edition.
Analyses in post-2008 sub-cohort limited to model of age, gender, cT-stage, cN stage and distance to AV (equivalent to model A in Table 3)
* For full cohort, the complete case multivariable model was based on 514 patienst, equivalent to model A in Table 3. For post-2008 cohort, the complete case multivariable model was based on 393 patients.
† Categorisation cut-off points for serum CEA and distance to AV were based on clinical reasons. Distance to AV of 6cm was taken as equivalent to that commonly used to define low-rectal cancers.
‡ Not included in multivariable model due to substantial proportion of missingness.
¶ Not appropriate due to coincidence of radiotherapy dose and study centre.
¶¶ Cut-off point of 13 weeks determined using spline approaches; equivalent to Model B in Table 3

Table 3 Outputs from frailty models clustering for centres and assessing changes in between-study heterogeneity (theta) for local regrowth, with and without covariates

	Covariates in model	No. of datasets	No. of patients	Mean theta, θ (se)	% difference in theta	Likelihood of theta = 0	AIC
TOTAL COHORT							
Model A							
No covariates	none	10	514	0.1190 (0.0954)		0.002	1673.7
With covariates	age, gender, cT stage, cN stage, distance to AV			0.1248 (0.1013)	4.8%	0.003	1680.2
Model B							
No covariates	none	8	337	0.1812 (0.1481)		0.001	981.5
With covariates	age, gender, cT stage, cN stage, distance to AV, time to W&W decision			0.2633 (0.2134)	45.3%	0.001	978.3
Model C							
No covariates	none	8	278	0.2662 (0.2054)		< 0.001	872.2
With covariates	age, gender, cT stage, cN stage, distance to AV, baseline serum CEA			0.2465 (0.1921)	7.4%	0.001	870.9
POST-2008 SUB-COHORT							
Model A							
No covariates	none	10	393	0.0964 (0.0776)		0.005	1234.4
With covariates	age, gender, cT stage, cN stage, distance to AV			0.1084 (0.0851)	12.4%	0.003	1233.9

W&W: watch and wait. Se: standard error. CEA: carcinoembryonic antigen. AV: anal verge. Distance to AV, time to W&W decision and serum CEA as continuous variables. Time to W&W decision as a spline pivoted as 13 weeks (determined from fractional polynomials)

Table 4 Treatment of 166 patients with local regrowth initially managed by W&W in the InterCoRe consortium

		Post-salvage surgery pathology findings			
	N (%)	Positive CRM	Positive DRM	ypT stage† T0/T1/T2/T3/T4/missing	ypN stage† N0/N+/ missing
No. of patients with local regrowth	166				
Non-surgical treatments	29* (17)				
Surgical treatments	137 (83)				
Operation types					
Abdomino-perineal resection	73 (52)	4	0	1/7/22/35/2/6	56/9/8
Anterior resection	29 (21)	0	0	3/5/6/14/0/1	20/8/1
Hartmann's procedure	4 (3)	0	1	0/0/0/3/0/1	2/1/1
Other radical operations	6 (4)	0	0	0/0/2/2/2/0	6/0/0
Transanal local excision or TEM	25 (18)	Not applicable	1	0/5/13/5/0/0	Not applicable
Totals		**4**	**2**	**4/17/43/59/4/8**	**84/18/10**
Total colostomies	80 (48)				

Values in parentheses are percentages and only cited if value greater than five.
TEM: transrectal endoscopic micro-dissection. CRM: circumferential resection margin. DRM: distal resection margin.
*Five patients had synchronous diagnoses of distant metastases.
† The Taiwan study did not contribute to the pathological T and N staging.

FIGURE LEGENDS

Figure 1 Forest plot of 11 datasets. Sorted by descending 2-year local regrowth cumulative incidences Summary estimate, 95% confidence intervals, and prediction intervals shown for random effects method, and restricted maximum likelihood estimators (reml).
UK: United Kingdom. DK: Denmark. NYU NYC: New York University, New York City. Arg: Argentina. US: United States. NL: The Netherlands

.

Figure 2 A; pooled analysis with local regrowth cumulative incidence from 1 to 5 years, with 95% CIs. B; 2-stage random-effect meta-analysis with summary estimates for 1- through 5-years, with 95% CIs, and predictive intervals in green.

References

1. Brenner H, Kloor M, Pox CP. Colorectal cancer. *Lancet* 2014; **383**(9927): 1490-502.

2. Renehan AG, Malcomson L, Emsley R, et al. Watch-and-wait approach versus surgical resection after chemoradiotherapy for patients with rectal cancer (the OnCoRe project): a propensity-score matched cohort analysis. *Lancet Oncol* 2016; **17**(2): 174-83.

3. Breugom AJ, van de Velde CJ. Is it time for watchful waiting for rectal cancer? *Lancet Oncol* 2015; **16**(8): 875-6.

4. Habr-Gama A, Gama-Rodrigues J, Sao Juliao GP, et al. Local recurrence after complete clinical response and watch and wait in rectal cancer after neoadjuvant chemoradiation: impact of salvage therapy on local disease control. *Int J Radiat Oncol Biol Phys* 2014; **88**(4): 822-8.

5. Habr-Gama A, Perez RO, Nadalin W, et al. Operative versus nonoperative treatment for stage 0 distal rectal cancer following chemoradiation therapy: long-term results. *Annals of surgery* 2004; **240**(4): 711-7; discussion 7-8.

6. Habr-Gama A, Sabbaga J, Gama-Rodrigues J, et al. Watch and wait approach following extended neoadjuvant chemoradiation for distal rectal cancer: are we getting closer to anal cancer management? *Dis Colon Rectum* 2013; **56**(10): 1109-17.

7. Maas M, Beets-Tan RG, Lambregts DM, et al. Wait-and-see policy for clinical complete responders after chemoradiation for rectal cancer. *J Clin Oncol* 2011; **29**(35): 4633-40.

8. Martens MH, Maas M, Heijnen LA, et al. Long-term Outcome of an Organ Preservation Program After Neoadjuvant Treatment for Rectal Cancer. *J Natl Cancer Inst* 2016; **108**(12).

9. Dossa F, Chesney TR, Acuna SA, Baxter NN. A watch-and-wait approach for locally advanced rectal cancer after a clinical complete response following neoadjuvant chemoradiation: a systematic review and meta-analysis. *Lancet Gastroenterol Hepatol* 2017; **2**(7): 501-13.

10. Dattani M, Heald RJ, Goussous G, et al. Oncological and Survival Outcomes in Watch and Wait Patients With a Clinical Complete Response After Neoadjuvant Chemoradiotherapy for Rectal Cancer: A Systematic Review and Pooled Analysis. *Annals of surgery* 2018.

11. van der Valk MJM, Hilling DE, Bastiaannet E, et al. Long-term outcomes of clinical complete responders after neoadjuvant treatment for rectal cancer in the International Watch & Wait Database (IWWD): an international multicentre registry study. *Lancet* 2018; **391**(10139): 2537-45.

12. Burke DL, Ensor J, Riley RD. Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. *Stat Med* 2017; **36**(5): 855-75.

13. Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ* 2010; **340**: c221.

14. Stewart LA, Clarke M, Rovers M, et al. Preferred Reporting Items for Systematic Review and Meta-Analyses of individual participant data: the PRISMA-IPD Statement. *Jama* 2015; **313**(16): 1657-65.

15. Heald RJ, Beets G, Carvalho C. Report from a consensus meeting: response to chemoradiotherapy in rectal cancer - predictor of cure and a crucial new choice for the patient: on behalf of the Champalimaud 2014 Faculty for 'Rectal cancer: when NOT to operate'. *Colorectal Dis* 2014; **16**(5): 334-7.

16. Habr-Gama A, Perez RO, Wynn G, Marks J, Kessler H, Gama-Rodrigues J. Complete clinical response after neoadjuvant chemoradiation therapy for distal rectal cancer: characterization of clinical and endoscopic findings for standardization. *Dis Colon Rectum* 2010; **53**(12): 1692-8.

17. Renehan AG, Malcomson L, Emsley R. Watch-and-wait approach for rectal cancer: concepts of a subject-specific method. *Lancet Gastroenterol Hepatol* 2017; **2**(9): 627.

18. Moga C, Guo B, Schopflocher D, Harstall C. Development of a quality appraisal tool for case series studies using a modified Delphi technique. Edmonton: Institute of Health Economics, 2012.

19. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002; **21**(11): 1539-58.

20. Hougaard P. Frailty models for survival data. *Lifetime Data Anal* 1995; **1**(3): 255-73.

21. Wille-Jorgensen P, Syk I, Smedh K, et al. Effect of More vs Less Frequent Follow-up Testing on Overall and Colorectal Cancer-Specific Mortality in Patients With Stage II or III Colorectal Cancer: The COLOFOL Randomized Clinical Trial. *Jama* 2018; **319**(20): 2095-103.

22. Lai CL, Lai MJ, Wu CC, Jao SW, Hsiao CW. Rectal cancer with complete clinical response after neoadjuvant chemoradiotherapy, surgery, or "watch and wait". *Int J Colorectal Dis* 2016; **31**(2): 413-9.

23. Royston P, Sauerbrei W. A new approach to modelling interactions between treatment and continuous covariates in clinical trials by using fractional polynomials. *Stat Med* 2004; **23**(16): 2509-25.

24. Pocock SJ, Ware JH. Translating statistical findings into plain English. *Lancet* 2009; **373**(9679): 1926-8.

25. MERCURY_study_group. Diagnostic accuracy of preoperative magnetic resonance imaging in predicting curative resection of rectal cancer: prospective observational study. *BMJ* 2006; **333**(7572): 779.

26. Sterne JA, Sutton AJ, Ioannidis JP, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 2011; **343**: d4002.

27. Ahmed I, Sutton AJ, Riley RD. Assessment of publication bias, selection bias, and unavailable data in meta-analyses using individual participant data: a database survey. *BMJ* 2012; **344**: d7762.

28. Smith FM, Al-Amin A, Wright A, Berry J, Nicoll JJ, Sun Myint A. Contact radiotherapy boost in association with 'watch and wait' for rectal cancer: initial experience and outcomes from a shared programme between a district general hospital network and a regional oncology centre. *Colorectal Dis* 2016; **18**(9): 861-70.

29. Appelt AL, Ploen J, Harling H, et al. High-dose chemoradiotherapy and watchful waiting for distal rectal cancer: a prospective observational study. *Lancet Oncol* 2015; **16**(8): 919-27.

30. Araujo RO, Valadao M, Borges D, et al. Nonoperative management of rectal cancer after chemoradiation opposed to resection after complete clinical response. A comparative study. *Eur J Surg Oncol* 2015; **41**(11): 1456-63.

31. Bitterman DS, Resende Salgado L, Moore HG, et al. Predictors of Complete Response and Disease Recurrence Following Chemoradiation for Rectal Cancer. *Front Oncol* 2015; **5**: 286.

32. Dalton RS, Velineni R, Osborne ME, et al. A single-centre experience of chemoradiotherapy for rectal cancer: is there potential for nonoperative management? *Colorectal Dis* 2012; **14**(5): 567-71.

33. Smith JJ, Chow OS, Gollub MJ, et al. Organ Preservation in Rectal Adenocarcinoma: a phase II randomized controlled trial evaluating 3-year disease-free survival in patients with locally advanced rectal cancer treated with chemoradiation plus induction or consolidation chemotherapy, and total mesorectal excision or nonoperative management. *BMC Cancer* 2015; **15**: 767.

34. Vaccaro CA, Yazyi FJ, Ojra Quintana G, et al. Locally advanced rectal cancer: Preliminary results of rectal preservation after neoadjuvant chemoradiotherapy. *Cir Esp* 2016; **94**(5): 274-9.

35. Kong JC, Guerra GR, Warrier SK, Ramsay RG, Heriot AG. Outcome and Salvage Surgery Following "Watch and Wait" for Rectal Cancer after Neoadjuvant Therapy: A Systematic Review. *Dis Colon Rectum* 2017; **60**(3): 335-45.

36. Habr-Gama A, Sao Juliao GP, Gama-Rodrigues J, et al. Baseline T Classification Predicts Early Tumor Regrowth After Nonoperative Management in Distal Rectal Cancer After Extended Neoadjuvant Chemoradiation and Initial Complete Clinical Response. *Dis Colon Rectum* 2017; **60**(6): 586-94.

37. Hupkens BJP, Maas M, Martens MH, et al. Organ Preservation in Rectal Cancer After Chemoradiation: Should We Extend the Observation Period in Patients with a Clinical Near-Complete Response? *Annals of surgical oncology* 2018; **25**(1): 197-203.

38. Sun Myint A, Smith FM, Gollins S, et al. Dose Escalation Using Contact X-ray Brachytherapy After External Beam Radiotherapy as Nonsurgical Treatment Option for Rectal Cancer: Outcomes From a Single-Center Experience. *Int J Radiat Oncol Biol Phys* 2018; **100**(3): 565-73.

39. Hupkens BJP, Martens MH, Stoot JH, et al. Quality of Life in Rectal Cancer Patients After Chemoradiation: Watch-and-Wait Policy Versus Standard Resection - A Matched-Controlled Study. *Dis Colon Rectum* 2017; **60**(10): 1032-40.

40. Smith RK, Fry RD, Mahmoud NN, Paulson EC. Surveillance after neoadjuvant therapy in advanced rectal cancer with complete clinical response can have comparable outcomes to total mesorectal excision. *Int J Colorectal Dis* 2015; **30**(6): 769-74.