

Meta-analysis of colorectal cancer follow-up after potentially curative resection

S. Mokhles¹, F. Macbeth², V. Farewell³, F. Fiorentino⁴, N. R. Williams⁵, R. N. Younes⁶, J. J. M. Takkenberg¹ and T. Treasure⁷

¹Department of Cardio-Thoracic Surgery, Erasmus University Medical Centre, Rotterdam, The Netherlands, ²Wales Cancer Trials Unit, Cardiff University, Cardiff, ³Medical Research Council Biostatistics Unit, Institute of Public Health, University of Cambridge, Cambridge, ⁴Division of Surgery and Cancer, and Imperial College Trials Unit, Imperial College London, ⁵Surgical and Interventional Trials Unit, Division of Surgery and Interventional Science, Faculty of Medical Sciences, University College London, London, UK, ⁶Oncology Centre, Hospital Alemão Oswaldo Cruz, Sao Paulo, Brazil, and ⁷Clinical Operational Research Unit, University College London, London, UK

Correspondence to: Professor T. Treasure, Clinical Operational Research Unit (UCL), 4 Taviton Street, London WC1H 0BT, UK (e-mail: tom.treasure@gmail.com)

Background: After potentially curative resection of primary colorectal cancer, patients may be monitored by measurement of carcinoembryonic antigen and/or CT to detect asymptomatic metastatic disease earlier.

Methods: A systematic review and meta-analysis was conducted to find evidence for the clinical effectiveness of monitoring in advancing the diagnosis of recurrence and its effect on survival. MEDLINE (Ovid), Embase, the Cochrane Library, Web of Science and other databases were searched for randomized comparisons of increased intensity monitoring compared with a contemporary standard policy after resection of primary colorectal cancer.

Results: There were 16 randomized comparisons, 11 with published survival data. More intensive monitoring advanced the diagnosis of recurrence by a median of 10 (i.q.r. 5–24) months. In ten of 11 studies the authors reported no demonstrable difference in overall survival. Seven RCTs, published from 1995 to 2016, randomly assigned 3325 patients to a monitoring protocol made more intensive by introducing new methods or increasing the frequency of existing follow-up protocols *versus* less invasive monitoring. No detectable difference in overall survival was associated with more intensive monitoring protocols (hazard ratio 0.98, 95 per cent c.i. 0.87 to 1.11).

Conclusion: Based on pooled data from randomized trials published from 1995 to 2016, the anticipated survival benefit from surgical treatment resulting from earlier detection of metastases has not been achieved.

Presented to the Third Preventing Overdiagnosis Conference, Bethesda, Maryland, USA, September 2015

Paper accepted 17 May 2016

Published online 4 August 2016 in Wiley Online Library (www.bjs.co.uk). DOI: 10.1002/bjs.10233

Introduction

A variety of monitoring strategies have been used in patients who have had potentially curative surgery for primary colorectal cancer. Their aim has been to detect active disease before it is symptomatic or clinically evident so that further treatment can be instigated. Five randomized trials published from 1995 to 1998 were the subject of a systematic review and meta-analysis published in 2002¹. Intensive follow-up was associated with significantly earlier detection by a mean of 8.5 months. The combined risk ratio was 0.81 (95 per cent c.i. 0.70 to 0.94) in favour of intensive follow-up. However, the authors found that methods were poorly reported and concluded that

‘large trials are required to identify which components of intensive follow up are most beneficial’. Since then, three large trials^{2–4} of intensified monitoring have reported. An updated search, systematic review and meta-analysis have been undertaken to examine the effect of these programmes on overall survival including all randomized studies identified.

Methods

A systematic review of literature on follow-up strategies for patients with colorectal cancer was conducted according to the PRISMA guidelines⁵ and is registered in PROSPERO

(CRD42015026835). This study was based on predefined eligibility criteria and conducted according to a predefined methodological approach.

Search strategy

An extensive search for published articles was conducted in collaboration with a medical librarian. The electronic databases of MEDLINE (Ovid), Embase, the Cochrane Library and Web of Science, Scopus, CINAHL (EBSCO), PubMed publisher, Google Scholar, LILACS, SciELO and ProQuest were searched. The searches identified four index terms: large intestinal cancer, surgery, periodical surveillance and mortality or survival. Appropriate thesaurus terms (for MEDLINE, Embase and CINAHL) and keywords in the title and/or abstract were combined by Boolean logical operators, and adapted to the appropriate syntax of each database. The reference lists of reviews and included studies were cross-checked.

Selection of studies

Papers were screened by two independent investigators, arbitrated by a third reviewer. Data were extracted from studies reporting randomly assigned groups of patients in surveillance protocols of differing intensity. Only studies conducted in humans and written in English were included. Studies with inadequate data on survival for meta-analysis were retained for textual summaries of the design, findings and conclusions.

Outcome measures

The primary outcome was the overall survival difference between the existing monitoring strategy compared with a more intensive monitoring strategy.

Quality control

Studies were checked independently for quality using the Cochrane risk of bias tool⁶. The authors of all studies were approached for further information⁷.

Data extraction

Data were extracted by one researcher and checked independently by a second reviewer. A third investigator resolved any discrepancies. Patient numbers, baseline characteristics, all-cause mortality, cancer-specific mortality and recurrence rates were retrieved for each study.

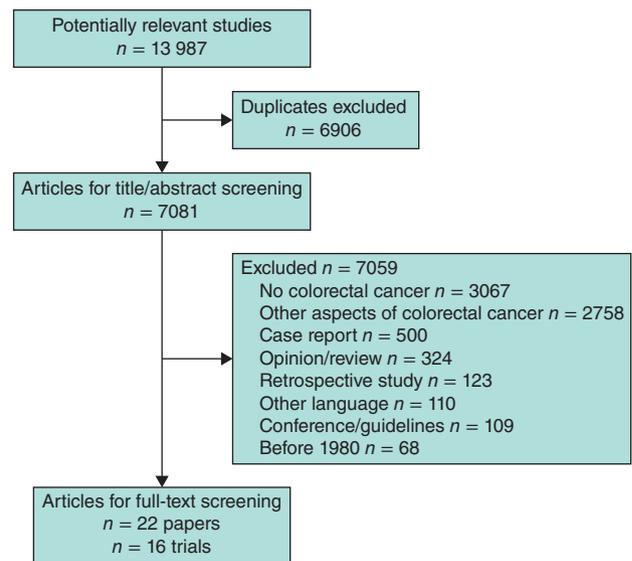


Fig. 1 Flow chart showing selection of trials for review

Overall survival data were extracted as event rates reported for more *versus* less intensive monitoring arms of all randomized comparisons. Odds ratios (ORs) and their variances were calculated. Hazard ratios (HRs) were derived from Kaplan–Meier curves. The method described by Williamson and colleagues⁸ was used to estimate a logarithmic HR with corresponding variance when the number of patients at risk was given at each time point. If these data were not provided, the method of Parmar *et al.*⁹ was used. The overall HR with 95 per cent c.i. was estimated using an inverse variance-weighted average¹⁰.

Statistical analysis

Review Manager (RevMan) for Windows[®] version 5.3 (Nordic Cochrane Centre, Cochrane Collaboration, Copenhagen, Denmark) was used for meta-analysis. Funnel plots were used to investigate publication bias. Heterogeneity among the included studies was analysed by means of the I^2 measure¹¹. A random-effects meta-analysis was performed after exclusion of trials with a high risk of bias.

ORs were also used to summarize observed effects, and a random-effects logistic regression model was used to provide an overall estimate of an effect for subsets of studies defined by the chosen method of enhanced detection. Subgroup analyses of outcome were performed to account for different diagnostic tests used during follow-up in different randomized trials. Studies were grouped as follows: any site of recurrence; endoscopically detected recurrence; or

Table 1 Studies excluded from meta-analysis

Trial	Start	End	Tests	No. of centres	No. of patients randomized	Reason for exclusion
Barillari <i>et al.</i> ¹⁸	1980	1990	Colonoscopy	1	212	Inadequate survival data
Schoemaker <i>et al.</i> ²¹	1984	1990	CT	1	325	Potential lack of allocation concealment
Secco <i>et al.</i> ²³	1988	1996	Colonoscopy CEA	1	337	Reviewers could not reconcile conclusions with randomized groups
COLOFOL ²⁸	2006	2011	Colonoscopy CEA CT	24	2571	Results not yet published
CEAwatch ³¹	2010	2012	Ultrasonography CEA	11	3223	No outcome data reported

CEA, carcinoembryonic antigen.

Table 2 Details of the seven trials included in meta-analysis

Trial	Start	End	Tests*	No. of centres	No. of patients randomized	Authors' conclusion
CEASL ³⁰	1982	1993	CEA	58	216†	'... highly unlikely that any survival advantage would be demonstrated for patients undergoing second-look surgery'
Ohlsson <i>et al.</i> ¹⁷	1983	1986	Endoscopy CT CEA	2	107	'Intense follow-up ... did not prolong survival in this study'
Pietra <i>et al.</i> ²²	1987	1990	CEA Ultrasonography CT	1	207	'Our data support use of an intense follow-up plan after primary resection of large-bowel cancer, at least in patients with rectal cancer'
Mäkelä <i>et al.</i> ¹⁶	1988	1990	Chest X-ray Colonoscopy CEA Chest X-ray CT	1	106	'Earlier detection of recurrent carcinoma by intensified follow-up does not lead to increased re-resectability or improved 5-year survival'
Rodriguez -Moranta <i>et al.</i> ²	1997	2001	CEA Colonoscopy CT Ultrasonography	3	259	'there was no difference in the probability of overall survival'
GILDA ⁴	1998	2006	Chest X-ray CEA Colonoscopy Chest X-ray Ultrasonography	41	1228	'early diagnosis of cancer recurrence is not associated with overall survival benefit'
FACS ³	2003	2009	CEA CT	39	1202	'The number of deaths was not significantly different in the combined intensive monitoring groups vs the minimum follow-up group'

*There were more tests, more frequent tests, or both in the group with more intensive monitoring. †If the carcinoembryonic antigen (CEA) level was raised according to study criteria, patients were randomized to have this revealed to the clinical team or not.

the clinical setting of follow-up. Sensitivity analyses were performed to identify studies that were estimated to have a high risk of bias.

Meta-analysis was undertaken and forest plots were constructed for all trials that reached the criteria for inclusion. Since the previous meta-analysis calling for large trials¹, there have been three large multicentre trials, published in 2006, 2014 and 2016²⁻⁴ relating to policies of earlier detection of patients suitable for the growing practice of metastasectomy¹²⁻¹⁴. The analysis was repeated in this subset of trials.

Results

Among 7081 publications, there were 22 relevant articles^{2-4,15-33} describing 16 randomized comparisons (*Fig. 1*). Text summaries of all 16 randomized trials are provided in *Appendix S1* (supporting information). Five studies were excluded because there were no survival data available for analysis^{28,31} or they had high risk of bias^{18,21,23} (*Table 1*). The remaining 11 studies provided data on overall survival suitable for meta-analysis and, of these, seven^{2-4,16,17,22,30} included methods that allowed detection

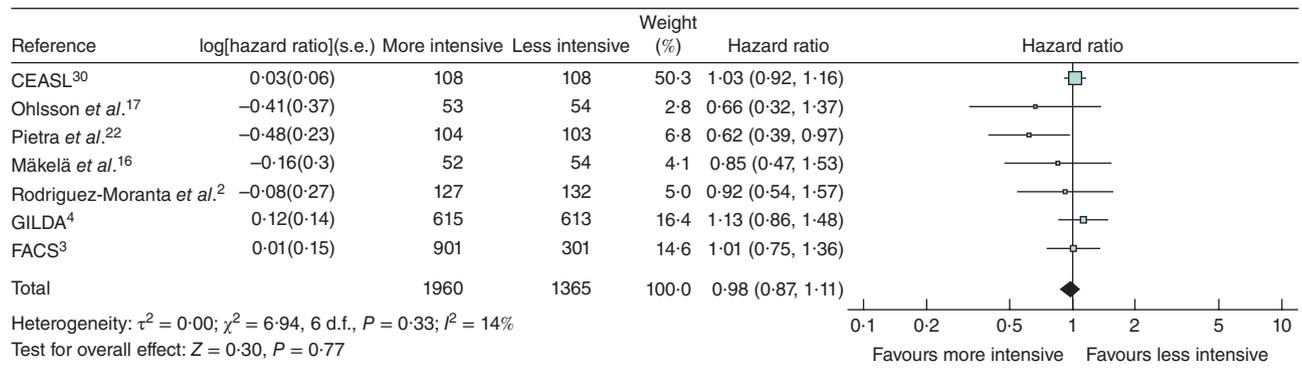


Fig. 2 Forest plot showing hazard ratios for death in seven randomized comparisons of more and less intensive follow-up from which hazard ratios could be derived. An inverse-variance random-effects model was used to produce an overall estimated hazards ratio. Hazard ratios are shown with 95 per cent confidence intervals. The studies are ordered according to the year of the start of the inclusion. CEASL is dominant because the weight of the study is dependent on the follow-up time, number of events and number of patients in each treatment arm. The Kaplan–Meier curve in CEASL is plotted up to 25 years. The point estimate in favour of more intensive monitoring in studies by Rodriguez-Moranta *et al.*² and Pietra and colleagues²² was attributed by the authors to detection by endoscopy and successful treatment of recurrent rectal carcinoma

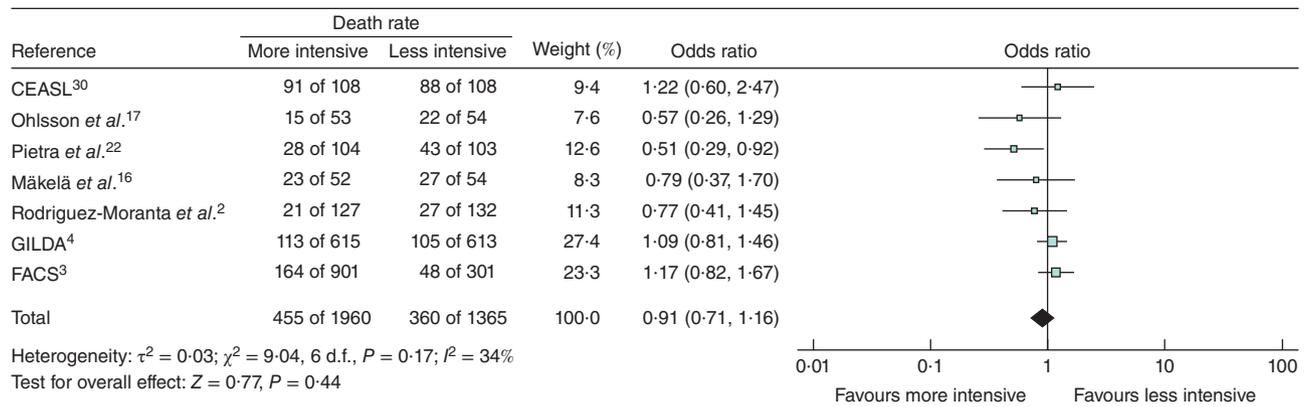


Fig. 3 Forest plot showing odds ratios for death in seven randomized comparisons of more and less intensive follow-up. A Mantel–Haenszel random-effects model was used to produce an overall estimated odds ratio. Odds ratios are shown with 95 per cent confidence intervals

Table 3 Details of single-centre trials confined to endoscopic methods of monitoring

Trial	Start	End	Tests	No. of patients	Authors' conclusions
Kjeldsen <i>et al.</i> ²⁰	1983	1994	Colonoscopy	597	'no improvement in overall survival or in cancer-related survival'
Wang <i>et al.</i> ²⁷	1995	2001	Colonoscopy	326	There was higher detection of asymptomatic recurrence and more operations but the authors concluded that more intensive colonoscopy 'did not improve overall survival'

Table 4 Risk of bias

Reference	Randomization method	Allocation concealment	Incomplete outcome assessment	Selective reporting
CEASL ³⁰	Low	Low	Low	Low
Ohlsson <i>et al.</i> ¹⁷	Unclear	Unclear	Low	Low
Pietra <i>et al.</i> ²²	Unclear	Unclear	Low	Low
Rodriguez-Moranta <i>et al.</i> ²	Low	Low	Low	Low
Mäkelä <i>et al.</i> ¹⁶	Unclear	Unclear	Low	Low
GILDA ⁴	Unclear	Low	Low	Low
FACS ³	Unclear	Unclear	Low	Low
Kjeldsen <i>et al.</i> ²⁰	Unclear	Unclear	High*	Low
Wang <i>et al.</i> ²⁷	Unclear	Low	Unclear	Low
Wattchow <i>et al.</i> ²⁴	Unclear	Low	Low	Low
Augestad <i>et al.</i> ²⁵	Low	Low	Low	Low
Schoemaker <i>et al.</i> ²¹	Low	High†	Unclear	Low
Secco <i>et al.</i> ²³	High‡	Unclear	High§	Low
CEAwatch ³¹	Low	Low	Low	Low

*Groups not balanced (290 : 307); drop-outs may not be included in assessment. †Cards not in envelopes; groups not balanced (167 : 158). ‡Method unclear; groups not balanced (108 : 84 and 84 : 61). §Patients excluded from survival rather than censored if lost to follow-up. COLOFOL has not reported so cannot be assessed; the methods of assignment were not described by Barillari *et al.*¹⁸, and this did not appear to be a true randomized comparison.

Table 5 Cancer recurrence rates and difference in time to detection in RCTs of monitoring strategies following potentially curative resection of colorectal cancer

Reference	Recruitment	Methods tested*	Recurrence (%)		Detection advance (months)
			Intensive	Control	
CEASL ³⁰	1982–1993	CEA	–†	–†	11
Ohlsson <i>et al.</i> ¹⁷	1983–1986	CEA, CT, endoscopy			4
Kjeldsen <i>et al.</i> ²⁰	1983–1984	Endoscopy	26	26	9
Mäkelä <i>et al.</i> ¹⁶	1988–1990	CT	42	39	5
Wang <i>et al.</i> ²⁷	1995–2001	Endoscopy	8	11	13
GILDA ⁴	1998–2006	CT, endoscopy, liver ultrasonography	15	13	6
Wattchow <i>et al.</i> ²⁴	1998–2001	Setting: surgeon- or GP-led			2
FACS ³					
CEA	2003–2009	CEA	22	14	24
CT	2004–2009	CT	20	14	30
CEA and CT	2005–2009	CEA, CT	17.5	14	24

*Only tests that were not the same in the two groups. †Only a minority of patients meeting the stringent carcinoembryonic antigen (CEA) criteria were randomized, so detection was similar in both randomized groups and effectively 100 per cent.

Table 6 All-cause mortality rates in randomized trials

	All-cause mortality	
	Intensive follow-up	Less intensive follow-up
CEASL ³⁰	91 of 108 (84.3)	88 of 108 (81.5)
Mäkelä <i>et al.</i> ¹⁶	23 of 52 (44)	27 of 54 (50)
Ohlsson <i>et al.</i> ¹⁷	15 of 53 (28)	22 of 54 (41)
Kjeldsen <i>et al.</i> ²⁰	88 of 290 (30.3)	100 of 307 (32.6)
Pietra <i>et al.</i> ²²	28 of 104 (26.9)	43 of 103 (41.7)
Schoemaker <i>et al.</i> ²¹	43 of 167 (25.7)	55 of 158 (34.8)
Secco <i>et al.</i> ²³	73 of 192 (38.0)	81 of 145 (55.9)
GILDA ⁴	113 of 615 (18.4)	105 of 613 (17.1)
Rodriguez-Moranta <i>et al.</i> ²	21 of 127 (16.5)	27 of 132 (20.5)
Wattchow <i>et al.</i> ²⁴	32 of 76 (42)	25 of 81 (31)
Wang <i>et al.</i> ²⁷	42 of 165 (25.5)	50 of 161 (31.1)
Augestad <i>et al.</i> ²⁵	1 of 55 (2)	4 of 55 (7)
FACS ^{3,34}	164 of 901 (18.2)	48 of 301 (15.9)
CEAwatch ³¹	n.r.	n.r.
COLOFOL ²⁸	n.r.	n.r.
Total	734 of 2905 (25.3)	675 of 2272 (29.7)

Values in parentheses are percentages. n.r., Not reported.

Table 7 Local recurrence rates in randomized trials

	Local recurrence	
	Intensive follow-up	Less intensive follow-up
CEASL ³⁰	n.r.	n.r.
Mäkelä <i>et al.</i> ¹⁶	12 of 52 (23)	11 of 54 (20)
Ohlsson <i>et al.</i> ¹⁷	11 of 53 (21)	8 of 54 (15)
Kjeldsen <i>et al.</i> ^{19,20}	49 of 290 (16.9)	42 of 307 (13.7)
Pietra <i>et al.</i> ²²	20 of 104 (19.2)	26 of 103 (25.2)
Schoemaker <i>et al.</i> ²¹	7 of 167 (4.2)	11 of 158 (7.0)
Secco <i>et al.</i> ²³	41 of 192 (21.4)	35 of 145 (24.1)
GILDA ⁴	36 of 615 (5.9)	32 of 613 (5.2)
Rodriguez-Moranta <i>et al.</i> ²	11 of 127 (8.7)	13 of 132 (10.1)
Wattchow <i>et al.</i> ²⁴	n.r.	n.r.
Wang <i>et al.</i> ²⁷	10 of 165 (6.1)	12 of 161 (7.5)
Augestad <i>et al.</i> ²⁵	6 of 55 (11)	8 of 55 (15)
FACS ^{3,34}	35 of 901 (3.9)	6 of 301 (2.0)
CEAwatch ³¹	31 of 316 (9.8)	13 of 1182 (1.1)
COLOFOL ²⁸	n.r.	n.r.
Total	269 of 3037 (8.9)	217 of 3265 (6.6)

Values in parentheses are percentages. n.r., Not reported.

Table 8 Distant recurrence rates in randomized trials

	Distant recurrence	
	Intensive follow-up	Less intensive follow-up
CEASL ³⁰	32 of 108 (29.6)	n.r.
Mäkelä <i>et al.</i> ¹⁶	10 of 52 (19)	10 of 54 (19)
Ohlsson <i>et al.</i> ¹⁷	9 of 53 (17)	12 of 54 (22)
Kjeldsen <i>et al.</i> ^{19,20}	34 of 290 (11.7)	48 of 307 (15.6)
Pietra <i>et al.</i> ²²	15 of 104 (14.4)	21 of 103 (20.4)
Schoemaker <i>et al.</i> ²¹	n.r.	n.r.
Secco <i>et al.</i> ²³	38 of 192 (19.8)	47 of 145 (32.4)
GILDA ⁴	59 of 615 (9.6)	42 of 613 (6.9)
Rodriguez-Moranta <i>et al.</i> ²	20 of 127 (15.7)	19 of 132 (14.4)
Wattchow <i>et al.</i> ²⁴	n.r.	n.r.
Wang <i>et al.</i> ²⁷	n.r.	n.r.
Augustad <i>et al.</i> ²⁵	3 of 55 (5)	4 of 55 (7)
FACS ^{3,34}	39 of 901 (4.3)	18 of 301 (6.0)
CEAwatch ³¹	n.r.	n.r.
COLOFOL ²⁸	n.r.	n.r.
Total	259 of 2497 (10.4)	221 of 1764 (12.5)

Values in parentheses are percentages. n.r., Not reported.

of metastases (Table 2, Figs 2 and 3). Two studies^{20,27} were confined to endoscopic examination following more or less intensive protocols (Table 3). Two studies^{24,25} followed the same protocol in each arm but were administered in hospital by a specialist or in a general practice setting.

Detailed protocols for investigations used for monitoring and their frequency are summarized in Appendix S2 (supporting information).

Quality of studies: risk of bias

Three studies were found to have a high risk of bias in at least one domain (Table 4). Blinding of participants

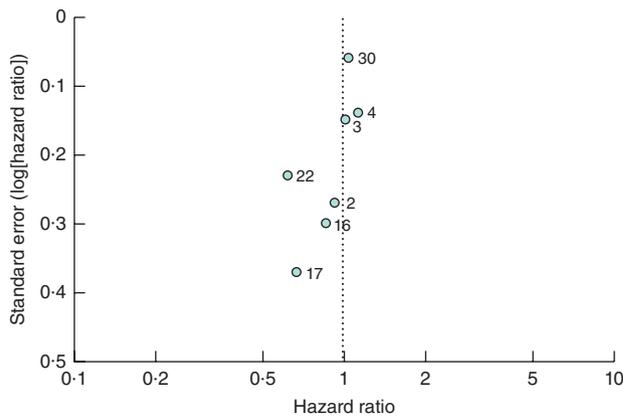


Fig. 5 Funnel plot of studies included in meta-analysis. Reference numbers are shown

and personnel (performance bias) was not possible and so there is a remaining risk of bias. Blinding of outcome assessment (detection bias) is not relevant for the main outcome measure, which is death/survival.

Effectiveness of more intensive monitoring in advancing detection

For studies in which the time difference in detection was given (9 of 16) the advance in diagnosis was 2–30 months, with a median of 10 (i.q.r. 5–24) months (Table 5).

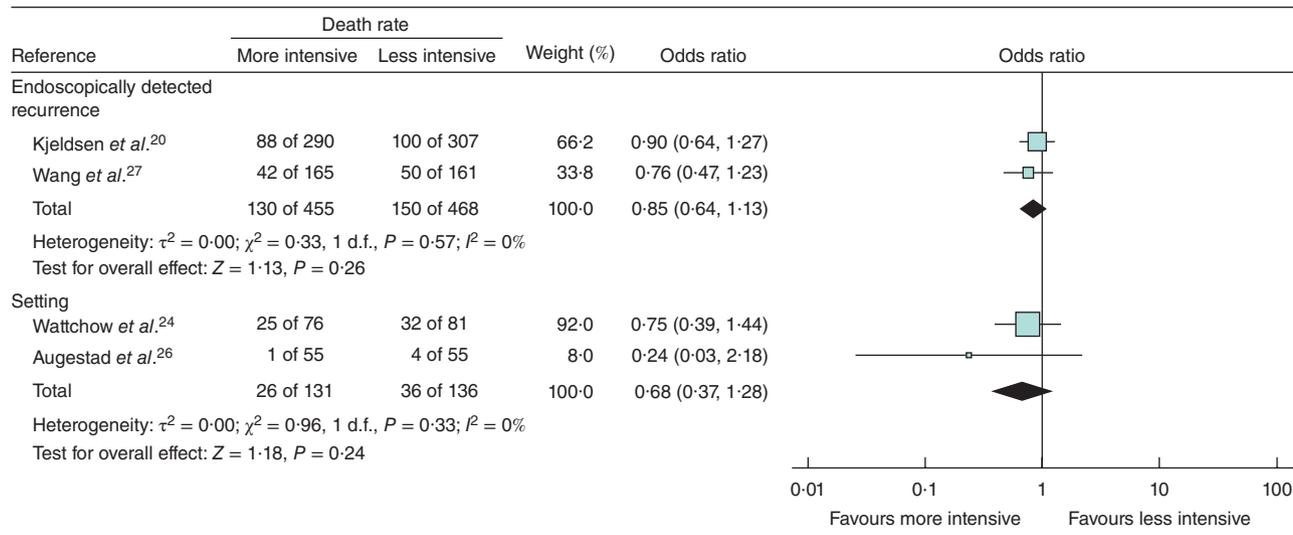


Fig. 4 Forest plot for death in two studies in which the difference in monitoring was confined to endoscopy, and two studies for which the difference was between a hospital/specialist setting and a general practice setting. A Mantel–Haenszel random-effects model was used to produce an overall estimated odds ratio. Odds ratios are shown with 95 per cent confidence intervals

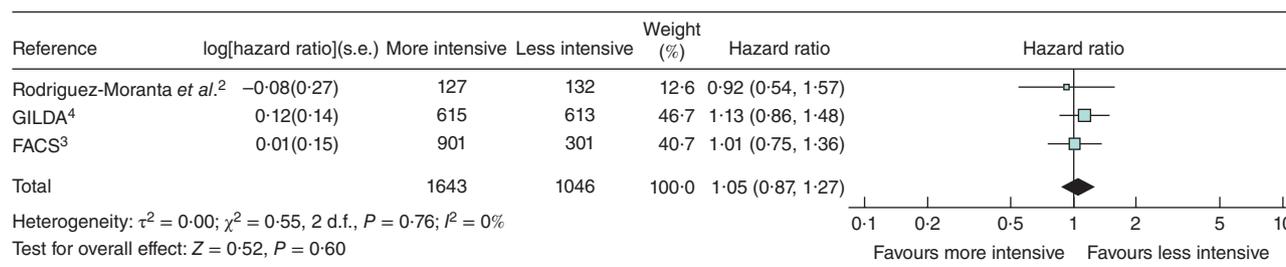


Fig. 6 Subset analysis of three large multicentre RCTs published in 2006, 2014 and 2016, which included 80.9 per cent of all patients in the full systematic review. A random-effects inverse-variance model was used to produce an overall estimated hazards ratio. Hazard ratios are shown with 95 per cent confidence intervals

Main outcome measure: effectiveness in improving survival

The numbers of randomized patients and the numbers of all detection events and deaths are given in *Tables 6–8*. The principal result derives from the meta-analysis of HRs based on trials from which these could be estimated. The summary HR estimate was 0.98 (95 per cent c.i. 0.87 to 1.11), with no evidence of significant heterogeneity ($I^2 = 14$ per cent) (*Fig. 2*).

The meta-analysis of simple ORs for death based on the same seven studies is shown in *Fig. 3*; the summary OR for death was 0.91 (95 per cent c.i. 0.71 to 1.16). ORs were derived from the percentage of deaths in each arm at the time of reporting, whereas the HR gives an estimate of the overall relative survival, which is more relevant when considering a time-to-event endpoint.

A meta-analysis of ORs for death is also shown for two studies in which the difference in monitoring was confined to endoscopy, and two studies for which the difference was between a hospital/specialist setting and a general practice setting (*Fig. 4*).

There is a residual possibility of publication bias as demonstrated in the asymmetry of the forest plot for the main analysis (*Fig. 2*) and the funnel plot (*Fig. 5*). Because of the proportion of the weight accredited by RevMan to CEASL (Carcino-Embryonic Antigen Second Look), and because in this study the monitoring was solely by CEA and not CT, a sensitivity analysis was performed after exclusion of CEASL. This did not alter the conclusion (HR 0.92, 0.76 to 1.36).

The three trials reporting from 2006 to 2016 were larger, multicentre, better quality studies and contributed 2689 (80.9 per cent) of the 3325 patients included in the overall meta-analysis (*Fig. 2*). The summary estimated HR from a meta-analysis of these trials was 1.05 (0.87 to 1.27) (*Fig. 6*).

Discussion

The question addressed by this systematic review is whether follow-up strategies involving more intensive monitoring, with more frequent investigation and/or additional methods of detection, lead to an improvement in overall survival. Meta-analyses of the ORs and derived HRs from seven RCTs including 3325 patients showed no survival benefit from successively intensified monitoring policies. The authors of ten^{2–4,16,17,20,24,25,27,30} of the 11 trials reported no survival benefit from more intensive monitoring. This could have been due to lack of power, but the meta-analysis is consistent with the findings of individual studies. Three studies^{4,24,26} that reported outcomes on quality of life found no differences with respect to this outcome.

A random-effects model and, where possible, HRs were used to quantify outcomes, as the outcomes of interest occur over time. Owing to data limitations, ORs were used for some comparisons. Where it was possible to do both analyses, there was no difference in the conclusion. A limitation of this systematic review is that publication bias may have affected the observed outcomes as unpublished data, abstracts and presentations were not included. However, consideration of this possible bias would likely make a survival benefit even less plausible.

Intensive follow-up was reported to show significantly improved survival by the authors of only one study²². It was attributed to reappearance of treatable residual disease after resection of rectal (as opposed to colonic) cancer. In exploratory subset analyses, other studies^{2,4} showed a survival difference in favour of more intensive monitoring benefit where local recurrence of rectal cancer was found endoscopically.

Earlier meta-analyses^{1,35} suggested a favourable effect on survival. This was not found in the present meta-analysis, which included larger and methodologically more robust

trials reported in the past couple of years. These showed poorer survival in the more intensively screened groups, although the results were not significant individually. The survival results of this updated meta-analysis show no benefit. Although not statistically significant, the point estimates consistently suggest an adverse effect on survival. What these later RCTs have in common is that they were multi-centre studies run from trial centres. It is possible that the commitment of physicians involved in follow-up has made a difference to outcomes in smaller institutional studies. Many of these patients will have had individualized treatment including systemic chemotherapy, but no overall benefit from monitoring and earlier detection has been shown in the meta-analysis. This analysis gives a coherent and trustworthy, but disappointingly negative, message about the hoped-for survival benefit of intensification of active monitoring after primary resection of colorectal cancer.

Acknowledgements

The authors thank W. M. Bramer (Erasmus Medical Centre) for performing the search; and C. Russell (Chair of the Pulmonary Metastectomy in Colorectal Cancer Trial Steering Committee) and C. Brew-Graves (leader of the Surgical and Interventional Trials Unit (SITU), University College London), who actively supported this work, some of which was carried out within the SITU. The following authors of papers reporting RCTs have corresponded with the review authors and agreed to be acknowledged: A. Castells (Rodriguez-Moranta *et al.*), C. Verberne (CEAwatch), D. Schoemaker (Schoemaker *et al.*), G. Rosati, R. Fossati (GILDA), K. M. Augestad (Augestad *et al.*), L. Roncorini (Pietra *et al.*) and P. A. Wille-Jørgensen (COLOFOL).

This study received no specific funding. F.F. is partly funded by the British Heart Foundation.

Disclosure: The authors declare no conflict of interest.

References

- 1 Renehan AG, Egger M, Saunders MP, O'Dwyer ST. Impact on survival of intensive follow up after curative resection for colorectal cancer: systematic review and meta-analysis of randomised trials. *BMJ* 2002; **324**: 813.
- 2 Rodriguez-Moranta F, Saló J, Arcusa A, Boadas J, Piñol V, Bessa X *et al.* Postoperative surveillance in patients with colorectal cancer who have undergone curative resection: a prospective, multicenter, randomized, controlled trial. *J Clin Oncol* 2006; **24**: 386–393.
- 3 Primrose JN, Perera R, Gray A, Rose P, Fuller A, Corkhill A *et al.* FACS Trial Investigators. Effect of 3 to 5 years of scheduled CEA and CT follow-up to detect recurrence of colorectal cancer: the FACS randomized clinical trial. *JAMA* 2014; **311**: 263–270.
- 4 Rosati G, Ambrosini G, Barni S, Andreoni B, Corradini G, Luchena G *et al.*; GILDA working group. A randomized trial of intensive *versus* minimal surveillance of patients with resected Dukes B2-C colorectal carcinoma. *Ann Oncol* 2016; **27**: 274–280.
- 5 Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009; **339**: b2535.
- 6 Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD *et al.* The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011; **343**: d5928.
- 7 Roberts I, Ker K, Edwards P, Beecher D, Manno D, Sydenham E. The knowledge system underpinning healthcare is not fit for purpose and must change. *BMJ* 2015; **350**: h2463.
- 8 Williamson PR, Smith CT, Hutton JL, Marson AG. Aggregate data meta-analysis with time-to-event outcomes. *Stat Med* 2002; **21**: 3337–3351.
- 9 Parmar MK, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. *Stat Med* 1998; **17**: 2815–2834.
- 10 Tierney JF, Stewart LA, Ghersi D, Burdett S, Sydes MR. Practical methods for incorporating summary time-to-event data into meta-analysis. *Trials* 2007; **8**: 16.
- 11 Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**: 557–560.
- 12 Barton MK. Metastectomy on the rise across several cancer types. *CA Cancer J Clin* 2015; **65**: 163–164.
- 13 Bartlett EK, Simmons KD, Wachtel H, Roses RE, Fraker DL, Kelz RR *et al.* The rise in metastectomy across cancer types over the past decade. *Cancer* 2015; **121**: 747–757.
- 14 Jawed I, Wilkerson J, Prasad V, Duffy AG, Fojo T. Colorectal cancer survival gains and novel treatment regimens: a systematic review and analysis. *JAMA Oncol* 2015; **1**: 787–795.
- 15 Makela J, Laitinen S, Kairaluoma MI. Early results of follow-up after radical resection for colorectal cancer. Preliminary results of a prospective randomized trial. *Surg Oncol* 1992; **1**: 157–161.
- 16 Mäkelä JT, Laitinen SO, Kairaluoma MI. Five-year follow-up after radical surgery for colorectal cancer. Results of a prospective randomized trial. *Arch Surg* 1995; **130**: 1062–1067.
- 17 Ohlsson B, Breland U, Ekberg H, Graffner H, Tranberg KG. Follow-up after curative surgery for colorectal carcinoma. Randomized comparison with no follow-up. *Dis Colon Rectum* 1995; **38**: 619–626.
- 18 Barillari P, Ramacciato G, Manetti G, Bovino A, Sammartino P, Stipa V. Surveillance of colorectal cancer: effectiveness of early detection of intraluminal recurrences

- on prognosis and survival of patients treated for cure. *Dis Colon Rectum* 1996; **39**: 388–393.
- 19 Kjeldsen BJ, Kronborg O, Fenger C, Jørgensen OD. The pattern of recurrent colorectal cancer in a prospective randomised study and the characteristics of diagnostic tests. *Int J Colorectal Dis* 1997; **12**: 329–334.
 - 20 Kjeldsen BJ, Kronborg O, Fenger C, Jørgensen OD. A prospective randomized study of follow-up after radical surgery for colorectal cancer. *Br J Surg* 1997; **84**: 666–669.
 - 21 Schoemaker D, Black R, Giles L, Toouli J. Yearly colonoscopy, liver CT, and chest radiography do not influence 5-year survival of colorectal cancer patients. *Gastroenterology* 1998; **114**: 7–14.
 - 22 Pietra N, Sarli L, Costi R, Ouchemi C, Grattarola M, Peracchia A. Role of follow-up in management of local recurrences of colorectal cancer: a prospective, randomized study. *Dis Colon Rectum* 1998; **41**: 1127–1133.
 - 23 Secco GB, Fardelli R, Gianquinto D, Bonfante P, Baldi E, Ravera G *et al*. Efficacy and cost of risk-adapted follow-up in patients after colorectal cancer surgery: a prospective, randomized and controlled trial. *Eur J Surg Oncol* 2002; **28**: 418–423.
 - 24 Wattchow DA, Weller DP, Esterman A, Pilotto LS, McGorm K, Hammett Z. General practice *vs* surgical-based follow-up for patients with colon cancer: randomised controlled trial. *Br J Cancer* 2006; **94**: 1116–1121.
 - 25 Augestad KM, Vonen B, Aspevik R, Nestvold T, Ringberg U, Johnsen R. Should the surgeon or the general practitioner (GP) follow up patients after surgery for colon cancer? A randomized controlled trial protocol focusing on quality of life, cost-effectiveness and serious clinical events. *BMC Health Serv Res* 2008; **8**: 137.
 - 26 Augestad KM, Norum J, Dehof S, Aspevik R, Ringberg U, Nestvold T *et al*. Cost-effectiveness and quality of life in surgeon *versus* general practitioner-organised colon cancer surveillance: a randomised controlled trial. *BMJ Open* 2013; **3**: pii: e002391.
 - 27 Wang T, Cui Y, Huang WS, Deng YH, Gong W, Li CJ. The role of postoperative colonoscopic surveillance after radical surgery for colorectal cancer: a prospective, randomized clinical study. *Gastrointest Endosc* 2009; **69**: 609–615.
 - 28 Wille-Jørgensen P, Laurberg S, Pählman L, Carriquiry L, Lundqvist N, Smedh K *et al*. An interim analysis of recruitment to the COLOFOL trial. *Colorectal Dis* 2009; **11**: 756–758.
 - 29 Hansdotter Andersson P, Wille-Jørgensen P, Horváth-Puhó E, Petersen SH, Martling A, Sørensen HT *et al*. The COLOFOL trial: study design and comparison of the study population with the source cancer population. *Clin Epidemiol* 2016; **8**: 15–21.
 - 30 Treasure T, Monson K, Fiorentino F, Russell C. The CEA Second-Look Trial: a randomised controlled trial of carcinoembryonic antigen prompted reoperation for recurrent colorectal cancer. *BMJ Open* 2014; **4**: e004385.
 - 31 Verberne CJ, Zhan Z, van den Heuvel E, Grossmann I, Doornbos PM, Havenga K *et al*. Intensified follow-up in colorectal cancer patients using frequent carcino-embryonic antigen (CEA) measurements and CEA-triggered imaging: results of the randomized ‘CEAwatch’ trial. *Eur J Surg Oncol* 2015; **41**: 1188–1196.
 - 32 Verberne CJ, Wiggers T, Grossmann I, de Bock GH, Vermeulen KM. Cost-effectiveness of a carcinoembryonic antigen (CEA)-based follow-up programme for colorectal cancer (the CEA Watch trial). *Colorectal Dis* 2016; **18**: O91–O96.
 - 33 Grossmann EM, Johnson FE, Virgo KS, Longo WE, Fossati R. Follow-up of colorectal cancer patients after resection with curative intent – the GILDA trial. *Surg Oncol* 2004; **13**: 119–124.
 - 34 Pugh SA, Shinkins B, Fuller A, Mellor J, Mant D, Primrose JN. Site and stage of colorectal cancer influence the likelihood and distribution of disease recurrence and postrecurrence survival: data from the FACS randomized controlled trial. *Ann Surg* 2016; **263**: 1143–1147.
 - 35 Jeffery GM, Hickey BE, Hider P. Follow-up strategies for patients treated for non-metastatic colorectal cancer. *Cochrane Database Syst Rev* 2002; (1)CD002200.

Supporting information

Additional supporting information may be found in the online version of this article:

Appendix S1 Text summaries of 16 randomized trials (Word document)

Appendix S2 Spreadsheet of all monitoring methods (Excel spreadsheet)

Editor's comment

Cancer is a frightening word that brings uncertainty to patients' lives and those of their loved ones. It is our role to dispel fear, establish trust, and deliver realistic expectations for the future. Tests give confidence to both parties if sensitive, specific, and comfortable, as well as cost-effective and generalizable. Clinical assessment, quality colonoscopy, serum carcinoembryonic antigen testing, and baseline CT fulfil all those targets at diagnosis. Repetition of scans every 6–12 months is in doubt if the findings presented here do not represent immortal time bias or other anomaly. If so, then we understand pitifully little of the biological interplay between host and disease that evade the (evidently) crude instruments of this age.

Cancer outcomes improved with specialized surgery, multidisciplinary management, and molecular-targeted therapy, so it is counterintuitive that earlier identification of metastatic disease does not improve survival. Intensive surveillance should prolong disease-specific survival by bringing forward the date of diagnosis and by rendering the disease burden easier to manage. Far from being the bleak nihilism some will choose to see, this paper will stimulate evolution (not cessation) of surveillance and metastasis-directed healthcare. It challenges us to determine how to follow patients optimally, to whom to offer intensive monitoring, and to better define the molecular complexity of tumour-immune interfaces. Fear of the unknown is a burden we have allowed patients to shoulder while we peer over the horizon of their personal voyage. If you are not alarmed that modern colorectal cancer care is suboptimal after reading this paper, please reflect on it again.

D. C. Winter
Editor, BJS

