Cochrane Library

Cochrane Database of Systematic Reviews

Follow-up strategies for patients treated for non-metastatic colorectal cancer (Review)

Jeffery M, Hickey BE, Hider PN, See AM

Jeffery M, Hickey BE, Hider PN, See AM.
Follow-up strategies for patients treated for non-metastatic colorectal cancer.
Cochrane Database of Systematic Reviews 2016, Issue 11. Art. No.: CD002200.
DOI: 10.1002/14651858.CD002200.pub3.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Follow-up strategies for patients treated for non-metastatic colorectal cancer

Mark Jeffery[1], Brigid E Hickey[2,3], Phil N Hider[4], Adrienne M See[2]

[1]Canterbury Regional Cancer and Haematology Service, Christchurch Hospital, Christchurch, New Zealand. [2]Radiation Oncology Mater Service, Princess Alexandra Hospital, Brisbane, Australia. [3]School of Medicine, The University of Queensland, Brisbane, Australia. [4]Department of Population Health, University of Otago, Christchurch, Christchurch, New Zealand

Contact address: Mark Jeffery, Canterbury Regional Cancer and Haematology Service, Christchurch Hospital, Private Bag 4710, Christchurch, 8140, New Zealand. mark.jeffery@cdhb.govt.nz.

Editorial group: Cochrane Colorectal Cancer Group.
Publication status and date: Edited (conclusions changed), published in Issue 11, 2016.
Review content assessed as up-to-date: 20 May 2016.

Citation: Jeffery M, Hickey BE, Hider PN, See AM. Follow-up strategies for patients treated for non-metastatic colorectal cancer. *Cochrane Database of Systematic Reviews* 2016, Issue 11. Art. No.: CD002200. DOI: 10.1002/14651858.CD002200.pub3.

ABSTRACT

Background

It is common clinical practice to follow patients with colorectal cancer (CRC) for several years following their curative surgery or adjuvant therapy, or both. Despite this widespread practice, there is considerable controversy about how often patients should be seen, what tests should be performed, and whether these varying strategies have any significant impact on patient outcomes. This is the second update of a Cochrane Review first published in 2002 and first updated in 2007.

Objectives

To assess the effects of intensive follow-up for patients with non-metastatic colorectal cancer treated with curative intent.

Search methods

For this update, we searched CENTRAL (2016, Issue 3), MEDLINE (1950 to May 20th, 2016), Embase (1974 to May 20th, 2016), CINAHL (1981 to May 20th, 2016), and Science Citation Index (1900 to May 20th, 2016). We also searched reference lists of articles, and handsearched the Proceedings of the American Society for Radiation Oncology (2011 to 2014). In addition, we searched the following trials registries (May 20th, 2016): ClinicalTrials.gov and the World Health Organization International Clinical Trials Registry Platform. We further contacted study authors. No language or publication restrictions were applied to the search strategies.

Selection criteria

We included only randomised controlled trials comparing different follow-up strategies for participants with non-metastatic CRC treated with curative intent.

Data collection and analysis

Two authors independently determined trial eligibility, performed data extraction, and assessed methodological quality.

Main results

We studied 5403 participants enrolled in 15 studies. (We included two new studies in this second update.) Although the studies varied in setting (general practitioner (GP)-led, nurse-led, or surgeon-led) and "intensity" of follow-up, there was very little inconsistency in the results.

Overall survival: we found no evidence of a statistical effect with intensive follow-up (hazard ratio (HR) 0.90, 95% confidence interval (CI) 0.78 to 1.02; I^2 = 4%; P = 0.41; high-quality evidence). There were 1098 deaths among 4786 participants enrolled in 12 studies.

Colorectal cancer-specific survival: this did not differ with intensive follow-up (HR 0.93, 95% CI 0.78 to 1.12; I^2 = 0%; P = 0.45; moderate-quality evidence). There were 432 colorectal cancer deaths among 3769 participants enrolled in seven studies.

Relapse-free survival: we found no statistical evidence of effect with intensive follow-up (HR 1.03, 95% CI 0.90 to 1.18; I^2 = 5%; P = 0.39; moderate-quality evidence). There were 1416 relapses among 5253 participants enrolled in 14 studies.

Salvage surgery with curative intent: this was more frequent with intensive follow-up (risk ratio (RR) 1.98, 95% CI 1.53 to 2.56; I^2 = 31%; P = 0.14; high-quality evidence). There were 457 episodes of salvage surgery in 5157 participants enrolled in 13 studies.

Interval (symptomatic) recurrences: these were less frequent with intensive follow-up (RR 0.59, 95% CI 0.41 to 0.86; I^2 = 66%; P = 0.007; moderate-quality evidence). Three hundred and seventy-six interval recurrences were reported in 3933 participants enrolled in seven studies.

Intensive follow-up did not appear to affect quality of life, anxiety, nor depression (reported in three studies).

Harms from colonoscopies did not differ with intensive follow-up (RR 2.08, 95% CI 0.11 to 40.17; moderate-quality evidence). In two studies, there were seven colonoscopic complications in 2112 colonoscopies.

Authors' conclusions

The results of our review suggest that there is no overall survival benefit for intensifying the follow-up of patients after curative surgery for colorectal cancer. Although more participants were treated with salvage surgery with curative intent in the intensive follow-up group, this was not associated with improved survival. Harms related to intensive follow-up and salvage therapy were not well reported.

PLAIN LANGUAGE SUMMARY

Follow-up strategies for participants treated for non-metastatic colorectal cancer

What is the issue?

Colorectal cancer affects about 1 in 20 people in developed countries. Most patients (about two thirds) have curable disease. Follow-up after curative treatment usually means visits to the doctor as well as having some tests. Many people believe that follow-up saves lives, but we are not sure how often the patient should see the doctor and what tests they should have, and when.

Why is it important?

Follow-up is expensive, it can make patients anxious around the time of their visit, and can be inconvenient. Tests are expensive and can have side effects. If tests find that cancer has come back in a person who feels well, but treatment cannot cure them, finding the recurrent cancer may not have helped that person or their family.

We asked...

We asked if follow-up (i.e. tests and doctor visits) after colorectal cancer has been treated curatively is helpful. We looked at all different kinds of follow-up: some versus none; more tests versus fewer tests; and follow-up done by surgeons, general practitioners (GPs), or nurses.

We found...

We found 15 studies, including 5403 participants. We found that follow-up did not improve overall survival (high-quality evidence), colorectal cancer-specific survival (moderate-quality evidence), or relapse-free survival (moderate-quality evidence). If patients have follow-up, they are much more likely to have surgery if the cancer is detected again (high-quality evidence). With follow-up, more

asymptomatic "silent" cancer relapses are likely to be found at planned visits (moderate-quality evidence). Harms from tests were not common, but only two studies reported them (moderate-quality evidence). We found very little data on quality of life or costs.

This means...

The information we have now suggests that there is little benefit from intensifying follow-up, but there is also little evidence about quality of life, harms, and costs. We do not know what is the best way to follow patients treated for non-metastatic colorectal cancer, or if we should at all. We know little about the costs of follow-up in this setting. However, we found four ongoing trials (which will enrol a further 4801 participants); they will look at quality of life, harms, and costs, and may reveal a better understanding of what is the best follow-up programme. Consumer needs and concerns with respect to the value of follow-up require further research.

SUMMARY OF FINDINGS FOR THE MAIN COMPARISON [Explanation]

Follow-up strategies for patients treated for non-metastatic colorectal cancer

Patient or population: colorectal cancer treated with curative intent
Setting: tertiary hospitals or cancer centres
Intervention: intensive follow-up
Comparison: conventional follow-up

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with conventional follow-up	Risk with intensive follow-up				
Overall survival (OS) Follow-up: range = 24 months to 105 months (median)	Study population 242 per 1000	225 per 1,000 (192 to 260)	HR 0.92 (0.77 to 1.09)	4786 (12 RCTs)	⊕⊕⊕⊕ HIGH[1, 2]	17 fewer deaths per 1000 (between 50 fewer and 18 more)
Colorectal cancer-specific survival (CC-SS) Follow-up: range = 24 months to 105 months (median)	Study population 143 per 1000	133 per 1,000 (113 to 158)	HR 0.93 (0.78 to 1.12)	3822 (7 RCTs)	⊕⊕⊕◯ MODERATE[1, 2, 3]	10 fewer death per 1000 (between 30 fewer and 16 more)
Relapse-free survival (R-FS) Follow-up: range = 48 months to 120 months (median)	Study population 275 per 1000	282 per 1,000 (252 to 316)	HR 1.03 (0.90 to 1.18)	5253 (14 RCTs)	⊕⊕⊕◯ MODERATE[1, 2, 4]	7 more per 100 (between 24 fewer and 41 more)
Salvage surgery (SS) Follow-up: range = 24 months to 105 months (median)	Study population		RR 1.98 (1.53 to 2.56)	5157 (13 RCTs)	⊕⊕⊕⊕ HIGH	60 more episodes of salvage surgery (between 33 more and 96 more)

	62 per 1000	122 per 1,000 (94 to 157)				
Interval recurrences (IR) assessed with: recurrent CRC diagnosed between scheduled follow-up visits Follow-up: range = 43 months to 79 months (median)	Study population		RR 0.59 (0.41 to 0.86)	3933 (6 RCTs)	⊕⊕⊕◯ MODERATE[3,6]	52 fewer interval recurrences (between 18 fewer and 75 fewer)
	127 per 1000	75 per 1,000 (52 to 109)				
Adverse effects	Study population		RR 2.08 (0.11 to 40.42)	1381 (1 RCT)	⊕⊕⊕◯ MODERATE[7,8]	-
	0 per 1,000	0 per 1,000 (0 to 0)				

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

CI: confidence interval; **CRC**: colorectal cancer; **HR**: hazard ratio; **RCT**: randomised controlled trial; **RR:** Risk ratio; **OR:** Odds ratio

GRADE Working Group grades of evidence
High quality: We are very confident that the true effect lies close to that of the estimate of the effect
Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

[1] > 300 events.
[2] Confidence intervals include 1 and exclude clinically meaningful benefit or harm.
[3] Downgraded because there were 192/468 (41%) events from studies deemed at high risk of bias because of incomplete follow-up.
[4] Downgraded because there were 383/1120 (34%) events from studies deemed at high risk for lack of blinding.
[5] Not downgraded because prespecified sensitivity analysis explained heterogeneity on the basis of study age.

[6] Downgraded because there were 95/178 (58%) events from studies deemed at high risk of bias because of incomplete follow-up.

[7] Total number of events is less than 300.

[8] Denominator was number of colonoscopies performed.

BACKGROUND

Description of the condition

Colorectal cancer (CRC) is a commonly diagnosed malignancy affecting about one person in 20 in most westernised countries (DevCan 2005). Approximately two-thirds of patients will present with potentially curable disease (by surgery plus or minus adjuvant therapies). Of these, 50% to 60% will relapse with metastatic disease (Lee 2007; Van Cutsem 2006; Yoo 2006). Surgeons, oncologists, and other health professionals caring for CRC patients have pursued a number of strategies to try to improve clinical outcomes. These have included population screening, better diagnostic testing, improved surgical and anaesthetic techniques, and more widespread utilisation of effective adjuvant therapies.

After definitive treatment is completed, clinician attention turns to follow-up strategies designed to detect recurrence at a stage when further curative procedures can be used. Follow-up strategies have also been developed in order to detect new curable metachronous (i.e. occurring at different times) primary tumours. There is overlap between these two strategies; this current review focuses on the former issue of strategies designed to detect curable recurrences of the original cancer. These include recurrences that are localised in either the lung, liver, abdomen, or pelvis and can be completely resected or ablated with curative intent.

Following patients after definitive treatment for cancer has become a traditional component of medical care (Edelman 1997). It is likely that clinicians follow patients after curative treatment for CRC at least in part to provide positive feedback on their management but also to assess the toxicities of treatment and to provide more accurate outcome data (Audisio 1996). Patients and their clinicians develop relationships during treatment and follow-up that can make it hard to discharge patients and return responsibility for care back to their primary physician in the community (Audisio 2000). As a result, a practising clinician can accumulate a large number of follow-up patients, and the surveillance of this cohort consumes significant resources.

The opportunity cost of the resources involved is considerable, limiting the care that the clinician can provide for other individuals. Few clinicians restrict CRC follow-up visits to clinical examination only, and the temptation to order routine investigations is often reinforced by patients who desire tests to "prove" that their disease is under control (Audisio 2000; Kievit 2000). Clinicians justify this approach by claiming that recurrences are being detected earlier than would otherwise occur and that patient outcomes are improved as a result (Kievit 2000).

Description of the intervention

Follow-up programmes in colorectal cancer should be based on the anatomic and temporal patterns of recurrence (Audisio 2000; Edelman 1997). The most important phase of follow-up is the first two to three years after primary resection, as during this time, the majority of recurrences will become apparent (Böhm 1993; Ovaska 1989). The liver is the most common site of metastases from colorectal cancer. A small proportion of these patients (10% to 20%) will have liver metastases that are distributed within the liver in such a fashion that makes them amenable to surgical resection or ablation (Alberts 2005; Muratore 2007). Published series' of patients undergoing such surgical interventions (with significant numbers of long-term survivors) encourage this approach (Choti 2002; Kanas 2012; Pawlik 2005). A number of strategies have been proposed to detect liver metastases at an early stage in order to identify such patients; these include the monitoring of blood tests (liver function, level of serum carcinoembryonic antigen (CEA)), and routine imaging of the liver and lung (Fleischer 1989; Sugarbaker 1987).

The psychological outcomes of follow-up programmes for patients with cancer can be positive or negative. Positive outcomes include reassurance and support. The negative outcomes include false reassurance, increased anxiety, fear and disappointment associated with early detection of an incurable recurrence, morbidity and mortality associated with procedures performed as a result of abnormal results, and distress caused by false-positive results. Appropriate quality of life measurements could provide information about these outcomes.

Follow-up can be comprised of clinic visits, examinations, and tests (blood tests and endoscopic and radiological examinations). Intensive follow-up may consist of an increased frequency of clinic visits, tests, and examinations in comparison with none or fewer clinic visits, tests, and examinations.

How the intervention might work

Follow-up programmes in colorectal cancer are thought to increase the early detection of recurrence at a stage when further curative procedures can be used, as well as new curable metachronous primary tumours, thereby, improving survival outcomes (GILDA 1998).

Why it is important to do this review

Whether systematic follow-up can alter long-term clinical outcomes for CRC remains controversial (Pfister 2004). Whilst some commentators have concluded that follow-up is worthwhile (Gerdes 1990), others have questioned its effectiveness (Kievit 2000; McArdle 2000). The variation in follow-up programmes, in terms of timing and frequency of clinician visits and the investigations undertaken by clinicians, is considerable (Collopy 1992; Connor 2001; Vernava 1994; Virgo 1995). Routine follow-up has the potential to create psychological harm in patients, and any such disadvantages need to be outweighed by improved clinical

outcomes (such as overall survival) that matter to patients. Data from follow-up studies in other cancers (e.g. breast cancer and overall survival) is not encouraging in this regard (Rojas 2005). Therefore, we conducted a systematic review of randomised controlled trials exploring questions relating to the effectiveness of follow-up strategies in CRC patients treated with curative intent.

OBJECTIVES

To assess the effect of follow-up programmes (follow-up versus no follow-up, follow-up strategies of varying intensity, and follow-up in different healthcare settings) on overall survival for patients with colorectal cancer treated with curative intent. Secondary endpoints included relapse-free survival, salvage surgery, interval recurrences, quality of life, and the harms and costs of surveillance and investigations.

METHODS

Criteria for considering studies for this review

Types of studies

Randomised controlled trials (RCTs) comparing different follow-up strategies for participants with colorectal cancer (CRC). These included comparisons of follow-up versus no follow-up, follow-up strategies of varying intensity (differing frequency or quantity of testing, or both), and follow-up in different healthcare settings (e.g. primary care versus hospital). Cluster-RCTs were eligible.

Types of participants

Males and females of any age with histologically proven adenocarcinoma of the colon or rectum, staged as T1-4N0-2M0 (Edge 2010), treated surgically with curative intent (plus or minus adjuvant treatment).

Types of interventions

Follow-up visits with health professionals, including symptom enquiry, clinical examination, and procedures and investigations (including but not limited to colonoscopy, blood tests, faecal analysis, and radiological examinations).

Types of outcome measures

Primary outcomes

1. Overall survival (measured from the time of randomisation in the study).

Secondary outcomes

1. Colorectal cancer-specific survival (measured from the time of randomisation in the study).
2. Relapse-free survival (measured from the time of randomisation in the study).
3. Salvage surgery (surgery performed with curative intent for relapse of CRC).
4. Interval recurrences (relapse of CRC detected between follow-up visits).
5. Quality of life (using trial-specific instruments, including but not limited to FACT (Functional Assessment of Cancer Therapy), EORTC QLQ-C30 (European Organization for Research and Treatment of Cancer Quality of Life), and EORTC-CRC) (Cella 1993; Sprangers 1993; Whistance 2009).
6. Harms, including but not limited to psychological harms, investigation-related complications, and waste of resources.
7. Costs of surveillance (including investigations).

Search methods for identification of studies

Electronic searches

We searched the following electronic databases with no language restriction.

- The Cochrane Central Register of Controlled Trials (CENTRAL; 2016, Issue 3) in the Cochrane Library using the strategy in Appendix 1;
- MEDLINE Ovid (from 1950 to 20 May 2016) using the strategy in Appendix 2;
- Embase Ovid (from 1974 to 20 May 2016) using the strategy in Appendix 3;
- CINAHL EBSCO (Cumulative Index to Nursing and Allied Health Literature; from 1981 to May 2016) using the strategy in Appendix 4; and
- Science Citation Index (from 1900 to May 2016) using the strategy in Appendix 5

For the Review first published in the Cochrane Library 2002 issue 1, we also searched the electronic database CANCERLIT, which stopped existing in 2003.

Searching other resources

Trial registries

We searched the following trials registries:

- US National Institutes of Health Ongoing Trials Register ClinicalTrials.gov (www.clinicaltrials.gov; searched May 2016); and
- World Health Organization International Clinical Trials Registry Platform (apps.who.int/trialsearch; searched May 2016).

Handsearching

We searched the following journals and conference proceedings:

- American Society for Clinical Oncology (ASCO) (1995 to 2010);
- European Society for Therapeutic and Radiation Oncology (1990, 1993, 2000 to 2010, 2012); and
- International Journal of Radiation Oncology Biology Physics: proceedings of the American Society for Radiation Oncology (ASTRO) (2011 to 2015).

We searched reference lists of published articles and previous systematic reviews and made personal contact with experts. We identified non-English and unpublished studies.

Grey literature

We searched OpenGrey (www.opengrey.eu) (May 20th, 2016).

Data collection and analysis

Selection of studies

Two authors (BEH and GMJ) checked the titles and abstracts identified from the databases. The authors obtained the full text of all studies of possible relevance for independent assessment, decided which trials met the inclusion criteria, and graded their methodological quality. Discussion between the review authors resolved any disagreement. We contacted authors of primary studies for clarification where necessary. The reported outcomes were not used as criteria for including studies. We included studies irrespective of their publication status. We documented the selection process using Covidence and presented the details of the search in a PRISMA diagram. Reasons for exclusion are presented in the 'Characteristics of excluded studies' table. We collated multiple reports of the same study so that each study, rather than the report, was the unit of interest in the review, and we identified the primary source.

Data extraction and management

Two review authors (BEH and GMJ) independently performed data extraction; we contacted the authors of trials to provide missing data where possible. We entered data into a previously piloted data form then into Covidence. One author (BEH) entered data into RevMan, which a second author (AS) checked. We resolved any disagreements by discussion. We extracted the following data when available:

1. number of participants;
2. the age and status of the participants;
3. inclusion and exclusion criteria;
4. setting;
5. treatment regimen;
6. follow-up details; and
7. survival, adverse events, and quality of life indices.

We collected data that were sufficient to populate a table of characteristics of included studies. For studies where only a subset of the participants recruited were eligible for inclusion, we included them if they reported data for that subgroup separately. Where a study had more than one study arm, such as FACS 2014, we combined those intervention study arms that met the inclusion criteria and compared them with the control arm; this ensured we did not double-count data. When subgroup analysis was performed for FACS 2014, we combined the two arms that had carcinoembryonic antigen (CEA) measured and the two arms in which computerised tomography (CT) was used. We compared the magnitude and direction of effects reported by studies with how they were presented in the review.

In order to report time-to-event data, we used the RevMan, RevMan 2014, calculator and a spreadsheet developed by Matthew Sydes, Tierney 2007, to derive observed (O) and log-rank expected events (E) (O-E) and variance. Tierney 2007 presents 11 methods for calculating a hazard ratio (HR) or associated statistics, or both, from published time-to-event-analyses into a practical, less statistical guide. The methods we used to do so were dependent on the available information in the texts, and we report them as follows.

Reports presenting HRs and 95% confidence intervals allowed application of method three in Tierney 2007 and were available for analysis as follows:

1. overall survival (Augestad 2013; FACS 2014; Strand 2011; Treasure 2014; Wang 2009) (please note that for Wang 2009, we used the RevMan calculator to derive the HR, because this agreed with the P value given in the text);
2. colorectal cancer-specific survival (Augestad 2013; FACS 2014; Rodríguez-Moranta 2006); and
3. relapse-free survival (FACS 2014; Mäkelä 1995; Ohlsson 1995; Pietra 1998; Schoemaker 1998; Treasure 2014; Rodríguez-Moranta 2006; Secco 2002; Strand 2011).

In reports with a P value, events in each arm, and where the randomisation ratio was 1:1, we used method seven in Tierney 2007 to derive O-E and variance. Such studies contributed to the fol-

lowing analyses:

1. overall survival (GILDA 1998; Kjeldsen 1997; Ohlsson 1995; Schoemaker 1998; Rodríguez-Moranta 2006) (for Rodríguez-Moranta 2006, we used the RevMan calculator to derive the HR, because the statistic presented in the text was adjusted for confounding (this was also the approach that the systematic review Pita-Fernández 2014 used for this study));
2. colorectal cancer-specific survival (Kjeldsen 1997; Ohlsson 1995); and
3. relapse-free survival (GILDA 1998; Kjeldsen 1997; Rodríguez-Moranta 2006).

In reports where we extracted data from the survival curve, assuming constant censoring, we used method 10 in Tierney 2007 for two trials contributing to the outcome of overall survival (Mäkelä 1995; Pietra 1998).

Assessment of risk of bias in included studies

Two study authors (BEH and GMJ) constructed and presented 'Risk of bias' tables using the Cochrane 'Risk of bias' tool, resolving any disagreements by discussion. We evaluated the following domains:

- random sequence generation;
- allocation concealment;
- blinding of participants and personnel;
- blinding of outcome assessment;
- incomplete follow-up (exclusions, attrition);
- selective reporting; and
- other bias, including but not limited to early stopping, inadequate duration of follow-up, or baseline imbalances.

We graded domains as at low risk of bias, high risk of bias, or unclear risk of bias (using the criteria in the *Cochrane Handbook for Systematic Reviews of Interventions*, Table 8.5d) (see Appendix 6), with our reasons and supporting evidence detailed in the tables (Higgins 2011). We summarised the risk of bias for each of the key study outcomes (overall survival, disease-specific survival, relapse-free survival, salvage surgery, interval recurrences, and complications of colonoscopy).

Measures of treatment effect

Where possible, we conducted time-to-event analyses for overall survival, colorectal cancer-specific survival, and relapse-free survival. We expressed the results as hazard ratios (HR) with 95% confidence intervals (CI) when the relevant information was available in the text or could be derived. Where necessary, we derived the HR using the RevMan calculator and calculated associated statistics using an Excel spreadsheet developed by Matthew Sydes (Cancer Division of the Medical Research Council (MRC) Clinical Trials Unit) in collaboration with the Meta-analysis Group of the MRC Clinical Trials Unit, London (Tierney 2007). We reported relative risks (RR) and 95% CIs for dichotomous outcomes and weighted mean differences (WMD) and 95% CI for continuous outcomes. We interpreted a statistically non-significant result (P value larger than 0.05) as a finding of uncertainty unless the confidence intervals were sufficiently narrow to rule out a potentially important magnitude of effect. We defined confidence intervals between 0.75 and 1.25 as excluding clinically meaningful benefits or harms.

Unit of analysis issues

All of the RCTs were parallel in design, with participants being the unit of randomisation; therefore, we had no unit of analysis issues.

Dealing with missing data

We contacted the authors of Secco 2002 to request the raw data. We were informed on 20 April 2015 that because of personnel changes, the authors were unable to retrieve the data, which meant we were not able to report overall survival data for this study. Because they plotted two curves for each study arm, it was not possible to extract data for use in time-to-event analysis. We were in contact with the study authors for GILDA 1998 on 18 February 2016, who kindly provided us with unpublished data, which we were able to include for the outcomes 'Interval recurrences' and 'Salvage surgery' (Fossati 2015). Therefore, all analyses were by intention-to-treat.

Assessment of heterogeneity

We assessed heterogeneity both visually and statistically using the Chi^2 test of heterogeneity, Altman 1992; Walker 1988, and I^2 statistic, Higgins 2002; Higgins 2011. The criterion for identification of heterogeneity is a P value less than 0.10 for the Chi^2 test (acknowledging the limitations of this process) and an I^2 statistic value of greater than 50%. Where we identified significant heterogeneity, we first checked the data to ensure it was not due to error, explored the potential causes of it, and made a cautious attempt to explain the heterogeneity.

Assessment of reporting biases

We assessed the potential impact of reporting biases by the use of a funnel plot for the three outcomes that included data from 10 or more studies (overall survival, relapse-free survival, and salvage surgery). Including 15 studies allowed us to visually assess whether small-study effects were present or not.

Data synthesis

We calculated a weighted treatment effect (using a random-effects model) across trials using the Cochrane statistical package, RevMan version 5.3.5 (RevMan 2014). Where O-E and variance were available, we used a log-rank approach and a fixed-effect

model to synthesise data. We summated data where we judged the participants, interventions, and outcomes to be sufficiently similar to ensure a clinically meaningful answer.

Subgroup analysis and investigation of heterogeneity

We used subgroup analyses to investigate possible differences in participant outcomes according to study variables that we believed could be effect modifiers. These included the use of CEA, CT, and PET/CT (positron emission tomography-computed tomography) in the intensive follow-up strategy when compared with no use or less frequent use (twice at most) in the control arm, and setting for follow-up (general practitioner (GP)- or nurse-led follow-up compared with hospital follow-up and "dose" of follow-up, i.e. studies that compared the use of more visits and tests with fewer visits and tests). These subgroup analyses may help identify which investigations are useful in follow-up for colorectal cancer and allow us to give specific guidance to clinicians. We used a formal statistical test to compare subgroups.

Sensitivity analysis

We performed prespecified sensitivity analyses to test the strength of our conclusions by excluding studies judged to be at high risk of bias for the particular outcome concerned (Kjeldsen 1997; Mäkelä 1995; Pietra 1998; Schoemaker 1998; Wang 2009), and by study age (excluding those studies that completed accrual by 1996) (Kjeldsen 1997; Mäkelä 1995; Ohlsson 1995; Pietra 1998; Schoemaker 1998; Treasure 2014).

We performed post hoc sensitivity analysis in response to reviewer suggestion by excluding one study (Ohlsson 1995), where the intensity of follow-up in the intensive arm was comparable with the intensity of follow-up in the control arm of other studies.

'Summary of findings' tables

We evaluated the quality of evidence using the GRADE approach for the following outcomes (Schünemann 2009):

- overall survival;
- colorectal cancer-specific survival;
- relapse-free survival;
- salvage surgery;
- interval recurrences; and
- harms associated with surveillance.

We used GRADEpro to present the quality of evidence for the aforementioned outcomes in 'Summary of findings' tables. We could downgrade the quality of the evidence by one (serious concern) or two levels (very serious concern) for the following reasons: risk of bias, inconsistency (unexplained heterogeneity, inconsistency of results), indirectness (indirect population, intervention, control, outcomes), imprecision (wide confidence intervals, single trial), and publication bias. We could also downgrade the quality by one level due to a large summary effect.

RESULTS

Description of studies

Results of the search

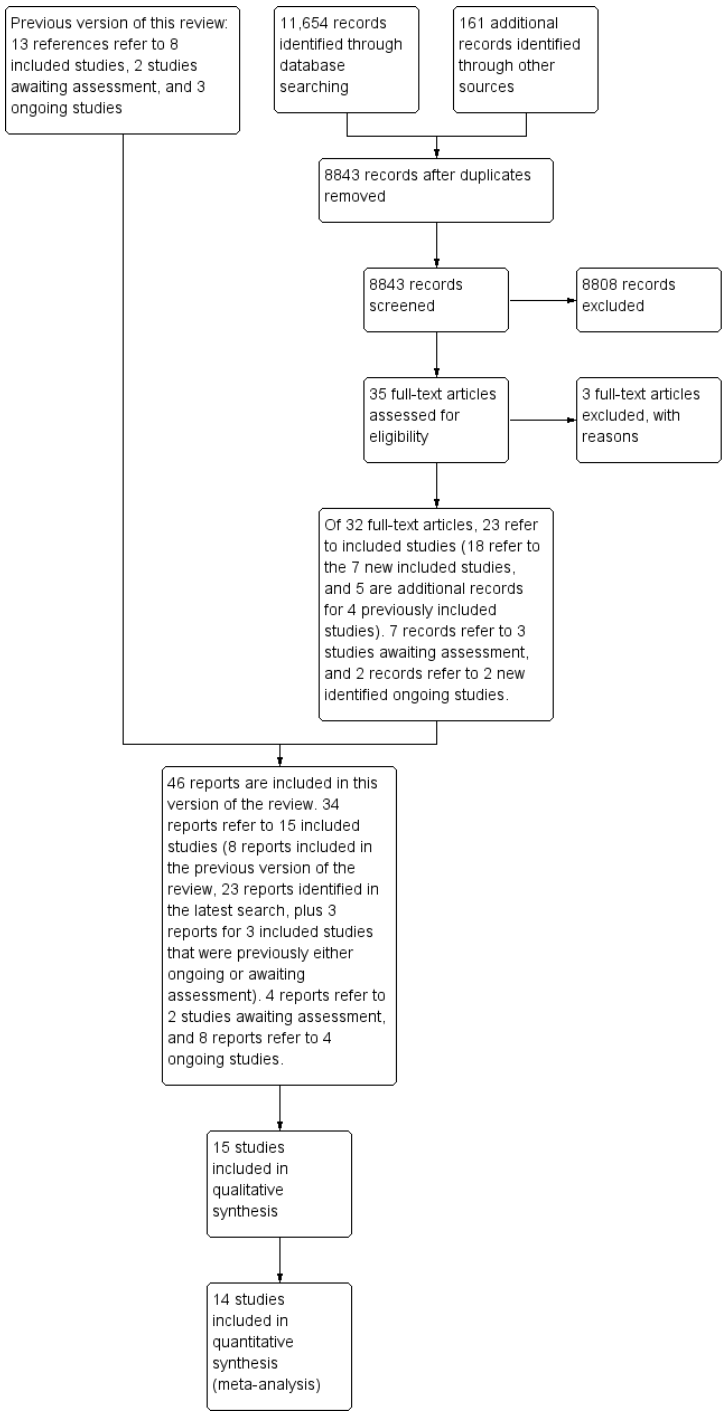
For this update, we screened 8843 references, of which, we assessed 35 references in full. We identified seven new studies for inclusion, comprised of 18 identified references from the last performed search (Augestad 2013; FACS 2014; GILDA 1998; Sobhani 2008; Strand 2011; Treasure 2014; Wang 2009), and we identified five additional references for four previously included studies (Kjeldsen 1997; Mäkelä 1995; Ohlsson 1995; Wattchow 2006). In this update, we moved three references referring to three studies that were previously either ongoing or awaiting assessment to included studies. Furthermore, we identified two additional references for an ongoing study presented in the last published version of this review (COLOFOL), and finally, we identified two new ongoing studies (NCT00995202; NCT01628211).

In summary, this updated version of the review now includes a total of 49 references.

- Thirty-four references refer to 15 included studies (Augestad 2013; FACS 2014; GILDA 1998; Kjeldsen 1997; Mäkelä 1995; Ohlsson 1995; Pietra 1998; Rodríguez-Moranta 2006; Schoemaker 1998; Secco 2002; Sobhani 2008; Strand 2011; Treasure 2014; Wang 2009; Wattchow 2006).
- Three references refer to three excluded studies (Barillari 1996; Kronborg 1981; Sano 2004).
- Seven references refer to three studies awaiting assessment (Jefford 2013; NCT00199654; Verberne 2015).
- Five references refer to three ongoing studies (COLOFOL; NCT00995202; NCT01628211).

There was considerable variation in the follow-up strategies employed by the 15 studies; both the frequency of, the setting for, and the investigations that were performed during follow-up visits were different in each study (see the 'Characteristics of included studies' tables and Figure 1).

Figure 1. Study flow diagram



Included studies

Similarities and differences between the included studies

Nine of the 15 studies were multicentred (Augestad 2013; FACS 2014; GILDA 1998; Kjeldsen 1997; Mäkelä 1995; Ohlsson 1995; Pietra 1998; Treasure 2014; Wattchow 2006). We identified no cluster-RCTs (randomised controlled trials).

Participants

Seven of the 15 studies included Dukes' stage A, B, and C colon and rectal cancer (Augestad 2013; FACS 2014; Kjeldsen 1997; Mäkelä 1995; Ohlsson 1995; Rodríguez-Moranta 2006; Wang 2009). Two studies excluded Dukes' A participants (GILDA 1998; Pietra 1998), two studies excluded participants with rectal cancer (Pietra 1998; Wattchow 2006), and one study included only rectal cancer participants (Strand 2011).

Interventions

The studies can be grouped into the following areas of assessment:

1. "dose" of follow-up: more visits and tests versus fewer visits and tests (Kjeldsen 1997; Mäkelä 1995; Pietra 1998; Rodríguez-Moranta 2006; Secco 2002; Treasure 2014; Wang 2009);
2. formal follow-up versus minimal/no follow-up (FACS 2014; Ohlsson 1995; Schoemaker 1998; Secco 2002);
3. more liver imaging versus less liver imaging (FACS 2014; GILDA 1998; Rodríguez-Moranta 2006; Schoemaker 1998);
4. carcinoembryonic antigen (CEA) versus no CEA (FACS 2014; Kjeldsen 1997; Ohlsson 1995; Treasure 2014); and
5. setting for follow-up (where frequency of visits and tests were identical in both arms): general practitioner (GP)-led follow-up, Augestad 2013; Wattchow 2006, or nurse-led follow-up, Strand 2011, compared with surgeon-led follow-up.

The included studies did not assess the quality of histopathology.

Outcomes

1. Twelve RCTs reported overall survival (Augestad 2013; FACS 2014; GILDA 1998; Kjeldsen 1997; Mäkelä 1995; Ohlsson 1995; Pietra 1998; Rodríguez-Moranta 2006; Schoemaker 1998; Strand 2011; Treasure 2014; Wang 2009).
2. Seven RCTs reported colorectal cancer-specific survival (measured from the time of randomisation in the study) (Augestad 2013; FACS 2014; GILDA 1998; Kjeldsen 1997; Ohlsson 1995; Rodríguez-Moranta 2006; Wang 2009).
3. Fourteen RCTs reported relapse-free survival (measured from the time of randomisation in the study) (Augestad 2013; FACS 2014; GILDA 1998; Kjeldsen 1997; Mäkelä 1995; Ohlsson 1995; Pietra 1998; Rodríguez-Moranta 2006; Schoemaker 1998; Secco 2002; Sobhani 2008; Strand 2011; Treasure 2014; Wang 2009).
4. Thirteen RCTs reported salvage surgery (surgery performed with curative intent for relapse of colorectal cancer (CRC)) (Augestad 2013; FACS 2014; GILDA 1998; Kjeldsen 1997; Mäkelä 1995; Ohlsson 1995; Pietra 1998; Rodríguez-Moranta 2006; Schoemaker 1998; Secco 2002; Sobhani 2008; Treasure 2014; Wang 2009).
5. Eight RCTs reported interval recurrences (relapse of CRC detected between follow-up visits or symptomatic recurrences) (FACS 2014; Kjeldsen 1997; Mäkelä 1995; Secco 2002; Sobhani 2008; Wang 2009; Wattchow 2006; Augestad 2013).
6. Four RCTs assessed quality of life (Augestad 2013; GILDA 1998; Kjeldsen 1997; Wattchow 2006). Augestad 2013 used various validated scales (EORTC QLQ C-30 (European Organization for Research and Treatment of Cancer quality of life questionnaire), EQ-5D (EuroQol five dimensions questionnaire)) (Dolan 1997; Sprangers 1993). Kjeldsen 1997 used the Nottingham Health Profile (Anderson 1996; Hunt 1980). Wattchow 2006 reported the short form (SF)-12 Physical and Mental Health component (Ware 1995) and Hospital Anxiety Depression Scale (Zigmond 1983). GILDA 1998 used the 12-item short version (SF-12) of SF-36 (Apolone 1998; Gandek 1998), which was validated in the Italian population, and the Psychological General Well-Being (PGWB) Index (Dupuy 1984).
7. Schoemaker 1998 and Wang 2009 reported harms.
8. Four studies evaluated costs of surveillance (including investigations) (Augestad 2013; Secco 2002; Rodríguez-Moranta 2006; Strand 2011). Rodríguez-Moranta 2006 and Augestad 2013 performed cost=minimisation analyses.

Study accrual dates spanned over three decades. Kjeldsen 1997; Ohlsson 1995; Mäkelä 1995; Pietra 1998; Schoemaker 1998; Secco 2002; and Treasure 2014 accrued in the 1980s and 1990s. GILDA 1998; Rodríguez-Moranta 2006; Sobhani 2008; and Wang 2009 accrued participants in the 1990s and early 2000s, and Augestad 2013 and FACS 2014 accrued participants from 2003 to 2011.

The variety of investigations used across the studies may affect the applicability of results. For example, Augestad 2013; Kjeldsen 1997; and Treasure 2014 did not use CT scanning.

Excluded studies

For this update, we applied the current recommendations from

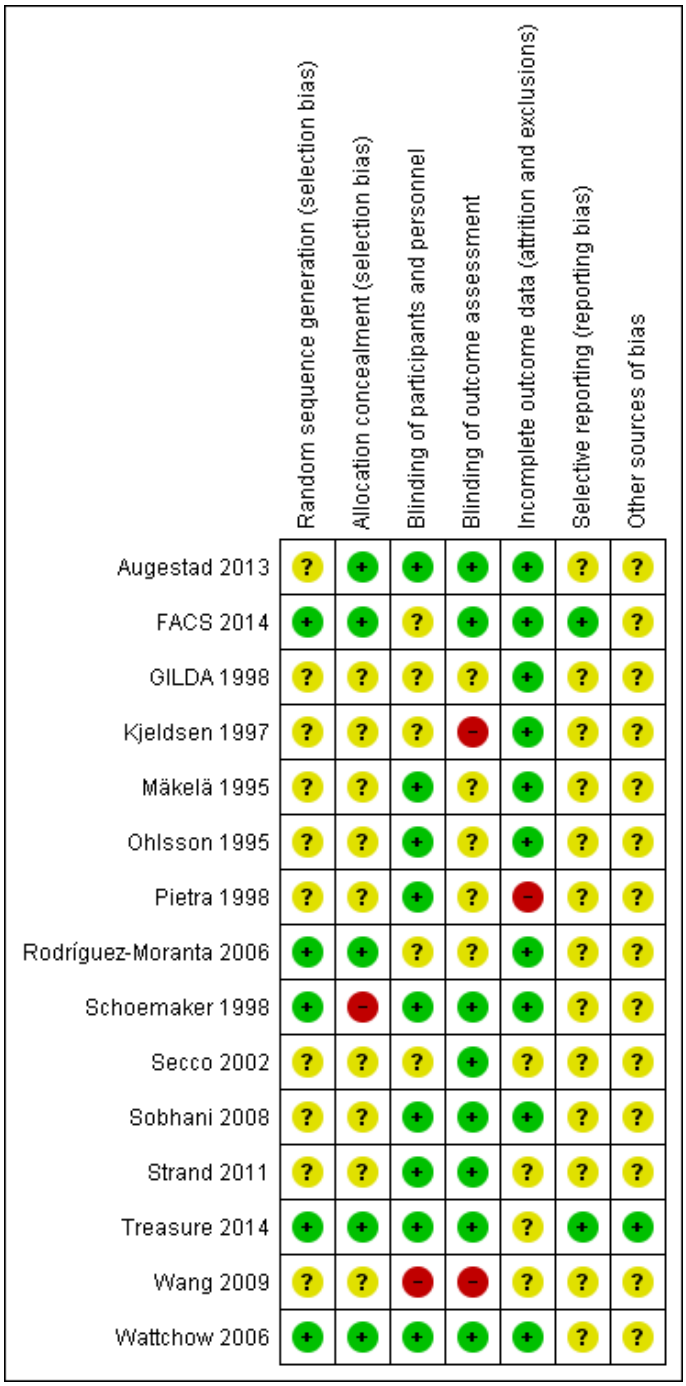
the *Cochrane Handbook for Systematic Reviews of Interventions* with respect to excluded studies and only classified studies as excluded if they were those that one might reasonably expect could have been eligible for inclusion.

We excluded three studies identified in the latest search (see the 'Characteristics of excluded studies' tables). We were unable to include data from Barillari 1996; of 607 participants enrolled, 212 were randomised, but data for randomised participants were not reported separately. Kronborg 1981 was a prospective, partly randomised trial, but we could not extract data relating to randomised participants from the paper. Sano 2004 was not eligible because participants had not had colorectal cancer.

Risk of bias in included studies

There was complete concordance between authors regarding the evaluation of trial methodology (Figure 2).

Figure 2. 'Risk of bias' summary: review authors' judgements about each 'Risk of bias' item for each included study



Allocation

Although all of the studies were reported to be randomised, only two explicitly reported that they concealed the allocation of participants to study groups(Mäkelä 1995; Wattchow 2006). We found that none of the studies were at high risk of bias with respect to allocation; we judged them all to be at low or uncertain risk of bias.

Blinding

Participant or clinician blinding was not possible. We judged several studies to be at high risk of bias for blinding of participants (Kjeldsen 1997; Mäkelä 1995; Schoemaker 1998; Wang 2009). One study used independent radiologists who were blinded to study group allocation to assess CT scans (Schoemaker 1998). We judged three studies to be at high risk of bias for blinding of outcome assessors (Kjeldsen 1997; Pietra 1998; Wang 2009).

Incomplete outcome data

Seven studies ensured that they obtained outcome data from more than 80% of the participants. Wattchow 2006 obtained outcome data for 77% of the participants. All studies conducted intention-to-treat analyses. We judged one study to be at high risk of bias for incomplete outcome data (Pietra 1998). Two studies examined compliance with the follow-up regimen (Rodríguez-Moranta 2006; Schoemaker 1998), but no study fully assessed contamination.

Selective reporting

We did not have access to the protocols for most studies (Augestad 2013; Kjeldsen 1997; Mäkelä 1995; Ohlsson 1995; Pietra 1998; Rodríguez-Moranta 2006; Schoemaker 1998; Secco 2002; Strand 2011; Wang 2009; Wattchow 2006), so we judged them to be at unclear risk of bias. With more information available we judged both FACS 2014 and Treasure 2014 to be at low risk of bias for this domain (see Characteristics of included studies).

Other potential sources of bias

We did not find other sources of bias (including inadequate follow-up duration and baseline imbalances on study populations). One study was stopped early (Treasure 2014), but we did not feel this was likely to introduce bias.

Overall survival

We judged four studies contributing data to this outcome to be at high risk of bias because blinding was not mentioned (Kjeldsen 1997; Mäkelä 1995; Secco 2002; Wang 2009). We did not feel this represented any risk of bias for this objective outcome. Although 10% (98/955) of the events contributing to this outcome came from studies deemed at high risk of bias for allocation concealment, Mäkelä 1995; Schoemaker 1998, and attrition bias, Pietra 1998, we did not feel that we needed to downgrade for risk of bias.

Colorectal cancer-specific survival

We judged Pietra 1998 to be at high risk of bias because the authors potentially excluded 15% of the participants randomised without explaining to which study arm they belonged. Kjeldsen 1997 and Pietra 1998 did not mention blinding and probably did not ensure it (this represented 192/468 (41%) of the events contributing to this outcome); because this could cause ascertainment bias for cause of death, we downgraded evidence quality for colorectal cancer-specific survival for risk of bias.

Relapse-free survival

For this outcome, Kjeldsen 1997 and Secco 2002 (which contributed 340/1340 (25%) of the events) did not mention blinding. We judged both Mäkelä 1995 and Schoemaker 1998 (which contributed 163/1340 (12%) of the events) to be at high risk of bias because of their allocation concealment. We judged this outcome to be at high risk of bias because of the lack of allocation concealment, so downgraded for risk of bias. A total of 340/1340 (25%) of the events contributing to this outcome came from studies at high risk of bias because of a lack of blinding.

Salvage surgery

We judged Schoemaker 1998 to be at high risk of bias for allocation concealment. We also deemed three other studies contributing to this outcome to be at high risk of bias: Kjeldsen 1997 and Wang 2009 for lack of blinding of outcome assessment, Pietra 1998 for incomplete outcome reporting, and Wang 2009 further did not blind participants and personnel. We did not downgrade for risk of bias despite these limitations, because Schoemaker 1998 contributed only 11/526 (0.05%) of the events for this outcome, lack of blinding was unlikely to have affected the outcome reporting of salvage surgery, and the incomplete outcome reporting in Pietra 1998 was related to other outcomes.

Interval recurrences

We did not downgrade this outcome for risk of bias. We judged FACS 2014 to be at low risk of bias for all domains. We judged both Secco 2002 and Wang 2009 to be at high risk of bias for the domain of blinding, but this was because blinding was not mentioned in Wang 2009, and in Secco 2002, there were prespecified follow-up schedules. We did not downgrade this outcome for risk of bias (Risk of bias in included studies and Summary of findings for the main comparison).

Effects of interventions

See: **Summary of findings for the main comparison** Follow-up strategies for patients treated for non-metastatic colorectal cancer

Intensive versus less intensive/minimalist follow-up (dose/frequency)

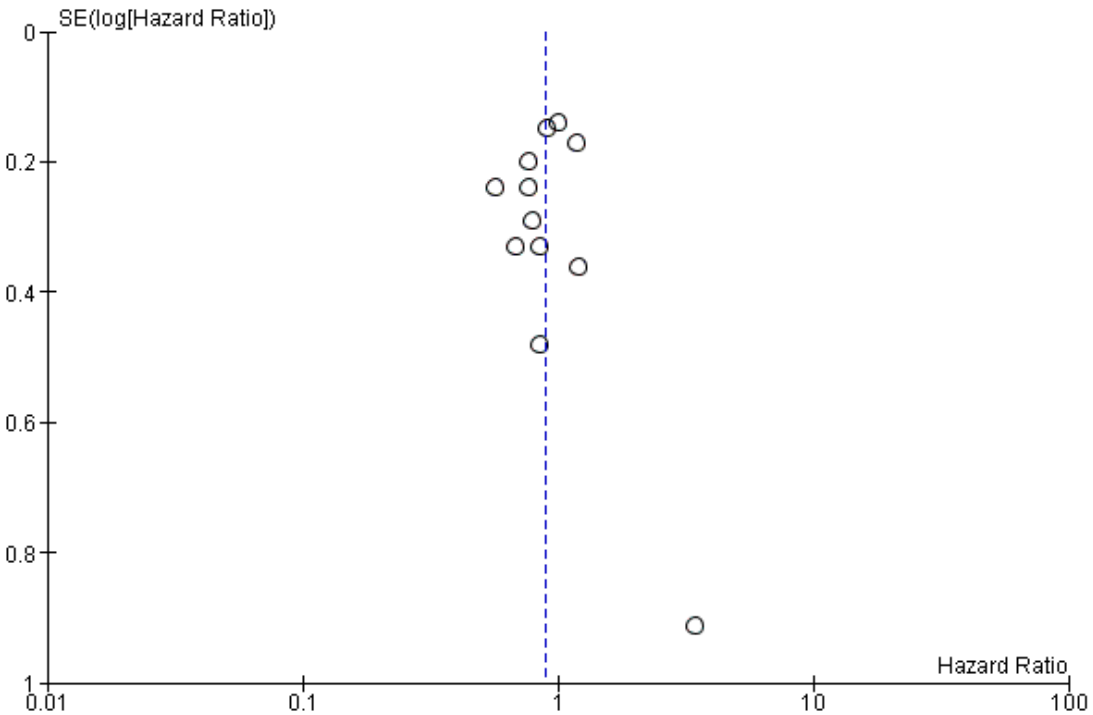
1. Primary outcome

1.1 Overall survival

We studied 1098 deaths in 4786 participants enrolled in 12 studies (duration of follow-up was greater than 48 months for 99% of the participants, which is adequate to record an event in most/all participants) (Augestad 2013; FACS 2014; GILDA 1998; Kjeldsen 1997; Mäkelä 1995; Ohlsson 1995; Pietra 1998; Rodríguez-Moranta 2006; Schoemaker 1998; Strand 2011; Treasure 2014; Wang 2009) (hazard ratio (HR) 0.90, 95% confidence interval (CI) 0.78 to 1.02 - the confidence intervals exclude both appreciable benefit or harm). Heterogeneity is not likely to be important: I^2 = 4%; P = 0.41 (Analysis 1.1).

The funnel plot did not show evidence of small-study effect (Figure 3).

Figure 3. Funnel plot of comparison: 1 Intensive follow-up versus minimal follow-up, outcome: 1.1 Overall survival



In absolute terms, the average effect of intensive follow-up on survival was associated with 17 fewer deaths per 1000 patients, but the true effect may have been between 50 fewer to 18 more deaths per 1000 patients. The GRADE assessment of evidence quality for this outcome was high.

Subgroup analyses

1. Studies comparing different health professionals did not find evidence of a clinically meaningful effect on overall survival (HR 0.90, 95% CI 0.78 to 1.02). Formal testing for subgroup differences was negative (Chi^2 = 0.44; P = 0.51; I^2 = 0%) when those studies that used different settings with general practitioner- or nurse-led follow-up, Augestad 2013; Strand 2011, were compared with those set in hospitals, FACS 2014; GILDA 1998; Kjeldsen 1997; Mäkelä 1995; Ohlsson 1995; Pietra 1998; Rodríguez-Moranta 2006; Schoemaker 1998; Treasure 2014; Wang 2009.
2. Studies that compared more visits and tests with fewer visits and tests, Kjeldsen 1997; Mäkelä 1995; Pietra 1998; Rodríguez-Moranta 2006; Treasure 2014; Wang 2009, and studies that compared follow-up with minimal or no follow-up, FACS 2014; Ohlsson 1995; Schoemaker 1998, did not find evidence of a clinically meaningful effect on overall survival (HR 0.86, 95% CI 0.74 to 1.00). Formal testing for subgroup differences was negative (Chi^2 = 0.24; P = 0.62; I^2 = 0%).
3. Studies using CEA in the intensive follow-up regimen did not find evidence of a statistically significant effect on overall survival (HR 0.99, 95% CI 0.81 to 1.21). We found little evidence of heterogeneity: I^2 = 0%; P = 0.40. Testing revealed no evidence of differences between the subgroups (Chi^2 = 0.52; P = 0.47).
4. Studies using CT in the intensive follow-up regimen did not find evidence of a statistically significant effect on overall survival (HR 0.92, 95% CI 0.77 to 1.09). Formal statistical testing revealed no evidence of differences between the subgroups (Chi^2 = 0.36; P = 0.55).
5. Studies using frequent CT scans in the intervention arm versus the use of two or fewer CT scans in the control arm did not find evidence of a statistically significant effect on overall survival (HR 0.87, 95% CI 0.73 to 1.05; Analysis 1.9). Formal statistical testing revealed no evidence of differences between the subgroups (Chi^2 = 0.99; P = 0.32).

Sensitivity analyses

Our findings for the outcome of overall survival were robust to sensitivity analyses.

1. Excluding studies at high risk of bias for this outcome (Schoemaker 1998; Pietra 1998), we found no statistical evidence of a survival advantage for the comparison of intensive versus less intensive follow-up (HR 0.95, 95% CI 0.83 to 1.10). We found no heterogeneity: I^2 = 0%; P = 0.68.
2. Excluding seven studies on the basis of study age (GILDA 1998; Kjeldsen 1997; Mäkelä 1995; Ohlsson 1995; Pietra 1998; Schoemaker 1998; Treasure 2014), we found no statistical evidence of a survival advantage for the comparison of intensive versus less intensive follow-up (HR 0.97, 95% CI 0.74 to 1.28, using random-effects). We found little evidence of heterogeneity: I^2 = 15%; P = 0.32.
3. Excluding one study where the intensity of the follow-up in the "intensive" arm was similar to that in the control arm of other studies (Ohlsson 1995), we found no evidence of a clinically meaningful effect on overall survival (HR 0.90, 95% CI 0.79 to 1.04). We found no evidence of heterogeneity: I^2 = 7%; P = 0.38.

2. Secondary outcomes

2.1 Colorectal cancer-specific survival (CRC-SS)

We were able to report on 432 colorectal cancer deaths in 2769 participants enrolled in seven studies (99.6% had a median follow-up of greater than 48 months): we found no evidence of effect of intensive versus less intensive follow-up on CRC-SS (HR 0.93, 95% CI 0.78 to 1.12). We found no evidence of heterogeneity: I^2 = 0%; P = 0.45 (Augestad 2013; FACS 2014; GILDA 1998; Kjeldsen 1997; Ohlsson 1995; Rodríguez-Moranta 2006; Wang 2009)(Analysis 1.2).

In absolute terms, the average effect of intensive follow-up on CRC-SS was 10 fewer CRC-SS deaths per 1000 patients, but the true effect could lie between 30 fewer to 16 more per 1000 patients. The GRADE assessment of evidence quality for this outcome was moderate.

Sensitivity analyses

Our findings for the outcome of CRC-SS were robust to the following sensitivity analyses.

1. Excluding studies at high risk of bias for this outcome (Kjeldsen 1997; Wang 2009), we found no statistical evidence of an effect on CRC-SS (HR 0.91, 95% CI 0.68 to 1.21) and no evidence of heterogeneity: I^2 = 24%; P = 0.26.
2. We found no statistical evidence of study age having an effect (HR 0.91, 95% CI 0.67 to 1.23) and no evidence of heterogeneity: I^2 = 30%; P = 0.22 (excluding GILDA 1998; Kjeldsen 1997; and Ohlsson 1995).

2.2 Relapse-free survival (R-FS)

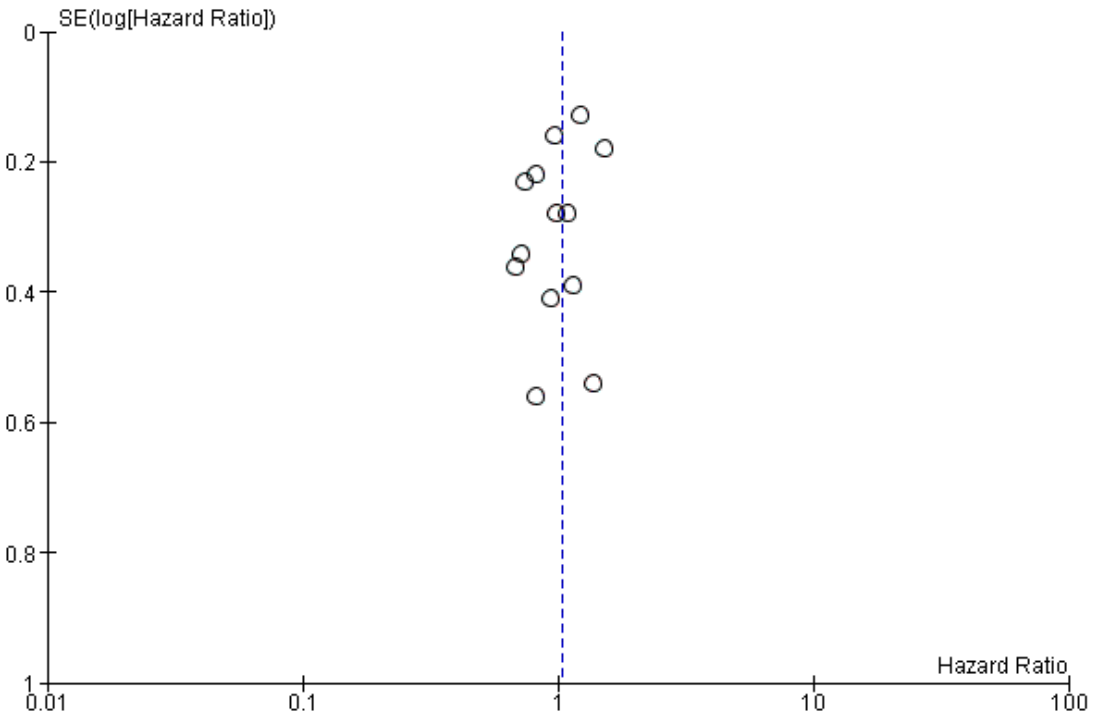
We were able to report on 1416 relapses in 5253 participants enrolled in 14 studies (with a median follow-up of greater than 48 months for 97.9% of participants studied) (Augestad 2013; FACS 2014; GILDA 1998; Kjeldsen 1997; Mäkelä 1995; Ohlsson 1995; Pietra 1998; Rodríguez-Moranta 2006; Schoemaker 1998;

Secco 2002; Sobhani 2008; Strand 2011; Treasure 2014; Wang 2009)(Analysis 1.3).

We found a small increase in R-FS (HR 1.03, 95% CI 0.90 to 1.18); the CIs however excluded both clinically meaningful benefits and harms. We found no evidence of heterogeneity: I^2 = 5%; P = 0.39.

The funnel plot did not show evidence of small-study effect (see Figure 4).

Figure 4. Funnel plot of comparison: 1 Intensive follow-up versus minimal follow-up, outcome: 1.3 Relapse-free survival



The average effect of intensive follow-up on relapse-free survival was seven more relapses per 1000 patients, but the true effect could lie between 24 fewer and 41 more per 1000 patients. The GRADE assessment of evidence quality for this outcome was moderate.

Sensitivity analyses

Our findings for the outcome of relapse-free survival were robust to the following sensitivity analyses.

1. Excluding studies at high risk of bias for this outcome (Kjeldsen 1997; Pietra 1998; Schoemaker 1998; Wang 2009), we found no statistical evidence of an effect (HR 1.14, 95% CI 0.97 to 1.33) and no clear evidence of heterogeneity: I^2 = 0%; P = 0.48.

2. With regard to study age (excluding GILDA 1998; Kjeldsen 1997; Mäkelä 1995; Ohlsson 1995; Pietra 1998; Schoemaker 1998; and Treasure 2014), we found no statistical evidence of an effect (HR 1.05, 95% CI 0.79 to 1.40) and little evidence of heterogeneity: I^2 = 30%; P = 0.21.

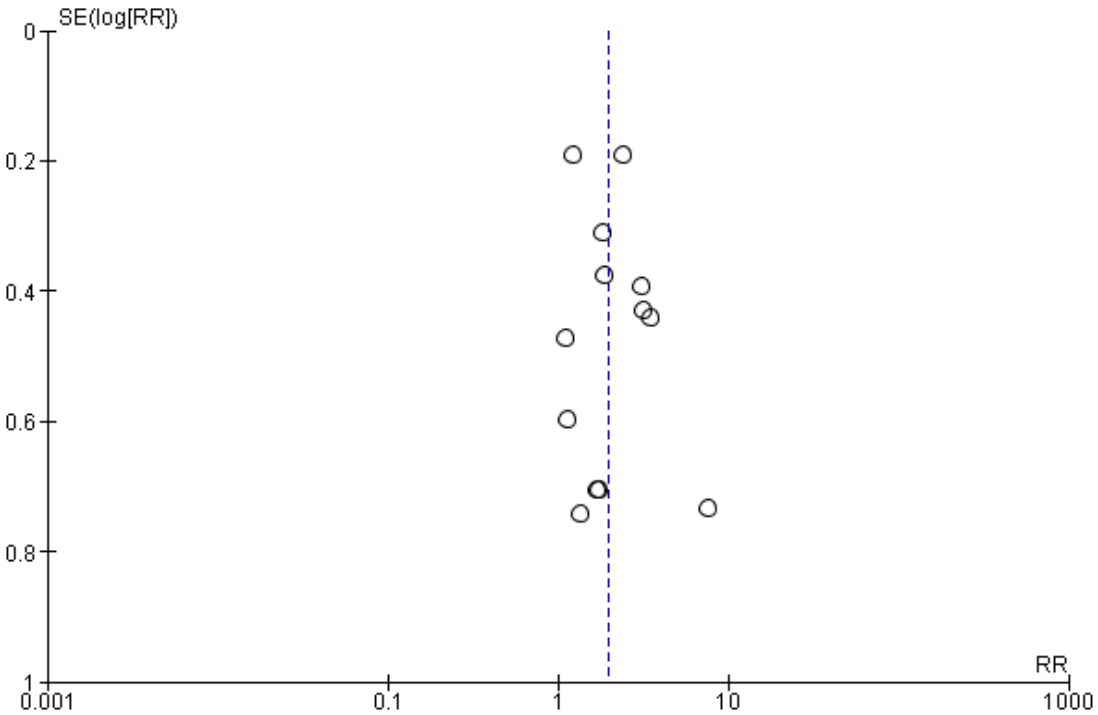
2.3 Salvage surgery

We were able to report on 457 episodes of salvage surgery in 5157 participants enrolled in 13 studies (with a follow-up duration of greater than 48 months in 90.6% of participants studied) (Augestad 2013; FACS 2014; GILDA 1998; Kjeldsen 1997; Mäkelä 1995; Ohlsson 1995; Pietra 1998; Rodríguez-Moranta 2006; Schoemaker 1998; Secco 2002; Sobhani 2008; Treasure 2014; Wang 2009)(Analysis 1.4).

We found an appreciable increase in episodes of salvage surgery with intensive follow-up for CRC (RR 1.98, 95% CI 1.53 to 2.56). The CIs included a range of clinically significant increases in salvage surgery. We found some non-significant evidence of heterogeneity: I^2 = 31%; P = 0.14.

The funnel plot did not show evidence of small-study effect (see Figure 5).

Figure 5. Funnel plot of comparison: 1 Intensive follow-up versus minimal follow-up, outcome: 1.4 Salvage surgery



In absolute terms, the effect of intensive follow-up on salvage surgery was 60 more episodes of salvage surgery per 1000 patients, but the true effect could lie between 33 to 96 more episodes per 1000 patients. The GRADE assessment of evidence quality for this outcome was high.

Sensitivity analyses

Our findings for the outcome of salvage surgery were robust to the following sensitivity analyses.

1. Excluding studies at high risk of bias for this outcome (Pietra 1998; Schoemaker 1998; Wang 2009) (RR 2.03, 95% CI 1.53 to 2.69), we found some non-significant heterogeneity: I^2 = 33%; P = 0.15.

2. With regard to study age (excluding GILDA 1998; Kjeldsen 1997; Mäkelä 1995; Ohlsson 1995; Pietra 1998; Schoemaker 1998; and Treasure 2014) (RR 2.04, 95% CI 1.35 to 3.09), we found that when we excluded the older studies, there was less heterogeneity: I^2 = 25%; P = 0.25.

2.4 Interval (symptomatic) recurrences

We found 376 interval recurrences reported in 3933 participants enrolled in seven studies (with a median follow-up duration of greater than 48 months for 100% of participants studied) (FACS 2014; Kjeldsen 1997; Mäkelä 1995; Secco 2002; Sobhani 2008; Wang 2009)(Analysis 1.5). There was an appreciable decrease in the number of interval recurrences (RR 0.59, 95% CI 0.41 to 0.86). The CIs included a range of clinically significant decreases in interval recurrences (Analysis 1.5). We detected heterogeneity: I^2= 66%; P = 0.007.

Intensive follow-up was associated with fewer interval recurrences (52 fewer per 1000 patients); the true effect is between 18 and 75 fewer per 1000 patients. The GRADE assessment of quality of evidence was moderate.

Sensitivity analyses

Our findings were robust to the following sensitivity analyses.

1. Excluding studies at high risk of bias for this outcome (Kjeldsen 1997; Wang 2009) (RR 0.61, 95% CI 0.37 to 1.02), we found evidence of heterogeneity: I^2= 75%; P = 0.003.

2. With regard to study age (excluding GILDA 1998; Kjeldsen 1997; and Mäkelä 1995) (RR 0.42, 95% CI 0.32 to 0.56), we found no evidence of heterogeneity: I^2 = 0%; P = 0.81.

2.5 Quality of life

Augestad 2013 reported no significant effect on quality of life main outcome measures. For EORTC QLQ-C30, significant effects in favour of GP-led follow-up were reported for role functioning (P = 0.02), emotional functioning (P = 0.01), and pain (P = 0.01). No significant differences in global health status were reported (Augestad 2013). Kjeldsen 1997 reported the influence of different follow-up strategies on quality of life for 350/597 Danish participants. They reported a small increase in quality of life (P < 0.05), as measured by the Nottingham Health Profile, associated with more frequent follow-up visits compared with virtually no follow-up. Wattchow 2006 assessed depression and anxiety, quality of life, and participant satisfaction in a cohort of participants randomised to follow-up of their colon cancer in different settings (see Included studies). They found that the study participants remained in the normal range for depression and anxiety with no difference between the two groups at either 12 or 24 months. Study participants (in each arm) had reduced physical quality of life at baseline, which improved as the study progressed, but there were no significant differences between the two groups. There were no differences between the two groups on the participant satisfaction scale, and both groups reported high levels of satisfaction with their care. There were no clinically significant differences among the three main quality of life scales (SF-12 mental component, SF-12 physical component, and PGWB Index) between the two study arms found in GILDA 1998.

2.6 Harms

Two studies reported adverse events associated with follow-up (Schoemaker 1998; Wang 2009). They reported three perforations and four gastrointestinal haemorrhages (requiring transfusion) from a total of 2112 (0.4%) colonoscopies. Intensive follow-up was not associated with increased risk of perforation (RR 2.08, 95% CI 0.11 to 40.17) (Wang 2009).

2.7 Costs of surveillance

Secco 2002 provided risk-adapted follow-up based on prognostic factors prospectively identified, and the authors commented that risk-adapted follow-up reduced costs for those with a better prognosis. Rodríguez-Moranta 2006 demonstrated that although the cost of intensive follow-up was higher, when resectability of recurrences was considered, the cost per resectable recurrence was lower in the intensively followed group. Augestad 2013 found the cost per participant for 24 months' follow-up was £9889 for surgeon-led follow-up and £8233 for GP-led follow-up (P < 0.001) (figures from text).

DISCUSSION

The results of our review suggest that there is no overall survival benefit for intensifying the follow-up of participants after curative surgery for colorectal cancer. The analyses did not show a significant difference in the incidence of recurrence between the participants in the intensively followed groups and the control groups. However, significantly more surgical procedures for recurrence were performed in the experimental arms of the trials. Recurrences in the more intensively followed groups may have been detected earlier allowing for effective salvage treatments, but this did not lead to better overall survival.

Each trial follow-up strategy combined a number of different components, including frequency of visits, type of clinical assessment, types and frequency of tests, and the setting in which follow-up was conducted. No trial compared the addition of one specific intervention, and the feasibility of comparing strategies with a variety of components and varying complexity becomes problematic.

The use of liver imaging does not appear to be associated with improved survival. A specific variation across the studies was the intensity of follow-up. For example, the follow-up intensity in the intensively followed group in Ohlsson 1995 was similar to the intensity of follow-up in the control groups of other studies in the review (Mäkelä 1995; Pietra 1998; Schoemaker 1998). Therefore, it was not possible to extract from these data a precise indication of the optimal combinations of frequency, type, and setting for follow-up investigations for these participants. Our findings were robust to sensitivity analysis when excluding Ohlsson 1995.

Most recurrences (about 90%) occur within the initial 36 months after initial therapy for colorectal cancer (Ryuk 2014), so to detect recurrences, follow-up duration should be at least 36 months for colorectal cancer. Patients with rectal cancer should have longer follow-up because liver and lung recurrences may be delayed. The use of adjuvant chemotherapy may further delay recurrence (Sadahiro 2003). For the outcomes included in this study, median follow-up duration was greater than 48 months for more than 90% of the participants studied.

This updated version of the review has substantially altered our conclusions: where we previously reported that "there was evidence that an overall survival benefit at five years exists for patients undergoing more intensive follow up" (Jeffery 2007), this update does not confirm these findings. In this update of the review, we were able to include data on an additional 3322 participants enrolled in seven more studies (Augestad 2013; FACS 2014; GILDA 1998; Sobhani 2008; Strand 2011; Treasure 2014; Wang 2009). Two of these studies had previously been unpublished (GILDA 1998; Treasure 2014), despite being completed. One of the six newly included studies was at high risk of bias; two assessed the intervention of setting (general practitioner (GP) or nurse versus surgeon). One study evaluated the utility of PET/CT (positron emission tomography-computed tomography); and one other large study, which contributed 212/1098 (19%) of the events for the survival outcome, was set in hospitals that used modern investigations and multidisciplinary teams to manage recurrences (FACS 2014).

Summary of main results

1. Overall survival: there was a lack of statistical evidence of effect for the use of intensive versus less intensive follow-up after curative treatment (hazard ratio (HR) 0.90, 95% confidence interval (CI) 0.78 to 1.02; Analysis 1.1).
2. Colorectal cancer-specific survival: there was a lack of statistical evidence of effect for the use of intensive versus less intensive follow-up after curative treatment) (HR 0.93, 95% CI 0.78 to 1.12; Analysis 1.2).
3. Relapse-free survival: there was a lack of statistical evidence of effect with intensive versus less intensive follow-up after curative treatment (HR 1.03, 95% CI 0.90 to 1.18; Analysis 1.3).
4. The use of salvage surgery was increased with intensive follow-up after curative treatment for colorectal cancer (risk ratio (RR) 1.98, 95% CI 1.53 to 2.56; Analysis 1.4).
5. Interval (symptomatic) recurrences were reduced with intensive follow-up after curative treatment for colorectal cancer (RR 0.59, 95% CI 0.41 to 0.86; Analysis 1.5).

Quality of life

Augestad 2013 reported no significant effect on quality of life main outcome measures. For EORTC QLQ-C30 (European Organization for Research and Treatment of Cancer-Quality of Life), significant effects in favour of GP-led follow-up were reported for pain, role functioning, and emotional functioning. No differences in global health status were reported (Augestad 2013), with intensive follow-up compared with less intense follow-up. In Kjeldsen 1997, a small increase in quality of life associated with more frequent follow-up visits compared with virtually no follow-up. Wattchow 2006 found that the study participants remained in the normal range for depression and anxiety with no difference between the two groups at either 12 or 24 months.

Harms and costs of surveillance (including investigations)

The trials reported three bowel perforations and four gastrointestinal haemorrhages (requiring transfusion) from a total of 2112 (0.4%) colonoscopies (Schoemaker 1998; Wang 2009).
Secco 2002 reported that risk-adapted follow-up reduced costs for those with a better prognosis. Rodríguez-Moranta 2006 demonstrated that although the cost of intensive follow-up was higher, when resectability of recurrences was considered, the cost per resectable recurrence was lower in the intensively followed group. Augestad 2013 found the cost per participant for 24 months' follow-up was £9889 for surgeon-led follow-up and £8233 for GP-led follow-up ($P < 0.001$) (figures from text).

Overall completeness and applicability of evidence

The evidence we report is directly relevant to the study question. The studies have been accrued over a prolonged time period, during which time there have been significant changes in cancer staging procedures, operative techniques (surgical metastasectomy or ablation of liver metastases), postoperative care, adjuvant therapies, and the investigations available to detect recurrence. Systemic adjuvant therapies and effective palliative chemotherapy drugs are now widely utilised, with significant prolongation of survival rates. All of these factors question the validity of applying the results of early studies to the modern surgical and oncological setting. A sensitivity analysis excluding studies that commenced accrual before 1996 did not however reveal any effect of study age on

the outcome of overall survival (HR 0.97, 95% CI 0.74 to 1.28). The lack of benefit from intensive follow-up persisted despite the inclusion of modern studies that used modern surgical techniques for resection of liver metastases, FACS 2014, and more sensitive investigations to detect recurrence, such as PET scanning, Sobhani 2008.

The argument for intensive follow-up has been up to now based on observational data that reports improved survival after hepatic metastasectomy. We did not find an improvement in overall survival with the use of liver imaging (HR 0.92, 95% CI 0.81 to 1.21).

The identified studies are sufficient to address our objectives; future updates of this review are likely to address those outcomes that we can not yet address in detail, when subsequent publications detail quality of life (GILDA 1998; NCT00995202), costs (NCT00995202; NCT00199654), and the effects of the addition of PET scanning into follow-up for these participants (NCT00199654; NCT00624260 (reference under Sobhani 2008)).

We were unable to obtain data for overall survival for the Secco 2002 study (despite contacting the authors) because of personnel changes at the institution concerned. The COLOFOL study is still ongoing.

The studies included the relevant participant population. The interventions assessed were very inclusive, addressing a variety of "doses" or intensities of follow-up, ranging from no follow-up to intense follow-up, and evaluated multiple investigations including the use of CT and PET scanning.

No study addressed any potential psychological harms, anxiety, or distress that may be associated with follow-up after treatment for colorectal cancer.

The studies included in this review did not well report the potential harms (physical, psychological) and costs of follow-up strategies. Two studies reported harms related to colonoscopy (Schoemaker 1998; Wang 2009). The rate of perforation (3/2112 or 0.14%) was consistent with other published series (Araghizadeh 2001; Bowles 2004). None of the study reports included specific details of any harms (mortality or morbidity) resulting from investigation or treatment of recurrences. These outcomes should be available in order to fully assess any net benefit or harm of follow-up.

Some researchers have investigated the psychological effects of follow-up (Augestad 2013; Kjeldsen 1997; Stiggelbout 1997; Wattchow 2006). These studies have reported mixed effects on quality of life measures, but no study has found a deterioration in quality of life. Some form of follow-up appeared superior to virtually no follow-up in terms of quality of life (Kjeldsen 1997). Different settings for follow-up (GP- versus surgeon-led) did not appear to affect anxiety or depression; both groups had a high and similar level of participant satisfaction (Wattchow 2006). Ongoing studies will address the effects of intensifying follow-up on quality of life in this population (NCT00995202).

Further research is required into the value that participants place on follow-up after their curative surgery. Any survival benefit (or lack of benefit) of follow-up would have to be considered along with the views of participants so that follow-up programmes are accessible, acceptable, and address all participants' needs and concerns.

Little useful data are available from the studies in this review on the cost-effectiveness of follow-up in this group of patients treated for non-metastatic colorectal cancer. It appears that GP-led follow-up is cheaper than surgeon-led follow-up (Augestad 2013); risk-adapted follow-up is cheaper for those with better prognosis disease (Secco 2002), and although the cost of intensive follow-up is higher, it makes the cost per resectable recurrence lower (Rodríguez-Moranta 2006). Without a better understanding of which of the specific follow-up interventions is responsible for the improvement in outcomes, it is not possible to even speculate on the potential cost-effectiveness of any one approach. Investigators have previously tried to project the costs of a single intervention such as serum carcinoembryonic antigen (CEA) testing (Audisio 1996; Moertel 1993), and the reported costs have appeared prohibitively large. In contrast, an incremental cost-effectiveness analysis based on five randomised controlled trials has reported costs of intensive follow-up, which appear acceptable in the setting of the National Health Service in the United Kingdom (Renehan 2004), although the authors do acknowledge a number of limitations of their study. An ongoing study will address the issue of costs so that the relative cost-effectiveness of follow-up can be viewed from an economic perspective as well as a clinical one (NCT00995202).

Quality of the evidence

The findings of this review allow robust conclusions, with minimal heterogeneity and low risk of publication bias (based on the use of funnel plots).

Overall survival (OS)

For the outcome of OS, we studied 4786 participants in 12 studies. We did not downgrade for risk of bias. We did not downgrade for inconsistency (I^2 = 4%; P = 0.41) or indirectness: 1312/4733 (27%) participants contributing to this outcome were accrued after 2003 (so used modern investigations and surgical salvage techniques). We did not downgrade for imprecision (there were > 300 events (955), and the confidence intervals excluded clinically meaningful benefit or harm). We did not downgrade for publication or other bias. The GRADE assessment of evidence quality for this outcome was high.

Colorectal cancer-specific survival (CR-SS)

For the outcome CRC-SS, we studied 2769 participants in seven studies. We downgraded because 41% of the events were from

studies deemed at high risk of bias for incomplete outcome reporting or lack of blinding. We did not downgrade for inconsistency (I^2 = 0%; P = 0.45). We did not downgrade for indirectness, despite the long time period over which studies accrued participants; 34% of the participants included in this outcome were enrolled in studies that accrued in the 2000s. We did not downgrade for imprecision (there were more than 300 events (432), and the confidence intervals excluded clinically meaningful benefit or harm). We did not downgrade for publication or other bias. The GRADE assessment of evidence quality for this outcome was moderate.

Relapse-free survival (R-FS)

For the outcome of R-FS, we studied 5253 participants in 14 studies. We downgraded for risk of bias, because 34% of the events were from studies deemed at high risk of bias from lack of blinding. We did not downgrade for inconsistency (I^2 = 5%; P = 0.39), indirectness, or imprecision (there were more than 300 events (1309), and the confidence intervals excluded clinically meaningful benefit or harm). We did not downgrade for publication or other bias. The GRADE assessment of evidence quality for this outcome was moderate.

Salvage surgery (SS)

For the outcome of SS, we studied 5157 participants in 13 studies. We did not downgrade for risk of bias. We did not downgrade for indirectness, because 1988/5854 (33%) of participants contributing to this outcome were enrolled in studies done in the 2000s. There was evidence of precision (with > 300 events (457) and optimum information size met). While there was some evidence of non-significant inconsistency (I^2 = 31%; P = 0.14), we felt the variation in the intensity of follow-up across the studies and the prolonged time period for accrual (which allowed for varied surgical techniques and degrees of surgical aggression) explained this because of clinical heterogeneity between the studies, so we did not downgrade for this. Prespecified sensitivity analysis based on study age supported this (I^2 = 25%; P = 0.25). We did not downgrade for publication bias. The GRADE assessment of evidence quality for this outcome was high.

Interval (asymptomatic) recurrences (IR)

For the outcome of IR, we studied 3933 participants in seven studies. We downgraded for risk of bias, because 58% of the events came from studies at high risk of bias because of lack of blinding. We did not downgrade for inconsistency (I^2= 66%; P = 0.007), because prespecified sensitivity analysis based on study age explained this heterogeneity (I^2 = 0%; P = 0.81). We did not downgrade for indirectness or imprecision (there were > 300 events (364)) or publication bias. The GRADE assessment of quality of evidence was moderate.

Adverse effects

For the outcome of adverse effects, we studied 651 participants in two studies. These were rare events (seven events reported in 2112 colonoscopies in two studies). We did downgrade for imprecision, so the GRADE assessment of quality of evidence was moderate.

Potential biases in the review process

The studies included in this review did not report the potential harms of both investigations and salvage treatments. It is possible that investigators excluded participants from study enrolment whom they felt to be at high risk of recurrence; if this did occur, it may have diluted the effect of follow-up strategies on survival.

Agreements and disagreements with other studies or reviews

We found five published systematic reviews relevant to our question. Our findings differ from those reported by Pita-Fernández 2014; Renehan 2002; and Tjandra 2007: they found that intensive follow-up for participants treated with curative intent for colorectal cancer improved survival. The other two systematic reviews did not report a quantitative meta-analysis (Augestad 2014; Baca 2011) (see Table 1).

Pita-Fernández 2014 included 11 randomised controlled trials (RCTs), with 4055 participants. Their search was limited to four bibliographic databases (search date: June 2014). They reported that overall survival increased with intensive follow-up (HR 0.75, 95% CI 0.66 to 0.86) (see Table 1). They included data relating to the GILDA 1998 study from an earlier paper with 14 months' median follow-up, published while the trial was still accruing participants (Grossman 2004). In addition, they included data from Secco 2002 and Wattchow 2006, which we were unable to extract. It appears that the data for overall survival reported for Secco 2002 may have been derived from the actuarial survival at five years (reported as percentages). We had contact with the study authors for Secco 2002 (detailed in the Methods section), who informed us that they could not give us any more information than was in the text due to personnel changes. As stated above, it appears the review authors incorrectly derived the overall survival data from the actuarial survival percentages reported.

Augestad 2014 searched PubMed and reference lists of published studies (no search date given). They reported no quantitative meta-analysis for the five studies they included. They commented that recent data did not report a survival advantage (FACS 2014), and they suggested that the potential survival benefits of surveillance should be weighed against possible negative effects.

Baca 2011 searched PubMed and reference lists (search date: June 2000 to June 2010). They included Secco 2002; GILDA 1998 (Grossman data), and Wang 2009, but did not present a quantitative meta-analysis. They included both randomised (n = 5) and

non-randomised (n = 11) studies. They concluded that recent literature is inconclusive with respect to the benefit of surveillance for colorectal cancer after curative treatment. The difference in our findings was likely to relate to our more recent search and inclusion of 10 more RCTs.

Renehan 2002 included five RCTs and 1342 participants. Their search was systematic (search date: April 2001). They reported improved survival with intensive follow-up in this setting. Again, the difference in our findings was likely to relate to our more recent search and inclusion of 10 more RCTs.

Tjandra 2007 included eight RCTs and 2923 participants. Their search was systematic (search date: June 2007). They reported improved survival with intensive follow-up. Once again, the difference in our findings was likely to relate to our more recent search and inclusion of seven more RCTs.

The previous iteration of this Cochrane Review also found improved survival in this setting. This updated version of the review (which now includes data from 15 studies, including 5403 participants) contradicts the previously reported effects of surveillance on survival for participants treated with curative intent for colorectal cancer.

AUTHORS' CONCLUSIONS

Implications for practice

The results of this review suggest that intensifying clinical follow-up for participants with colorectal cancer (CRC) after curative treatment does not improve survival outcomes. The exact details of the optimal follow-up regimen still need clarification, but limiting follow-up intensity does not seem to be disadvantageous.

Implications for research

Clinicians are encouraged to enrol their participants in any ongoing trials in this field. Such trials may reflect advances in imaging and surgical technique and the use of adjuvant therapies. All investigators are encouraged to explicitly document any harms relating to follow-up and subsequent interventions.

Separate research programmes should explore patient needs and concerns relating to the value of follow-up, incorporating other study designs, using qualitative as well as quantitative methods.

ACKNOWLEDGEMENTS

Susan Bidwell, who both designed and performed the original search.

New Zealand Health Technology Assessment and Cochrane Colorectal Cancer assisted with retrieval of articles.

REFERENCES

References to studies included in this review

Augestad 2013 ***{published data only}***

* 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**(4):e002391. [3026260; NCT00572143]

Augestad KM, Vonen B, Aspevik R, Nestvold T, Ringberg U, Johnsen R, et al. 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 Services Research* 2008;**8**:137. [3026261]

FACS 2014 ***{published data only}***

Corkhill A, Primrose JN, Mant D, The FACS Study Group. Follow up after colorectal cancer (the FACS trial). *British Journal of Cancer* 2004;**91**(Suppl 1):S78. [3026263]

Mant D, Perera R, Gray A, Rose P, Fuller A, Corkhill A, et al. Effect of 3-5 years of scheduled CEA and CT follow-up to detect recurrence of colorectal cancer: FACS randomized controlled trial. *Journal of Clinical Oncology* 2013;**31**:Suppl; abstr 3500. [3026264]

NCT00560365. Follow-up study of patients who have undergone surgery for stage I, stage II, or stage III colorectal cancer [A randomised controlled trial to assess the cost–effectiveness of intensive versus no scheduled follow–up in patients who have undergone resection for colorectal cancer with curative intent (FACS – Follow–up After Colorectal Surgery)]. clinicaltrials.gov/ct2/show/NCT00560365?term=NCT00560365&rank=1 (first received November 14th 2012). [3026265]

Primrose JN, Fuller A, Rose P, Perera-Salazar R, Mellor J, Corkhill A, et al. Follow-up after colorectal cancer surgery: preliminary observational findings from the UK FACS trial. *Journal of Clinical Oncology, 2011 ASCO Annual Meeting Abstracts Part 1* 2011;**29, No 15**(Suppl (May 20)):3521. [3026267]

Primrose JN, Perera R, Gray A. Effect of 3 to 5 years of scheduled CEA and CT follow-up to detect recurrence of colorectal cancer: the FACS randomized clinical trial. *Diseases of the Colon and Rectum* 2014;**57**:e421–2. [3026269]

* Primrose JN, Perera R, Gray A, Rose P, Fuller A, Corkhill A, et al. 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**(3):

263–70. [3026266]
Primrose JN, Perera R, Gray 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. http://jama.jamanetwork.com/ (First received January 14th 2014). [3026268; DOI: 10.1001/jama.2013.285718]
Reibetanz J, Germer CT. Influence of structured CEA determination and CT-scan to detect recurrence of colorectal cancer: results of the randomized clinical FACS study. *Der Chirurg; Zeitschrift für alle Gebiete der operativen Medizen* 2014;**85**(7):642. [3026270]

GILDA 1998 ***{unpublished data only}***
Fossati R, Andreoni A, Di Costanzo B, Fossati D, Johnson G, Liberati L, et al. Preliminary results of the Italian study "GILDA" [Preliminari dello studio I risultati preliminari dello studio italiano GILDA]. Bologna 2005. [3026272]
Grossmann EM, Johnson FE, Virgo KS, Longo WE, Fossati R. Follow-up of colorectal cancer patients after resection with curative intent-the GILDA trial. *Surgical Oncology* 2004;**13**(2-3):119–24. [3026273]
Johnson FE, Virgo KS, Grossmann EM, Longo WE, Fossati R. Colorectal cancer patient follow-up following surgery with curative intent: the GILDA trial. *Journal of Clinical Oncology, 2004 ASCO Annual Meeting Proceedings (Post-meeting Edition)* 2004;**22**:14S. [3026274]
* Rosati G, Ambrosini G, Barni S, Andreoni B, Corradini G, Luchena G, et al. A randomized trial of intensive versus minimal surveillance of patients with resected Dukes B2-C colorectal carcinoma. *Annals of Oncology* 2016;**27**(2):274–80. [3026275; DOI: 10.1093/annonc/mdv541]

Kjeldsen 1997 ***{published data only}***
* Kjeldsen BJ, Kronborg O, Fenger C, Jørgensen OD. A prospective randomised study of follow-up after radical surgery for colorectal cancer. *British Journal of Surgery* 1997;**84**(5):666–9. [3026278]
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. *International Journal of Colorectal Disease* 1997;**12**(6):329–34. [3026279]
Kjeldsen BJ, Thorsen H, Whalley D, Kronborg O. Influence of follow-up on health-related quality of life after radical surgery for colorectal cancer. *Scandinavian Journal of Colorectal Diseases* 1999;**34**(5):509–15. [3026277]

Mäkelä 1995 ***{published data only}***
Mäkelä J, Laitinen S, Kairaluoma MI. Early results of follow-up after radical resection for colorectal cancer. Preliminary results of a prospective randomized trial. *Surgical Oncology* 1992;**1**(2):157–61. [3026281]
* Mäkelä JT, Laitinen SO, Kairaluoma MI. Five-year follow-up after radical surgery for colorectal cancer. Results of a prospective randomized trial. *Archives of Surgery* 1995;**130**(10):1062–7. [3026282]

Ohlsson 1995 ***{published data only}***
Kronborg O, Fenger C, Deichgräber E, Hansen L. Follow-up after radical surgery for colorectal cancer. Design of a randomized study. *Scandinavian Journal of Gastroenterology* 1988;**23**(Suppl 149):159–62. [3026284]
* Ohlsson B, Breland U, Ekberg H, Graffner H, Tranberg KG. Follow-up after curative surgery for colorectal carcinoma. Randomized comparison with no follow-up. *Diseases of the Colon and Rectum* 1995;**38**(6):619–26. [3026285]

Pietra 1998 ***{published data only}***
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. *Diseases of the Colon and Rectum* 1998;**41**(9):1127–33. [3026287]

Rodríguez-Moranta 2006 ***{published data only}***
Rodríguez-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. *Journal of Clinical Oncology* 2006;**24**(3):386–93. [3026289]

Schoemaker 1998 ***{published data only}***
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. *Gasteroenterology* 1998;**114**(1):7–14. [3026291]

Secco 2002 ***{published data only}***
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. *European Journal of Surgical Oncology* 2002;**28**(4):418–23. [3026293]

Sobhani 2008 ***{published data only}***
NCT00624260. Impact of PET Scan on the curative strategy of colo-rectal cancers: a randomized study (ITEP). clinicaltrials.gov/ct2/show/NCT00624260?term=NCT00624260&rank=1 (First received November 14th 2012). [3026295]
* Sobhani I, Tiret E, Labtahi R, Aparicio T, Itti E, Montravers F, et al. Early detection of recurrence by 18FDG-PET in the follow-up of patients with colorectal cancer. *British Journal of Cancer* 2008;**98**(5):875–80. [3026296]

Strand 2011 ***{published data only}***
Strand E, Nygren I, Bergkvist L, Smedh K. Nurse or surgeon follow-up after rectal cancer: a randomized trial. *Colorectal Disease* 2011;**13**(9):999–1003. [3026298]

Treasure 2014 ***{published data only}***
Lennon T, Houghton J, Northover J, on behalf of CRC/NIH CEA Trial Working Party. Post-operative CEA monitoring and second-look surgery in colorectal cancer: trial results. British Journal of Cancer. Proceedings of Ninth Annual Scientific Meeting of the British Oncological

Association. New Jersey: Macmillan Publishers Ltd, 1994: 16. [3026300]
Northover J, Houghton J, Lennon T. CEA to detect recurrence of colon cancer. *JAMA* 1994;**272**(1):31. [3026301]
* 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**(5):e004385. [3026302; DOI: 10.1.1136/bmjopen-2013-004385]

Wang 2009 *{published data only}*
Wang T, Cui Y, Huang WS, Deng YH, Gong W, Li CJ, et al. The role of postoperative colonoscopic surveillance after radical surgery for colorectal cancer: a prospective, randomized clinical study. *Gastrointestinal Endoscopy* 2009; **69**(3 Pt 2):609–15. [3026304]

Wattchow 2006 *{published data only}*
Wattchow DA, Weller D, Pilotto L, Estermann A. Randomized controlled trial (RCT) comparing hospital and general practice based follow-up of patients with colon cancer. *Australian and New Zealand Journal of Surgery* 2002; **72 Suppl**:A22. [3026306]
* Wattchow DA, Weller DP, Esterman A, Pilotto LS, McGorm K, Hammett Z, et al. General practice vs surgical-based follow-up for patients with colon cancer: randomised controlled trial. *British Journal of Cancer* 2006;**94**(8): 1116–21. [3026307]

References to studies excluded from this review

Barillari 1996 *{published data only}*
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. *Diseases of the Colon and Rectum* 1996;**39**(4):388–93. [3026309]

Kronborg 1981 *{published data only}*
Kronborg O, Hage E, Deichgraeber E. A prospective, partly randomized study of the effectiveness of repeated examination of the colon after polypectomy and radical surgery for cancer. *Scandinavian Journal of Gasteroenterology* 1981;**16**(7):879–84. [3026311]

Sano 2004 *{published data only}*
Sano Y, Fujii T, Oda Y, Matsuda T, Kozu T, Kudo S, et al. A multicenter randomized controlled trial designed to evaluate follow-up surveillance strategies for colorectal cancer: The Japan Polyp Study. *Digestive Endoscopy* 2004; **16**(4):376–8. [3026313]

References to studies awaiting assessment

Jefford 2013 *{published data only}*
Butow PN, Sze M, Dugal-Beri P, Mikhail M, Eisenbruch M, Jefford M, et al. From inside the bubble: migrants' perceptions of communication with the cancer team. *Psychooncology* 2010;**19**(Suppl 2):S193, P-39. [3026330]
* Jefford M, Aranda S, Gough K, Lotfi-Jam K, Butow P, Krishnasamy M, et al. Evaluating a nurse-led survivorship care package (SurvivorCare) for bowel cancer survivors: study protocol for a randomized controlled trial. *Trials* 2013;**14**(1):260. [3026329]
Jefford M, Gough K, Drosdowsky A, Russell L, Aranda S, Butow P, et al. A randomized controlled trial of a nurse-led supportive care package (SurvivorCare) for survivors of colorectal cancer. *The Oncologist* 2016;**21**(8):1014–23. [4456933]

NCT00199654 *{unpublished data only}*
NCT00199654. Follow-up care with or without CEA assessments in patients who have undergone surgery for stage II or stage III colorectal cancer [Follow–up of fully resected stage II or III colorectal cancer. Phase III multicentric prospective randomised study]. clinicaltrials.gov/ct2/show/NCT00995202?term=NCT00995202&rank=1 (First received November 14th 2012). [3026325]

Verberne 2015 *{unpublished data only}*
Verberne C, Doornbos PM, Grossmann I, De Bock GH, Wiggers T. Intensified follow-up in colorectal cancer patients using frequent carcino-embryonic antigen (CEA) measurements and CEA-triggered imaging. *European Journal of Cancer* 2013;**49 Suppl 2**:1188–96. [3026317]
* 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. *European Journal of Surgical Oncology* 2015;**41**(9): 1188–96. [3026318]
Zhan Z, van den Heuvel ER, Doornbos PM, Burger H, Verberne CJ, Wiggers T, et al. Strengths and weaknesses of a stepped wedge cluster randomized design: its application in a colorectal cancer follow-up study. *Journal of Clinical Epidemiology* 2014;**67**(4):454–61. [3026319]

References to ongoing studies

COLOFOL *{unpublished data only}*
* NCT00225641. Assessment of frequency of surveillance after curative resection in patients with stage II and III colorectal cancer [COLOFOL – a pragmatic study to assess the frequency of surveillance tests after curative resection in patients with stage II and III colorectal cancer – a randomised multicentre trial]. clinicaltrials.gov/ct2/show/NCT00225641?term=NCT00225641&rank=1 (First received November 14th 2012). [3026321]
Wille-Jørgensen P, Laurberg S, Påhlman L, Carriquiry L, Lundqvist N, Smedh K, et al. An interim analysis of recruitment to the COLOFOL trial. *Colorectal Disease* 2009;**11**(7):756–8. [3026322]
Wille-Jørgensen P, Laurberg S, Sørensen HT, Pa
° hlman L, on behalf of the COLOFOL study group. COLOFOL study protocol A pragmatic study to assess

the frequency of surveillance tests after curative resection in patients with stage II and III colorectal cancer - a randomised multicentre trial. COLFOL protocol April 30th 2008, issue 7.5. [3026323]

NCT00995202 ***{unpublished data only}***

NCT00995202. Follow-up care with or without CEA assessments in patients who have undergone surgery for stage II or stage III colorectal cancer [Follow–up of fully resected stage II or III colorectal cancer. Phase III multicentric prospective randomised study]. clinicaltrials.gov/ct2/show/NCT00995202?term=NCT00995202&rank=1 (First received November 14th 2012). [3026327]

NCT01628211 ***{unpublished data only}***

NCT01628211. Second look laparoscopy in colorectal cancer (HIPEC) [Randomized phase 2 study comparing second look laparoscopy to standard follow up in patients with no radiologic evidence of disease at 6 months after complete resection of colorectal mucinous carcinoma]. clinicaltrials.gov/ct2/show/NCT01628211?term=NCT01628211&rank=1 (First received August 20th 2014). [3026315]

Additional references

Alberts 2005

Alberts SR, Horvath WL, Sternfeld WC, Goldberg RM, Mahoney MR, Dakhil SR, et al. Oxaliplatin, fluorouracil, and leucovorin for patients with unresectable liver-only metastases from colorectal cancer: a North Central Cancer Treatment Group phase II study. *Journal of Clinical Oncology* 2005;**23**(36):9243–9.

Altman 1992

Altman DG. *Practical statistics for medical research*. London: Chapman and Hall, 1992.

Anderson 1996

Anderson RT, Aaronson NK, Bullinger M, McBee WL. A review of the progress towards developing health-related quality of life instruments for international clinical studies and outcomes research. *PharmacoEconomics* 1996;**10**(4): 336–55.

Apolone 1998

Apolone G, Mosconi P. The Italian SF-36 Health Survey: translation, validation and norming. *Journal of Clinical Epidemiology* 1998;**51**(11):1025–36.

Araghizadeh 2001

Araghizadeh FY, Timmcke AE, Opelka FG, Hicks TC, Beck DE. Colonoscopic perforations. *Diseases of the Colon and Rectum* 2001;**44**(5):713–6.

Audisio 1996

Audisio RA, Setti-Carraro P, Segala M, Capko D, Andreoni B, Tiberio G. Follow-up in colorectal cancer patients: a cost-benefit analysis. *Annals of Surgical Oncology* 1996;**3**(4): 349–57.

Audisio 2000

Audisio RA, Robertson C. Colorectal cancer follow-up: perspectives for future studies. *European Journal of Surgical Oncology* 2000;**26**(4):329–37.

Augestad 2014

Augestad KM, Rose J, Crawhsaw B, Cooper G, Delaney C. Do the benefits outweigh the side effects of colorectal cancer surveillance? A systematic review. *World Journal of Gastrointestinal Oncology* 2014;**6**(5):104–11.

Baca 2011

Baca B, Beart RW, Etzioni DA. Surveillance after colorectal cancer resection: a systematic review. *Diseases of the Colon and Rectum* 2011;**54**(8):1036–48.

Bowles 2004

Bowles CJ, Leicester R, Romaya C, Swarbrick E, Williams CB, Epstein O. A prospective study of colonoscopy practice in the UK today: are we adequately prepared for national colorectal cancer screening tomorrow?. *Gut* 2004;**53**(2): 277–83.

Böhm 1993

Böhm B, Schwenk W, Hucke HP, Stock W. Does methodic long-term follow-up affect survival after curative resection of colorectal carcinoma?. *Diseases of the Colon and Rectum* 1993;**36**(3):280–6.

Cella 1993

Cella DF, Tulsky DS, Gray G, Sarafian B, Linn E, Bonomi A, et al. The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. *Journal of Clinical Oncology* 1993;**11**(3):570–9.

Choti 2002

Choti MA, Sitzmann JV, Tiburi MF, Sumetchotimetha W, Rangsin R, Schulick RD, et al. Trends in long-term survival following liver resection for hepatic colorectal metastases. *Annals of Surgery* 2002;**235**(6):759–66.

Collopy 1992

Collopy BT. The follow-up of patients after resection for large bowel cancer, May 1992. Colorectal Surgical Society of Australia. *Medical Journal of Australia* 1992;**157**(9): 633–4.

Connor 2001

Connor S, Frizelle FA, Bagshaw PF. Follow-up after attempted curative surgery for colorectal cancer; postal survey of New Zealand surgeons' practice. *New Zealand Medical Journal* 2001;**114**(1129):151–3.

Covidence [Computer program]

Veritas Health Innovation. Covidence systematic review software. Melbourne, Australia. Available at www.covidence.org: Veritas Health Innovation, accessed date 23 October 2015.

DevCan 2005 [Computer program]

Statistical Research and Applications Branch National Cancer Institute. DevCan: Probability of developing or dying of cancer Software. Version 6.7.0. Statistical Research and Applications Branch National Cancer Institute, 2005.

Dolan 1997

Dolan P. Modeling valuations for EuroQol health states. *Medical Care* 1997;**35**:1095–108.

Dupuy 1984

Dupuy H. The Psychological General Well-Being (PGWB) Index. In: Wenger N, Mattson M, Furberg D, Elinson

J editor(s). *Assessment of quality of life in clinical trials of cardiovascular therapies*. New York: Le Jacq Publishing, 1984.

Edelman 1997
Edelman MJ, Meyers FJ, Siegel D. The utility of follow-up testing after curative cancer therapy: a critical review and economic analysis. *Journal of General Internal Medicine* 1997;**12**(5):318–31.

Edge 2010
Edge SB, Byrd SR, Compton CC. In: Edge SB, Byrd SR, Compton CC, et al. editor(s). *AJCC Cancer Staging Manual*. 7th Edition. New York: Springer Verlag, 2010.

Fleischer 1989
Fleischer DE, Goldberg SB, Browning TH, Cooper JN, Friedman E, Goldner FH, et al. Detection and surveillance of colorectal cancer. *JAMA* 1989;**261**(4):580–5.

Fossati 2015
Fossati R. Colorectal cancer follow-up review update overdue [personal communication]. Email to: GM Jeffery 04 May 2015.

Gandek 1998
Gandek B, Ware JE, Aaronson NK, Apolone G, Bjorner JB, Brazier JE, et al. Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results form the IQOLA Project. International Quality of Life Assessment. *Journal of Clinical Epidemiology* 1998;**51**(11):1171–8.

Gerdes 1990
Gerdes H. Surveillance after colon cancer: is it worthwhile?. *Gastroenterology* 1990;**99**(6):1849–51.

GRADEpro [Computer program]
GRADEpro GDT. GRADEpro GDT. Hamilton (ON): GRADE Working Group, McMaster University, 2015.

Grossman 2004
Grossmann EM, Johnson FE, Virgo KS, Longo WE, Fossati R. Follow-up of colorectal cancer patients after resection with curative intent-the GILDA trial. *Surgical Oncology* 2004;**13**(2-3):119–24.

Higgins 2002
Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* 2002;**21**(11):1539–58.

Higgins 2011
Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Higgins 2016
Higgins JPT, Lasserson T, Chandler J, Tovey D, Churchill R. In: Higgins JPT, Lasserson T, Chandler J, Tovey D, Churchill R. Standards for the conduct of new Cochrane Intervention Reviews.. *Methodological Expectations of Cochrane Intervention Reviews*. London: Cochrane, 2016.

Hunt 1980
Hunt SM, McKenna SP, McEwen J, Backett EM, Williams J, Papp E. A quantitative approach to perceived health status: a validation study. *Journal of Epidemiology and Community Health* 1980;**34**(4):281–6.

Kanas 2012
Kanas GP, Taylor A, Primrose JN, Langeberg WJ, Kelsh MA, Mowat FS, et al. Survival after liver resection in metastatic colorectal cancer: review and meta-analysis of prognostic factors. *Clinical Epidemiology* 2012;**4**:283–301.

Kievit 2000
Kievit J. Colorectal cancer follow-up: a reassessment of empirical evidence on effectiveness. *European Journal of Surgical Oncology* 2000;**26**(4):322–8.

Lee 2007
Lee WS, Yun SH, Chun HK, Lee WY, Yun HR, Kim J, et al. Pulmonary resection for metastases from colorectal cancer: prognostic factors and survival. *International Journal of Colorectal Diseases* 2007;**22**(6):699–704.

McArdle 2000
McArdle C. ABC of colorectal cancer: effectiveness of follow-up. *BMJ* 2000;**321**(7272):1332–5.

Moertel 1993
Moertel CG, Fleming TR, Macdonald JS, Haller DG, Laurie JA, Tangen C. An evaluation of the carcinoembryonic antigen (CEA) test for monitoring patients with resected colon cancer. *JAMA* 1993;**270**(8):943–7.

Muratore 2007
Muratore A, Zorzi D, Bouzari H, Amisano M, Massucco P, Sperti E, et al. Asymptomatic colorectal cancer with un-resectable liver metastases: immediate colorectal resection or up-front systemic chemotherapy?. *Annals of Surgical Oncology* 2007;**14**(2):766–70.

Ovaska 1989
Ovaska JT, Järvinen HJ, Mecklin JP. The value of a follow-up programme after radical surgery for colorectal carcinoma. *Scandinavian Journal of Gastroenterology* 1989;**24**(4):416–22.

Pawlik 2005
Pawlik TM, Scoggins CR, Zorzi D, Abdalla EK, Andres A, Eng C, et al. Effect of surgical margin status on survival and site of recurrence after hepatic resection for colorectal metastases. *Annals of Surgery* 2005;**241**(5):715–22.

Pfister 2004
Pfister DG, Benson AB, Somerfield MR. Clinical practice. Surveillance strategies after curative treatment of colorectal cancer. *The New England Journal of Medicine* 2004;**350**(23):2375–82.

Pita-Fernández 2014
Pita-Fernández S, Alhayek-Aí M, González-Martin C, López-Calviño B, Seoane-Pillado T, Pértega-Díaz S. Intensive follow-up strategies improve outcomes in nonmetastatic colorectal cancer patients after curative surgery: a systematic review and meta-analysis. *Annals of Oncology Advance Access* 2014;**00**:1–12.

Renehan 2002
Renehan A, Egger M, Saunders MP, O'Dwyer ST. Impact on survival of intensive follow-up after curative resection for

colorectal cacnerL systematic review and meta-analysis of randomised trials. *BMJ* 2002;**324**:1–8.

Renehan 2004
Renehan AG, O'Dwyer ST, Whynes DK. Cost effectiveness analysis of intensive versus conventional follow up after curative resection for colorectal cancer. *BMJ* 2004;**328** (7431):81–5.

RevMan 2014 [Computer program]
The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). Version 5.3, Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Rojas 2005
Rojas MP, Telaro E, Russo A, Fossati R, Palli D, Rosselli Del, et al. Follow-up strategies for women treated for early breast cancer. *Cochrane Database of Systematic Reviews* 2005, Issue 1. [DOI: 10.1002/14651858.CD001768.pub2]

Ryuk 2014
Ryuk JP, Choi GS, Park JS, Kim HJ, Park SY, Yoon GS, et al. Predictive factors and the prognosis of recurrence of colorectal cancer within 2 years after curative resection. *Annals of Surgical Treatment and Research* 2014;**86**(3): 143–51.

Sadahiro 2003
Sadahiro S, Suzuki T, Ishikawa K, Nakamura T, Tanaka Y, Masuda T, et al. Recurrence patterns after curative resection of colorectal cancer in patients followed for a minimum of ten years. *Hepatogastroenterology* 2003;**50**(53):1362–6.

Schünemann 2009
Schünemann H, Brozek J, Oxman A, editors. GRADE handbook for grading quality of evidence and strength of recommendation Version 3.2 (updated March 2009). The GRADE Working Group, 2009. Available from www.cc-ims.net/gradepro.

Sprangers 1993
Sprangers MA, Cull A, Bjordal K, Groenvold M, Aaronson NK. The European Organization for Research and Treatment of Cancer. Approach to quality of life assessment: guidelines for developing questionnaire modules. EORTC Study Group on Quality of Life. *Quality of Life Research* 1993;**2**(4):287–95.

Stiggelbout 1997
Stiggelbout AM, de Haes JC, Vree R, van de Velde CJ, Bruijninckx CM, van Groningen K, et al. Follow-up of colorectal cancer patients: quality of life and attitudes towards follow-up. *British Journal of Cancer* 1997;**75**(6): 914–20.

Sugarbaker 1987
Sugarbaker PH, Gianola FJ, Dwyer A, Neuman NR. A simplified plan for follow-up of patients with colon and rectal cancer supported by prospective studies of laboratory and radiologic test results. *Surgery* 1987;**102**(1):79–87.

Tierney 2007
Tierney JF, Stewart LA, Ghersi D, Burdett S, Sydes MR. Practical methods for incorporating summary time-to-event data into meta-analysis. Trials 2007; Vol. 8:16.

Tjandra 2007
Tjandra JJ, Chan MK. Follow-up after curative resection of colorectal cancer: a meta-analysis. *Diseases of the Colon and Rectum* 2007;**50**(11):1783–99.

Van Cutsem 2006
Van Cutsem EM, Nordlinger B, Adam R, Köhne CH, Pozzo C, Poston G, et al. European Colorectal Metastases Treatment Group. Towards a pan-European consensus on the treatment of patients with colorectal liver metastases. *European Journal of Cancer* 2006;**42**(14):2212–21.

Vernava 1994
Vernava AM, Longo WE, Virgo KS, Coplin MA, Wade TP, Johnson FE. Current follow-up strategies after resection of colon cancer. Results of a survey of members of the American Society of Colon and Rectal Surgeons. *Diseases of the Colon and Rectum* 1994;**37**(6):573–83.

Virgo 1995
Virgo KS, Vernava AM, Longo WE, McKirgan LW, Johnson FE. Cost of patient follow-up after potentially curative colorectal cancer treatment. *JAMA* 1995;**273**(23):1837–41.

Walker 1988
Walker AM, Martin-Moreno JM, Artalejo FR. Odd man out: a graphical approach to meta-analysis. *American Journal of Public Health* 1988;**78**(8):961–6.

Ware 1995
Ware J, Kosinski M, Keller S. *SF-12: How to Score the SF-12 Physical and Mental Health Summary Scales*. 2nd Edition. Boston Massachusetts: The Health Institute, New England Medical Centre, 1995.

Whistance 2009
Whistance RN, Conroy T, Chie W, Costantini A, Sezer O, Koller M, et al. European Organisation for the Research and Treatment of Cancer Quality of Life Group. Clinical and psychometric validation of the EORTC QLQ-CR29 questionnaire module to assess health-related quality of life in patients with colorectal cancer. *European Journal of Cancer* 2009;**45**(17):3017–26. [DOI: 10.1016/j.ejca.2009.08.014.]

Yoo 2006
Yoo PS, Lopez-Soler RI, Longo WE, Cha CH. Liver resection for metastatic colorectal cancer in the age of neoadjuvant chemotherapy and bevacizumab. *Clinical Colorectal Cancer* 2006;**6**(3):202–7.

Zigmond 1983
Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica* 1983;**67** (6):361–70.

References to other published versions of this review

Jeffery 2007

Jeffery M, Hickey BE, Hider PN. Follow-up strategies for patients treated for non-metastatic colorectal cancer. *Cochrane Database of Systematic Reviews* 2007, Issue 1. [DOI: 10.1002/14651858.CD002200.pub2]

* *Indicates the major publication for the study*

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Augestad 2013

Methods	RCT Accrual: 2007 to 2011 Median follow-up: 24 months
Participants	110 participants (65 men and 45 women) surgically treated for colon cancer Dukes' A: 24 Dukes' B: 55 Dukes' C: 32 Country: Norway Setting: hospital and community
Interventions	Experimental arm: surgeon follow-up Control: GP follow-up The follow-up intervals were the same.
Outcomes	• Quality of life (measured using EORTC QLQ-C30 and EuroQol-5D (EuroQol five dimensions questionnaire: EQ-5D) • Cost-effectiveness • Time to cancer diagnosis
Notes	National follow-up guidelines were applied in both study arms, and participants were followed for 2 years

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "Patients are randomised to follow-up either by their GP (intervention) or at the surgical clinic (controls)." Comment: the study did not report a description of the method used to generate randomisation sequence
Allocation concealment (selection bias)	Low risk	Quote: "...using a web-based randomisation service and managed by the Norwegian University of Science and Technology. " Comment: because the study used a web-based method, we assumed that it was truly concealed

Augestad 2013 (*Continued*)

Blinding of participants and personnel	Low risk	Quote (page 3): "Recruited patients were not informed about the other patients recruited in the same trial. Similarly, no information regarding trial progress and allocation was revealed to the participating GPs or surgeons. However, as GP-organised follow-up represented a new practice, blinding was not possible in the intervention arm." Comment: we judged this domain to be at low risk of bias.
Blinding of outcome assessment	Low risk	Quote (page 3): "The local trial investigator was not involved in the subsequent follow-up appointments in any way."Comment: this indicated that the assessors were blinded to the treatment arm; therefore, we judged this domain to be at a low risk of bias
Incomplete outcome data (attrition and exclusions)	Low risk	The study reported no exclusions, but detailed information with respect to attrition (detailed by arm and with reasons given) ensured that we judged this domain to be at low risk of bias
Selective reporting (reporting bias)	Unclear risk	**Outcomes specified in the objectives** • EORTC QLQ-C30 • EQ-5D • EQ-VAS • Cost-effectiveness • Time to diagnosis of relapse The paper reported on all of these. We did not have access to the protocol, so judged this outcome to be at unclear risk of bias
Other sources of bias	Unclear risk	We detected no other bias.

FACS 2014

Methods	RCT (1:1:1:1:1 minimisation algorithm, 2 x 2 randomised trial) Accrual: 2003 to 2009 Stratified for adjuvant chemotherapy, age, and sex Mean follow-up: 40.8 months Setting: tertiary centres
Participants	1202 participants (736 men and 466 women) treated with curative surgery for primary colorectal cancer Dukes' A: 254

FACS 2014 (*Continued*)

	Dukes' B: 553 Dukes' C: 354 Colon primary: 811 Rectal primary: 359 Country: United Kingdom
Interventions	• CEA testing every 3 months for 2 years, then every 6 months for 3 years with a single CT scan of the chest/abdomen/pelvis if requested at study entry by clinician • CT scan of the chest/abdomen/pelvis every 6 months for 2 years, then annually for 3 years, plus colonoscopy at 2 years • CEA and CT follow-up: both blood and imaging as above, plus colonoscopy at 2 years • Minimum follow-up: no scheduled follow-up except a single CT scan of the chest/abdomen/pelvis if requested at study entry by a clinician
Outcomes	• Recurrence (loco-regional, distant metastases, interval recurrences) • New cancers • Surgical salvage • Survival • DFS • Compliance
Notes	Eligible participants were those with no residual disease (confirmed by a CT scan of the chest and liver or a MRI of the liver), microscopically clear margins, and postoperative CEA $\leq$ 10 μ/L following surgery or completion of adjuvant therapy as indicated. All participants had colonoscopy at trial entry to ensure there was no residual intraluminal disease and were offered an end-of-trial colonoscopy

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote (page 265): "Randomization to 1 of 4 groups (Figure 1) on a 1:1:1:1 ratio was performed centrally at the Oxford Clinical Trials Unit using a minimization algorithm to balance patient characteristics within each centre based on 3 variables: adjuvant chemotherapy, sex, and age group." Comment: we judged this domain to be at low risk of bias.
Allocation concealment (selection bias)	Low risk	Quote (page 265): "Study nurses contacted the Oxford Clinical Trials Unit by telephone to enter a patient in the trial, reporting the relevant patient characteristics; they were then told the trial group to which the patient had been allocated." Comment: we judged this domain to be at

		low risk of bias.
Blinding of participants and personnel	Unclear risk	Quote (page 265): "Because this was a pragmatic open trial, it was not possible to conceal the allocation group from either participants or clinicians." Comment: we judged this domain to be at unclear risk of bias
Blinding of outcome assessment	Low risk	Quote: "However, the research staff who abstracted outcome data from clinical notes were employed by the local National Cancer Research Network teams independent of the investigators. The analysis program was undertaken first using dummy variables for the allocation groups and the code was not broken until the precise procedures for analysis were agreed on." Comment: we judged this domain to be at low risk of bias.
Incomplete outcome data (attrition and exclusions)	Low risk	Quote (page 265): "Exclusions: • Intensive arm: 6 exclusions, 1 withdrew consent, 6 had residual disease • Minimal arm: 3 excluded (1 entered a conflicting study, 2 had residual disease)" Comment: there was complete information given to explain the exclusions, and they were reported by arm and given the 3:1 randomisation ratio; the numbers were similar in each arm. We therefore judged this domain to be at low risk of bias (figure 1). Attrition was reported; it did not occur.
Selective reporting (reporting bias)	Low risk	**Primary outcome measures (recorded on isctrn.org)** Current primary outcome measure amended as of 11 February 2009: Number of recurrences in each group treated surgically with curative intent, analysed at study end (5 years) Previous primary outcome measure: Overall survival by intention-to-treat analysis **Secondary outcome measures** Current secondary outcome measures as of 11 February 2009: 1. Overall survival by intention-to-treat analysis, reviewed at study end (5 years) 2. Quality of life in survivors, assessed

FACS 2014 (*Continued*)

		at baseline and then at the end of study years 1 to 5 by the following: ○ 2. 1. EuroQol-5D (EQ-5D) ○ 2. 2. European Organization for Research and Treatment of Cancer Quality of Life Questionnaire for Cancer patients (EORTC QLQ-C30) ○ 2. 3. Hospital Anxiety and Depression Scale (HADS) ○ 2. 4. Modified form of a College of Health Questionnaire ○ 2. 5. A small number of items from the 7-item questionnaire used by Kjeldsen 1997 3. Cost of NHS services utilised (data collected at the end of study years 1 to 5 for all participants) 4. NHS cost per life-year saved, assessed at study end (5 years) **Outcome measures reported** Primary outcome: • Surgical treatment of recurrence with curative intent Secondary outcomes: • Overall survival • Colorectal cancer-specific survival • Time to detection of recurrence • Survival after treatment of recurrence with curative intent • Recurrences • Compliance We assumed that this was the initial publication and that subsequent publications will present the quality of life data, so judged this to be at low risk of bias
Other sources of bias	Unclear risk	We detected no other bias.

GILDA 1998

Methods	RCT Accrual dates: 1998 to 2006 Multicentered, international study Median follow-up: 62 months Setting: not stated
Participants	1228 participants (746 men and 482 women) with histopathologic diagnosis of adenocarcinoma of the colon or rectum, Dukes Astler-Coller stage B2-C, treated with curative intent (radical excision plus or minus adjuvant radio/chemotherapy)

	Participant must be free of known cancer prior to entry, attested by normal endoscopy, US, CXR, and CEA **Exclusion criteria** • Inability to undergo testing (disability, allergy to contrast, etc.) or geographically not amenable to follow-up • Enrolment in any other protocol requiring specific follow-up practice • History of any previous malignancy in the last 10 years (other than CIS of the cervix or non-melanoma skin cancer) • No informed consent
Interventions	**Experimental group programme** • 4, 8, 12, 16, 20, 24, 30, 36, 42, 48, and 60 monthly office visits and history and clinical examination, FBC, CEA, and CA 19-9 • Colonoscopy and CXR at 12, 24, 36, 48, and 60 months • Liver US at 4, 8, 12, 16, 24, 36, 48, and 60 months • For rectal participants, pelvic CT at 4, 12, 24, and 48 months **Control group programme** • 4, 8, 12, 16, 20, 24, 30, 42, 48, and 60 monthly office visits, including history, examination, and CEA • Colonoscopy at 12 and 48 months • Liver ultrasound at 4 and 16 months • Rectal cancer participants in addition had rectoscopy at 4 months, CXR at 12 months, and liver US at 8 and 16 months. A single pelvic CT was allowed if a radiation oncologist required it as baseline following adjuvant treatment
Outcomes	**Principal endpoints** • Overall survival • Specific mortality **Secondary endpoints** • Quantify lead time due to intensive programme • Treatment of recurrences with curative intent • Sensitivity of follow-up regimens • Compliance with follow-up regimen • Quality of life HRQoL self-assessed at baseline and at 12, 24, 36, 48, and 60 months • Relapsed participants were monitored every two months for 1 year using EORTC QLQ-C30
Notes	-

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote (abstract, page 274): "Colon cancer patients were randomised." Comment: as no details were given, we deemed this domain to be at unclear risk of bias

GILDA 1998 (*Continued*)

Allocation concealment (selection bias)	Unclear risk	Quote (paragraph 5, page 6): "Randomisation was performed centrally via telephone at the...Italy." Comment: this implied that randomisation was remote and potentially concealed, but the details were not given
Blinding of participants and personnel	Unclear risk	There was no mention of blinding of either participants or personnel, so this was probably not done.
Blinding of outcome assessment	Unclear risk	There was no mention of blinding with regard to the outcome assessors, so this was probably not done
Incomplete outcome data (attrition and exclusions)	Low risk	Quote (figure 2, Rosati 2016): "3 patients from 3 centres that ceased collaboration soon after randomisation, four patients not eligible..." Comment: the study gave details of reasons for exclusions and attrition by arm, so we judged this domain as not at high risk of bias
Selective reporting (reporting bias)	Unclear risk	The study reports on the primary outcome and selected secondary outcomes (quality of life measures). We did not have access to the study protocol, so we judged this domain to be at unclear risk of bias
Other sources of bias	Unclear risk	We detected no other bias.

Kjeldsen 1997

Methods	RCT (random numbers) Accrual: 1983 to 1994 Stratified for Dukes' stage and location Country: Denmark Setting: not stated Follow-up: not stated
Participants	597 participants, (326 men and 271 women) treated with primary radical surgery for CRC, no residual neoplasia **Inclusion criteria** • Aged less than 76 years • No complicating disease making follow-up impossible • No other major cancer within the past 5 years

Kjeldsen 1997 (*Continued*)

	• Permanent residency within the county of Funen Dukes' A 138 Dukes' B: 293 Dukes' C: 166 Colon primary: 314 Rectal primary: 283
Interventions	The experimental group had follow-up examinations at 6, 12, 18, 30, 36, 48, 60, 120, 150, and 180 months after radical surgery. The control group had examinations at 60, 120, and 180 months. Examinations included medical history, clinical examination, digital rectal examination (DRE), gynaecological examination, Haemoccult-II test, colonoscopy, CXR, haemoglobin level, erythrocyte sedimentation rate, and liver enzymes
Outcomes	• Local recurrence • Metachronous colorectal cancer • Overall survival • Cancer-related survival
Notes	Definition of radical surgery: no residual neoplasia detected by the following examinations: complete colonoscopy or incomplete colonoscopy plus double-contrast barium enema, CXR (2 views), histological evaluation of all surgical margins, biopsy of suspicious lesions (lymph nodes), inspection and palpation of liver during surgery Local recurrence was defined as growth in the region of the primary radical operation, including the surgical wound, and demonstrated clinically or by imaging techniques, but not necessarily verified by biopsy. New lesions were called metachronous when diagnosed at least 12 months after primary cancer

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote (page 666): "After surgery, the patients were allocated to one of two follow-up programmes (groups 1 and 2) by random numbers." Comment: the use of random numbers may be adequate, but there is not enough description to be certain; therefore, we graded this as unclear
Allocation concealment (selection bias)	Unclear risk	Quote (page 666): "After surgery, the patients were allocated to one of two follow-up programmes (groups 1 and 2) by random numbers." Comment: there was no description of concealment of allocation, so it was probably not done

Blinding of participants and personnel	Unclear risk	Quote (page 666): "All patients were also instructed to visit their general practitioner if they developed abdominal pain or changing bowel habits lasting more than 2 weeks after the immediate postoperative period." Comment: Participants: not mentioned, although it is possible that those participants in the minimal follow-up group may have put more weight on their symptoms in the knowledge that they had fewer planned investigations. We judged that this domain was probably not at risk of bias Assessors: not mentioned, but there is a risk that in the minimal follow-up group, personnel may have put more weight on reported symptoms in the knowledge that they had fewer planned investigations. However, there were prespecified follow-up schedules
Blinding of outcome assessment	High risk	Quote (page 666): "Local recurrence was defined as growth in the region of the primary radical operation, including the surgical wound, and demonstrated clinically or by imaging techniques, but not necessarily verified by biopsy." Quote (page 666): "Group 1 had follow-up examinations at 6, 12, 18, 24, 30, 36, 48, 60, 120, 150 and 180 months after radical surgery, while group 2 had examinations at 60, 120 and 180 months." 1. Objective outcomes: blinding not mentioned (likely to be a source of bias) 2. Subjective outcomes: not measured
Incomplete outcome data (attrition and exclusions)	Low risk	Quote (page 668): "In all, 88 of 290 patients in group 1 and 100 of 307 in group 2 have died." Comment: this implies that they had followed all participants
Selective reporting (reporting bias)	Unclear risk	Quote (page 666): "The main purpose of the present randomised study was to evaluate the possible influence of follow-up upon survival." "Recurrence (and/or distant spread)" Comment: the primary outcome of survival was reported on in the results. There

Kjeldsen 1997 (*Continued*)

		were multiple outcomes reported, but as we were not able to review the protocol, we judged the risk of bias for this domain as "unclear" Outcomes reported included the following: • recurrence (local, distant, symptomatic, and asymptomatic) • "cancer-free time" • curative salvage surgery • metachronous primaries • colorectal cancer deaths
Other sources of bias	Unclear risk	We detected no other bias.

Mäkelä 1995

Methods	RCT Accrual: 1988 to 1990 Single centre trial Follow-up: 60 months Setting: tertiary centre
Participants	106 participants (52 men, 54 women) who had "radical primary surgery for CRC" at the Oulu University Hospital (1988 to 1990) Dukes' A: 28 Dukes' B: 48 Dukes' C: 30 Colon primary: 75 Rectal primary: 31 Country: Finland
Interventions	Experimental group: participants who had rectal or sigmoid cancers had flexible sigmoidoscopy with video imaging every 3 months, colonoscopy at 3 months (if it had not been done pre-operation), then annually. They also had ultrasound of the liver and primary site at 6 months, then annually. Control group: participants who had rectal and sigmoid cancers had rigid sigmoidoscopy and barium enema annually
Outcomes	• Local recurrence • Regional recurrence • Time to detection of recurrence • Recurrence rates • Method of detection of recurrence • Mode of recurrence • Resectability • Overall survival

Notes	Radical resection: macroscopic removal of, with microscopically negative margins Local recurrence: restricted to anastomosis and its surrounds Regional recurrence: invasion beyond the site of the primary without distant metastases All participants reviewed at 3, 6, 9, 12, 15, 21, 24, 30, 36, 42, 48, 54, and 60 months At each visit: history, examination, FBC, faecal occult blood test, CEA, CXR performed

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	The study provided no more details regarding sequence allocation; therefore, we judged this domain to be at unclear risk of bias
Allocation concealment (selection bias)	Unclear risk	The study reported no details of allocation concealment, so we judged this domain to be at unclear risk of bias
Blinding of participants and personnel	Low risk	Participants: not mentioned, but the clear instructions to the participants (see paragraph 4, page 620) about when they were to contact the surgical department meant we judged this domain to be low risk of bias. Assessors: not mentioned, but the protocol for intensive follow-up was prespecified, as was the process for the minimal group. This would have reduced the risk of bias
Blinding of outcome assessment	Unclear risk	Participants: not mentioned, but probably not done. This could have been done relatively easily, but was unlikely to have introduced bias Assessors: not mentioned, but probably not done. This may have introduced bias because the personnel were aware that there were few planned investigations in the minimal group; this may have meant greater weight was placed in reported symptoms in this group. This could be a source of bias
Incomplete outcome data (attrition and exclusions)	Low risk	The study did not report incomplete outcome data, so we judged this domain to be at low risk of bias (both exclusions and attrition)

Mäkelä 1995 (*Continued*)

Selective reporting (reporting bias)	Unclear risk	Specified in the methods: recurrence (regional and anastomotic), time to detection of recurrence, method of detection of recurrence, surgery for recurrence, survival, synchronous adenomas detected during surveillance Actually reported: time to recurrence, recurrences (local, regional, and distant), method that detected recurrence most frequently, presence of symptoms at recurrence, method of detection of recurrence, surgery for recurrence, survival, survival after radical surgery of recurrence, adenomas detected during surveillance We did not have access to the protocol, so judged this domain to be an unclear risk of bias
Other sources of bias	Unclear risk	We detected no other bias.

Ohlsson 1995

Methods	RCT Accrual: 1983 to 1986 Single centre trial Setting: tertiary Follow-up: 66 to 105.6 months
Participants	107 participants (51 men, 56 women) undergoing resection with curative intent for CRC at the departments of surgery in Lund and Helsingborg, Sweden, from 1983 to 1986 **Exclusion criteria** • Local excision only • Distant metastases • Participants in whom age or severe illness was considered to preclude treatment of recurrent disease • Inability to co-operate • Crohn's disease • Ulcerative colitis • Familial polyposis • Incomplete colonoscopy together with uncertain findings at barium enema examination Dukes' A: 19 Dukes' B: 47 Dukes' C: 41 Colon primary: 71 Rectal primary: 36 Country: Finland

Ohlsson 1995 (*Continued*)

	Setting: hospital
Interventions	The experimental group were seen at 3-, 6-, 9-, 12-, 15-, 18-, 21-, 24-, 30-, 36-, 42-, 48-, and 60-month intervals. Performed at each visit were clinical exam, rigid proctosigmoidoscopy, CEA, alkaline phosphatase, gamma-glutaryl transferase, faecal haemoglobin, and CXR. Examination of anastomosis (flexible sigmoidoscopy or colonoscopy, as dictated by the lesion) was performed at 9, 21, and 42 months. Colonoscopy was performed at 3, 15, 30, and 60 months. CT of the pelvis was performed at 3, 6, 12, 18, and 24 months. The control group had no follow-up visits planned. They received written instructions recommending that they leave faecal samples with the district nurse for examination every third month during the first 2 years after surgery then once a year. They were instructed to contact the surgical department if they had any symptoms
Outcomes	• Overall survival • Local recurrence • Anastomotic recurrence • Symptomatic recurrence • Resection with curative attempt • Time to first recurrence • Protocol compliance
Notes	Local recurrence: recurrence within the initial bed, operative field, anastomosis, or structures contiguous or adherent to the primary (included relapse in the abdominal wound, drain site, pelvis, or perineum) Anastomotic recurrence: intraluminal recurrence within 5 cm of the anastomosis Symptomatic: when symptoms could be related to the participant's initial illness and when they resulted in or would have resulted in the participant seeking advice Resection with curative attempt: all visible removed, microscopic-free margins Time to first recurrence: interval between primary surgery and unequivocal demonstration of recurrence at laparotomy, imaging, or autopsy

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Quote: "Patients were randomised to..." Comment: the study provided no details about sequence generation; therefore, we judged this as unclear
Allocation concealment (selection bias)	Unclear risk	The paper provided no details, so we judged this as unclear.
Blinding of participants and personnel	Low risk	Quote (page 620): "No follow-up visits were planned for patients in the control group. They received written instruction, recommending they leave fecal samples with the district nurse for examination of

Ohlsson 1995 (*Continued*)

		haemoglobin every third month during the two first years after surgery and then once a year. They were also instructed to contact the surgical department as soon as they experienced any problems with the colostomy, abdominal or perineal pain, altered bowel movements, change in fecal colour, micturition problems, or weight loss. Protocol for active follow-up is given in Table 1." Comment: although it was not mentioned, participants received clear instructions about when they were to contact the surgical department, and the follow-up protocols for both groups were prespecified, which would have reduced the risk of bias
Blinding of outcome assessment	Unclear risk	The study did not mention blinding of outcome assessment; it would have been possible to blind those reporting the investigations treatment arm, but not having done so is unlikely to have introduced bias
Incomplete outcome data (attrition and exclusions)	Low risk	Quote (page 623): "Twenty-two of 54 patients in the control group and 15 of 53 patients in the F-U group were dead at the end of the study..." Comment: all participants reported on attrition; there was no attrition
Selective reporting (reporting bias)	Unclear risk	We did not have access to the study protocol, so we judged this domain at unclear risk of bias
Other sources of bias	Unclear risk	We detected no other bias.

Pietra 1998

Methods	RCT Accrual: 1987 to 1990 Single centre trial Setting: university hospital Follow-up: 60 months
Participants	207 consecutive participants (111 men, 96 women) who had curative resections for large bowel cancer; all had colonoscopy at 3 months post operation if had not been done preoperatively **Exclusion criteria**

	• Dukes' A • Liver metastases • Severe concurrent illness precluding follow-up or treatment of recurrent disease Dukes' A: 0 Dukes' B: 122 Dukes' C: 85 Colon primary: 139 Rectal primary: 68 Country: Italy Setting: university hospital
Interventions	The experimental group were seen at 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 42, 48, 54, and 60 months, then annually thereafter. There was clinical examination, ultrasound, CEA, and CXR at each visit. Annual CT of the liver and colonoscopy were performed. The control group were seen at 6 and 12 months, then annually. At each visit, clinical examination, CEA, and ultrasound were performed. They had annual CXR, yearly colonoscopy, and CT scan
Outcomes	• Local recurrence • Intramural recurrence • Overall survival
Notes	Local recurrence: all local disease detectable at follow-up, either alone or in conjunction with generalised recurrence Local recurrences were divided into extramural recurrences, where regrowth was located in and around the bed, including the pericolic fat, adjoining mesentery, or lymph nodes Intramural recurrence: regrowth involving only the anastomosis A local recurrence was considered resected when no macroscopic/microscopic disease remained after surgery

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Quote (paragraph 3, page 1128): "were randomly assigned" The study reported no details of the method of sequence generation, which makes this domain at unclear risk of bias
Allocation concealment (selection bias)	Unclear risk	The study reported no details, so we judged this domain to beat unclear risk of bias
Blinding of participants and personnel	Low risk	Participant: not mentioned, unlikely to have introduced bias Assessors: quote (paragraph 3, page 1128) : "The same clinical and instrumental tests CT included were performed whenever a patient of either group had symptoms sug-

		gestive of a possible recurrence of the disease (abdominal or perineal pain, altered bowel movements, change in fecal colour, or weight loss)." The prespecified follow-up schedules and the lists symptoms to be investigated make this domain at low risk of bias
Blinding of outcome assessment	Unclear risk	Blinding of outcome assessment was not mentioned. Local recurrence was the primary outcome measure so susceptible to bias
Incomplete outcome data (attrition and exclusions)	High risk	Quote (page 1128): "Nine patients (3.8 per cent with Stage A s were excluded from the study because our previous reports demonstrated a low rate of recurrences in these cases. Other exclusion criteria were the presence of liver metastases (4 patients) , even though these had been removed in apparently radical fashion during surgery on the primary, and the presence of severe illness that precluded intense follow-up or treatment of recurrent disease (10 patients) . The remaining 207 patients were enrolled in this study..." Exclusions: the exclusions were not reported by study arm. It is not clear from the report whether these exclusions occurred before randomisation. This means that the study potentially excluded 37/239 (15%) of those randomised. As little information has been provided, we have judged this to be at high risk of bias. Attrition: none reported, so we judged this domain at low risk of bias
Selective reporting (reporting bias)	Unclear risk	**Outcomes specified in the methods** 1. Detection of local recurrence (LR) 2. Resectability of LR 3. Survival **Outcomes actually reported in the paper** 1. LR (any LR, isolated LR, combined LR, interval LR, site of LR) 2. Curative resection 3. Metachronous primaries 4. DFS for those having curative resections at recurrence 5. Survival for all and for those who

Pietra 1998 (*Continued*)

		had curative resection of recurrence 6. DFS We did not have access to the protocol, so judged this outcome to be at unclear risk of bias
Other sources of bias	Unclear risk	We detected no other bias.

Rodríguez-Moranta 2006

Methods	RCT, Accrual: 1997 to 2001 Multicentred, stratified for centre, location, and stage Setting: hospital Follow-up: 48 months
Participants	259 participants (161 men and 98 women), stage II and III colon and rectal cancer Country: Spain
Interventions	The experimental group were seen with history, examination, and bloods (including CEA), US/CT, CXR, and colonoscopy. The control group were seen with history, examination, and bloods (including CEA)
Outcomes	• Local recurrence • Curative reoperation rates • Overall survival
Notes	**Experimental group** • Seen with history, examination, and bloods (including CEA) at 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36, 39, 42, 45, 48, 51, 54, 57, and 60 months • US/CT at 6, 12, 18, 24, 30, 36, 42, 48, and 56 months • CXR and colonoscopy at 12, 24, 36, 48, and 56 months **Control group** • Seen with history, examination, and bloods (including CEA) at 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36, 39, 42, 45, 48, 51, 54, 57, and 60 months

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: "Patients were stratified according to centre, location (colon/rectum), and TNM stage (II/III); thereafter, patients were randomly allocated to either simple or intensive surveillance strategies by means of sealed envelopes containing computer-generated random numbers." Comment: we judged this method of se-

		quence generation to be at low risk of bias
Allocation concealment (selection bias)	Low risk	Quote: "Patients were randomly allocated to either simple or intensive surveillance strategies by means of sealed envelopes containing computer-generated random numbers. Random assignment was centralized at the Hospital Clinic." Comment: sequence generation was reported to be remote, but as the study gave no further details, we rated this domain as at unclear risk of bias
Blinding of participants and personnel	Unclear risk	Participants: blinding to treatment arm was not mentioned, but this would have been difficult to do. As history and examination was performed for both arms, it is unlikely to have been a cause of bias for objective outcomes. The study reported no subjective outcomes. Personnel: not mentioned, unlikely to have been done. The follow-up schedule was specified (see Table 3, page 387). This would reduce the risk of bias. Knowledge of study arm could influence clinical decisions made on the basis of history and clinical findings, to influence further investigations, which could introduce potentially introduce bias. We therefore judged this domain to be at high risk of bias
Blinding of outcome assessment	Unclear risk	The study did not mention blinding of outcome assessment, but it is unlikely to have been a source of bias. Therefore, we judged this domain to be at low risk of bias
Incomplete outcome data (attrition and exclusions)	Low risk	Quote: "During the study period, 270 patients were included. Eleven patients (4%) were excluded after random assignment because of inadequate initial assessment of tumor stage (eight had distant metastases and three had a stage I tumor). Consequently, 259 participants constitute the basis of this study." Exclusions: these exclusions were reported by arm: • Intensive: 2 with Stage I s, 4 with Stage IV s

Rodríguez-Moranta 2006 (*Continued*)

		• Minimal: 1 with Stage I, 4 with Stage IV s This is unlikely to have introduced bias, as the reasons are similar for exclusions in each arm and the numbers excluded in each arm are similar. Attrition: quote (page 388): "No patient was lost to follow-up" Comment: this is unlikely to have introduced bias.
Selective reporting (reporting bias)	Unclear risk	**Outcomes specified in the methods** 1. Survival 2. Resectable recurrence **Outcomes actually reported in the paper** 1. Overall survival, survival by stage and location 2. Cause-specific survival 3. Recurrence 4. Time to relapse 5. Type of relapse 6. Resectable recurrences 7. Resectable recurrences by stage and location 8. Method of detecting first recurrence 9. Metachronous recurrences 10. Cost of follow-up 11. Cost per resectable recurrence We were not able to review the protocol, so we judged this domain to be at unclear risk of bias
Other sources of bias	Unclear risk	We detected no other bias.

Schoemaker 1998

Methods	RCT Accrual: 1984 to 1990 2 centres in the trial Stratified according to site (colon or rectum) and Dukes' stage Setting: tertiary centres Follow-up: 60 months
Participants	325 participants (207 men and 118 women) who had curative resection of newly diagnosed CRC **Exclusion criteria** • Medical comorbidity making follow-up difficult or 5-year survival unlikely • Residence in a remote area • Aged > 85 years

	• Refusal to participate in the trial • Evidence of residual or distant malignancy Dukes' A: 71 Dukes' B: 153 Dukes' C: 101 Colon primary: 238 Rectal primary: 87 Country: Australia
Interventions	Participants in the experimental arm underwent yearly CXR, CT of the liver, and colonoscopy. These investigations were only performed in the control group if indicated on clinical grounds or after screening test abnormality, and at 5 years of follow-up, to exclude a reservoir of undetected recurrences
Outcomes	• Overall survival
Notes	Both groups had regular clinical review, including history, examination, and screening investigation at 3, 6, 9, 12, 15, 21, 24, 30, 36, 42, 48, 54, and 60 months or until a major endpoint was reached. A nurse research assistant performed a review at each visit, and a consultant surgeon, on at least alternate visits. Clinical signs and symptoms were recorded on a structured ProForma. Screening investigations at each visit comprised of FBC, LFTs, CEA, and faecal occult blood testing using the Haemoccult-II test (without hydration) on 3 faecal samples. All screening or clinical abnormalities were investigated on merit. The only exception was CEA - an isolated rise in CEA was not used to trigger further investigations

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote (page 8): "The patients were then randomized to either standard or intensive follow-up by choosing the next card from a box of cards indicating the type of follow-up. The cards had been previously randomized using random tables." Comment: this is an adequate method of sequence generation.
Allocation concealment (selection bias)	High risk	Quote (page 8): "The patients were then randomized to either standard or intensive follow-up by choosing the next card from a box of cards indicating the type of follow-up." Comment: the study did not report details about how this was done, so we conclude that it makes this domain at high risk of bias

Schoemaker 1998 (*Continued*)

Blinding of participants and personnel	Low risk	Quote (page 8): "Review was performed by a nurse research assistant at each visit and by a consultant surgeon on at least alternate visits. Clinical symptoms and signs were obtained and recorded on a structured pro forma." Comment: the study did not mention blinding of participants and personnel, but it would have been difficult to do and unlikely to have introduced bias. The use of the pro-forma for data collection on symptoms and signs would reduce the risk of bias from lack of personnel and participant blinding
Blinding of outcome assessment	Low risk	Quote (page 8): "CXR and CT scans were interpreted by an independent senior radiologist. Colonoscopies were performed or supervised by recognized accredited colonoscopists and aimed to examine the entire residual colon to identify recurrence, metachronous carcinoma, and polyps." Comment: blinding of outcome assessors was likely to reduce the risk of bias
Incomplete outcome data (attrition and exclusions)	Low risk	Quote (page 8): "Eighteen patients withdrew from the study (standard, 8; intensive, 10). Three patients from each group were lost to follow-up at intervals ranging from 9 to 54 months because they moved to another state. Five patients in the standard group and 7 in the intensive group withdrew at intervals ranging from 3 to 54 months because of development of other medical illnesses that precluded further structured follow-up." Comment: the study did not report post-trandomisation exclusions. This attrition has been reported by study arm and the reasons are similar for the 2 arms, it is thus unlikely to be a source of bias
Selective reporting (reporting bias)	Unclear risk	**Outcomes specified in the methods** 1. Survival **Outcomes actually reported in the paper** 1. Withdrawals 2. Survival 3. Recurrences 4. Metachronous primaries

Schoemaker 1998 (*Continued*)

		5. Investigations i) Colonoscopies plus complications ii) CT scans iii) CXR 6. Lung and liver recurrences and survival after resection of these We did not have access to the study protocol, so we judged the risk of bias for this domain to be unclear
Other sources of bias	Unclear risk	We detected no other bias.

Secco 2002

Methods	RCT (2 "studies" within 1 publication) Accrual: 1988 to 1996 Single centre trial Setting: not stated Follow-up: 61.5 months (high-risk group) and 48 months (low-risk group)
Participants	337 participants (163 men and 174 women) who had curative surgery alone for colorectal cancer Participants were stratified into the following: 1. n = 200 high-risk: (adenocarcinoma rectum treated by low anterior resection, left colon adenocarcinoma modified Dukes B2 or T3, preoperative serum CEA greater than or equal to 7.5 ng/ml, Dukes stage C, poorly differentiated grade, mucinous adenocarcinoma, or signet ring cells) 2. n = 158 low-risk: participants had none of these characteristics Country: Italy
Interventions	108 high-risk participants were randomised to "intensive follow-up" (experimental arm) ; they had clinic visits and serum CEA, abdomen/pelvic US scans, and CXR. Participants with rectal carcinoma had rigid sigmoidoscopy and CXR. 84 high-risk participants were randomised to a "minimal follow-up programme performed by physicians"
Outcomes	• Overall survival (actuarial at 5 years) • Recurrence • Costs • Curative reoperations
Notes	Curative surgery: "macroscopic excision of primary, peri-rectal tissues and nodes" Experimental arm: clinic visits and serum CEA measured at 3, 6, 9, 12, 15, 18, 21, 24, 28, 32, 36, 42, 48, 54, and 60 months; abdomen/pelvic US scans at 6, 12, 18, 24, 30, 36, 48, and 60 months; CXR at 12, 24, 36, 48, and 60 months. Participants with rectal carcinoma had rigid sigmoidoscopy and CXR at 12, 24, 36, 48, and 60 months. Participants with rectal carcinoma had rigid sigmoidoscopy and CXR at 12, 24, 36, 48, and 60 months

	All participants received education regarding follow-up and the signs and symptoms of possible recurrence. All were expected to phone the surgical team every 6 months

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote (page 419): "Patients of each group were randomly included..." Comment: the study provided no details regarding sequence generation; therefore, we judged this domain to be at unclear risk of bias
Allocation concealment (selection bias)	Unclear risk	The paper provided no description of allocation concealment.
Blinding of participants and personnel	Unclear risk	The study did not mention blinding of participants and personnel. As participants in each arm were educated about signs and symptoms of a possible recurrence, it is possible that those allocated to the minimal arm might be more likely to report symptoms in the knowledge that they would not have any investigations performed in the absence of symptoms. This may have introduced bias; therefore, we judged this at high risk of bias. The lack of blinding for personnel is less likely to have caused bias, because the follow-up schedule was prespecified for both arms
Blinding of outcome assessment	Low risk	The study did not mention blinding of outcome assessment; it would have been possible to do so, but unlikely to have introduced bias
Incomplete outcome data (attrition and exclusions)	Unclear risk	Quote (paragraph 2, page 419): "Of the initial 358 patients...definitive randomisation of 337 patients" Comment: it is not clear whether these were prerandomisation or postrandomisation exclusions; we judged this domain to be at unclear risk of bias. Attrition: quote (paragraph 1, page 419): "Twenty-one (5.8%) patients dropped out over the first 13 months: eight cases from group 1 and 13 from Group 2." Comment: although the reasons for attri-

Secco 2002 (*Continued*)

		tion were not reported, those who dropped out were reported by study arm, and as the numbers were similar, we judged this domain to be at low risk of bias
Selective reporting (reporting bias)	Unclear risk	We did not have access to the study protocol, so we judged this domain to be at unclear risk of bias
Other sources of bias	Unclear risk	We detected no other bias.

Sobhani 2008

Methods	Randomised, single institution study, which enrolled participants from 7 French teaching hospitals Follow-up: 24 months or until death
Participants	N = 130 participants who had R0 (complete resections) surgery for colon or rectal cancer
Interventions	Experimental arm: PET performed at 9 and 15 months and conventional follow-up Conventional arm: conventional follow-up
Outcomes	• Recurrence • Time to recurrence • Time to second-line therapy • Surgical salvage (curative or not)
Notes	All participants followed the same schedule: 6 visits that included physical examination; CEA or CA 19-9, or both; ultrasound scan every 3 months (except at 9 and 15 months' follow-up); CXR every 6 months; and abdominal CT at 9 and 15 months' follow-up

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Quote (paragraph 2, page 876): "Patients were randomly divided..." Comment: the study reported no details of the process of sequence generation
Allocation concealment (selection bias)	Unclear risk	The paper gave no details about allocation concealment, so we judged this domain to be at unclear risk of bias
Blinding of participants and personnel	Low risk	The paper did not mention blinding of participants and personnel; it was probably not done, but because of prespecified protocols, we deemed this domain to be at low risk of

Sobhani 2008 (*Continued*)

		bias
Blinding of outcome assessment	Low risk	Quote (paragraph 3, page 876): "Physicians were unaware of the findings of the CT scan" Comment: the blinding of outcome assessors was not described, although because the finding of recurrence was dependent on biopsy and determined in a multidisciplinary clinic, we deemed this domain to be at low risk of bias
Incomplete outcome data (attrition and exclusions)	Low risk	Quote (paragraph 2, page 877): "One hundred and thirty patients (65 in each group) were evaluated in a ITT analysis." Comment: because the paper reported that all randomised participants were included in an ITT analysis, we deemed this domain to be at low risk of bias
Selective reporting (reporting bias)	Unclear risk	**Outcomes in the objectives** • Recurrence at 9 and 15 months • Time to recurrence • Time to second-line therapy **Reported outcomes** • Recurrences • Time to recurrence • Asymptomatic recurrences • Surgical salvage • Curative (R0) surgery • Number who had chemotherapy • Ddeaths We did not have access to the study protocol, so deemed this outcome to be at unclear risk of bias
Other sources of bias	Unclear risk	We detected no other bias.

Strand 2011

Methods	RCT Accrual: 2002 to 2005 Setting: tertiary centre Follow-up: 60 months
Participants	110 participants (59 men and 51 women) curatively operated on for colorectal cancer Country: Sweden Setting: hospital

Strand 2011 (*Continued*)

Interventions	Experimental arm: surgeon-led follow-up Control arm: nurse-led follow-up
Outcomes	• Participant satisfaction • Resource use • Medical safety
Notes	A nurse and surgeon performed follow-up in the same way: 6 monthly visits for 3 years, then annually up to 5 years. Symptom enquiry occurred at each visit (bloods and CEA as indicated) Abdomen US and CXR (replaced by CT in latter half of the study) at 1 and 3 years If "clean" colon was established preoperatively, colonoscopy at 5 years

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "Randomization to nurse (ES) or surgeon (KS) was performed using closed envelopes in blocks of four." Comment: the paper reported no more details, so we judged this domain to be at unclear risk of bias
Allocation concealment (selection bias)	Unclear risk	The stoma therapist (IN) provided written and verbal information and conducted the randomisation. We judged this domain to be at unclear risk of bias
Blinding of participants and personnel	Low risk	The study did not mention blinding of participants and personnel; it was probably not done, but unlikely to have introduced bias
Blinding of outcome assessment	Low risk	The study did not mention blinding of outcome assessment, but it was unlikely to have been done
Incomplete outcome data (attrition and exclusions)	Unclear risk	Quote (page 1001): "All patients completed the questionnaires." Comment: the paper reports that there was no attrition. Quote (page 1001): "One hundred and thirteen (113) consecutive patients were asked to participate in the study. Of these, three patients refused to participate, 56 were allocated to surgeon follow-up and 54 to nurse-led follow-up." Comment: it is not clear if these were pos-

Strand 2011 (*Continued*)

		trandomisation or prerandomisation exclusions
Selective reporting (reporting bias)	Unclear risk	**Outcomes prespecified** • Participant satisfaction • Resource utilisation • Medical safety **Actually reported** • Resource utilisation • Participant satisfaction • Medical safety and costs As we did not have access to the study protocol, we judged this domain to be at unclear risk of bias
Other sources of bias	Unclear risk	We detected no other bias.

Treasure 2014

Methods	RCT Accrual: 1982 to 1993 Multicentered trial Setting: not stated Follow-up: not stated
Participants	216 (128 men and 88 women) participants who have had potentially curative resection of colorectal cancer Dukes' A: 10 Dukes' B: 95 Dukes' C: 74 Country: UK Setting: hospital
Interventions	If a CEA rise occurred, the participants were randomised to the "aggressive" arm or "conventional" arm. In the "aggressive" arm, a CEA rise triggered the "second-look" surgery, with intention to remove any recurrence discovered
Outcomes	• Survival
Notes	All participants had identical clinical follow-up: every 3 months for the first 2 years, then every 6 months for the next 3 years. CEA was monthly for 3 years, then every 3 months for 2 years

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**

Random sequence generation (selection bias)	Low risk	Quote (page 5): "Patients were randomised equally between the two arms (1:1). Patients whose compliance was between 50% and 70% or whose immediate postoperative sample had not been received within the 4-6-week guideline were randomised in a separate stratum. Randomisation was also stratified by participating clinician. A block size of two was used in order to maintain as close a balance as possible between the two treatment arms." Comment: while the report did not describe the method of sequence generation, the authors who wrote this publication (not the original investigators) stated that the study was well performed, so we judged this domain to be at low risk of bias
Allocation concealment (selection bias)	Low risk	Quote (page 4): "The trial was coordinated (initially) from the Cancer Research Campaign Clinical Trials Centre at King's College Hospital." Comment: while the method of allocation concealment was not described, the study was co-ordinated remotely, so we judged this domain to be at low risk of bias
Blinding of participants and personnel	Low risk	Quote (page 6): "By the nature of the trial design, the clinician was blind as to whether such patients had been randomised to the 'Conventional' arm of the trial or had not been randomised because the CEA had failed to denote the presence of recurrent disease." Participants were blinded to treatment allocation. With regard to personnel, clinicians were blinded to which arm their participant had been allocated to; we judged this domain to be at low risk of bias
Blinding of outcome assessment	Low risk	The study did not mention blinding of outcome assessment. It was most likely not done. but is at little risk of introducing bias because of the blinding of participants and personnel

Incomplete outcome data (attrition and exclusions)	Unclear risk	Exclusions: not reported Attrition: not reported We judged this domain to be at unclear risk of bias.
Selective reporting (reporting bias)	Low risk	**Outcomes stated in the protocol** • A definitive answer concerning the effectiveness of CEA-prompted second-look surgery to improve survival • An accurate picture of the 'lead time' produced by CEA compared to clinical pick up of patients with recurrence. **Outcomes reported in the paper** • Deaths • Recurrences • Second-look laparotomy and subsequent surgery • Lead time for CEA detection of recurrence The third and fourth outcomes were thought to be preplanned subanalyses, but problems with data formatting meant they were not able to be reported. We judged this domain to be at low risk of bias
Other sources of bias	Low risk	Quote (page 8): "A Data Monitoring Sub-Committee (DMSC) composed of Working Party members not entering patients into the trial was asked to review the data after the first 100 patients had been randomised, which occurred in January 1988, and again after 200 patients had been randomised in February 1993. At this point it was recommended by the Data Monitoring Committee that the trial be stopped since it was very unlikely that any clinically important advantage would be demonstrated for patients undergoing second-look surgery." Comment: early stopping occurred; this was recommended by the trial monitoring committee, and was unlikely to have introduced bias

Wang 2009

Methods	RCT Accrual dates: Jan 1995 to March 2001 Setting: teaching hospital Follow-up: 64 to 79 months
Participants	326 participants (177 men and 149 women) who had had curative resection for colorectal cancer Stratified for location (colon or rectum) and Dukes' stage. Dukes' A: 53 Dukes' B: 186 Dukes' C: 93 Country: China
Interventions	Experimental arm: colonoscopy at each visit Control arm: colonoscopy at six months, , 30 months, and 60 months from randomisation
Outcomes	• Overall survival (5-year survival and HR) • Postoperative cancer (anastomotic, extraluminal recurrence, and metachronous primaries) • CRC deaths • Salvage surgery • Asymptomatic recurrences • Major colonoscopy complications
Notes	All participants had clinic visits 3/12 for 12/12, 6/12 for 24/12, then 12/12 for 24/12. At each visit, history and examinationwas performed, as well as CEA, CXR, and liver imaging (CT or US). In each group, more investigations and examinations were performed if symptomatic Curative resection: no macroscopic residual and clear pathological margins All recurrences were confirmed histologically. Local recurrence were divided into anastomotic (intraluminal recurrence within 5 cm of the initial primary) and extraluminal Metachronous: second primary colorectal cancer after exclusion of a synchronous primary by a preoperative colonoscopy or within 6 months postoperatively Salvage surgery was considered curative when all macroscopic was removed and pathological margins were clear doi: 10.1016/j.gie.2008.05.017

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote (page 610): "The patients were then randomized to either the RCS group or the ICS group by means of sealed envelopes containing cards printed with ICS or RCS within each stratum." Comment: the paper gave insufficient de-

		tails of the randomisation process
Allocation concealment (selection bias)	Unclear risk	Quote (page 610): "Sealed envelopes containing cards printed with ICS or RCS within each stratum..." Comment: there is insufficient detail given to be sure that allocation concealment was truly concealed; it is not stated who performed the randomisation and if the envelopes were opaque
Blinding of participants and personnel	High risk	The study did not mention blinding of participants or personnel. Participants in the control arm may have been more likely to report symptoms knowing that they were not having colonoscopies
Blinding of outcome assessment	High risk	Quote (page 610): "More complete and systematic examinations were performed whenever a patient in either group had symptoms suggestive of a possible recurrence of the disease (e.g. abdominal or perineal pain, altered bowel movements, change in fecal colour, weight loss)." Comment: lack of blinding may have introduced bias with assessment of reported symptoms or signs, perhaps making further investigation more likely in the control group
Incomplete outcome data (attrition and exclusions)	Unclear risk	Quote (page 611): "Seven patients (ICS, 4; RCS, 3) were lost during the follow-up period." Comment: the paper gave no details about why they were lost to follow-up, which may introduce bias
Selective reporting (reporting bias)	Unclear risk	**Outcomes stated in the paper** • Survival • Local recurrence • Distant metastases • Metachronous CRC • Anastomotic recurrences • Intraluminal recurrences • Extraluminal recurrences **Outcomes reported** • OS • CRC deaths • Postoperative CRC (local recurrences

Wang 2009 (*Continued*)

		and metachronous) • Time to relapse • Asymptomatic recurrence • Salvage surgery • Adverse outcomes We did not have access to the study protocol, so rated the risk of bias for this outcome as unclear
Other sources of bias	Unclear risk	We detected no other bias.

Wattchow 2006

Methods	RCT Accrual: 1998 to 2001 Multicentred trial Follow-up: 24 months Randomisation method: remote and concealed (random numbers) Single blinded (researchers) Baseline characteristics: balanced other than trend to higher education levels in surgeon follow-up group Power calculation: power of 80% (2-sided) significant at 0.05, based on primary outcome measures (quality of life, anxiety and depression, and participant satisfaction). Number of participants required was 64, set target of 100 participants in each arm
Participants	203 participants (117 men and 86 women) who had undergone curative surgery for Dukes' A, B, or C colon cancer who had completed any postsurgical chemotherapy (rectal cancer excluded because of requirement for sigmoidoscopy in follow-up) Follow-up by general practitioners and surgeons had to be available, and informed consent given. Participants were randomised at either postsurgical visit or at completion of chemotherapy Country: Australia Setting: primary versus secondary care
Interventions	Setting and environment of follow-up (primary versus secondary care) Follow-up guidance was based on current clinical practice, and guidance was provided that suggested follow-up visits as follows: every 3 months for the first 2 years postoperatively, then every 6 months for the next 3 years. Each visit incorporated asking a list of set questions about symptoms, physical examination, annual faecal occult blood testing, and colonoscopy every 3 years
Outcomes	**Primary** • Quality of life • Depression and anxiety • Participant satisfaction **Other** • Number and type of investigations

	• Number and time to detection of recurrences • Deaths from all causes at 2 years
Notes	Quality of life assessment was based on SF-12 physical and mental health component scores at baseline and 12 and 24 months. Depression and anxiety assessment was based on the Hospital Anxiety and Depression scale, measured at baseline and 12 and 24 months. Participant satisfaction was based on Patient Visit-Specific Questionnaire measured at 24 months

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote (page 1117): "Consenting patients were then randomly allocated to either 'GP-led' or 'surgeon-led' follow-up using an Excel random number generator." Comment: we judged this domain to be at low risk of bias.
Allocation concealment (selection bias)	Low risk	Quote (page 1117): "Randomisation was conducted by the study researchers, who were not involved in the design of the study or the clinical care of the patients, and was concealed until the interventions were assigned." Comment: we judged this domain to be at low risk of bias.
Blinding of participants and personnel	Low risk	Quote (page 1117): "The study was single-blinded. Patients were reviewed by GPs in their practice rooms and surgeons in their surgical clinics." Comment: we assumed this means that they were blinded to treatment arm
Blinding of outcome assessment	Low risk	Quote (page 1117): "Researchers at all times were unaware of the patient allocation until after the randomisation process. " Quote (page 1118): "Analysis was blinded. .." Comment: we assumed this means outcome assessors were blinded
Incomplete outcome data (attrition and exclusions)	Low risk	Quote (page 1118): "Withdrawal was viewed as non-completion of questionnaires (primary outcome measures) - data

		on deaths were still collected. Reasons given for withdrawing were participant commitment (10), concern over the time involved (4), lack of understanding of the study (1) and one did not 'wish to be reminded of their illness'. The remaining patients gave no explanation, but the withdrawals were equally distributed between the groups. There were 76 patients in the GP group, and 81 in the surgical group after 24 months of follow-up, meeting the numbers required for statistical validity. Analysis was on an 'intention to treat' basis." Comment: reasons for withdrawal were reported by study arm and reasons were given
Selective reporting (reporting bias)	Unclear risk	**Outcomes in the methods** • SF-12 PCS and MCS • HADS • Number and type of investigations • Number and time to detraction of recurrence • Deaths (all causes) **Outcomes reported** • SF-12 PCS and MCS • HADS • PSVQ • Number and type of investigation • Number and timing of recurrence • Deaths all causes We did not have access to the protocol, so judged this outcome to be at unclear risk of bias
Other sources of bias	Unclear risk	We detected no other bias.

CA 19-9: cancer antigen 19-9.
CEA: carcinoembryonic antigen.
CIS: carcinoma in situ.
CRC: colorectal cancer.
CT: computerised tomography.
CXR: chest X-ray.
DFS: disease-free survival.
DRE: digital rectal examination.
EORTC QLQ-C30: European Organization for Research and Treatment of Cancer Quality of Life Questionnaire for Cancer patients.
EQ VAS: EuroQol visual analogue scale.
EuroQol-5D (EQ-5D): EuroQol five dimensions questionnaire.
FBC: full blood count.
F-U: follow-up.

GP: general practitioner.
HADS: Hospital Anxiety and Depression Scale.
HR: hazard ratio.
HRQoL: health-related quality of life.
ITT: intention-to-treat.
LFTs: liver function tests.
LR: local recurrence
MCS: mental component summary.
MRI: magnetic resonance imaging.
NHS: National Health Service.
OS: overall survival.
PCS: physical component summary.
PET: positron emission tomography.
PSVQ: Patient Visit-Specific Questionnaire.
RCT: randomised controlled trial.
SF-12: short form-12.
US: ultrasound.

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Barillari 1996	607 participants were enrolled in this trial; 212 were randomised, but the trial did not report data for these randomised participants separately
Kronborg 1981	This was a prospective, partly randomised trial. We were unable to extract the data relating to randomised participants from the paper
Sano 2004	The study was ineligible; participants had not had colorectal cancer

Characteristics of studies awaiting assessment *[ordered by study ID]*

Jefford 2013

Methods	Multicentre study Randomised controlled trial
Participants	Eligible participants have completed treatment for potentially cured CRC. Other eligibility criteria include stage I to III disease, age greater than 18 years, and adequate understanding of English
Interventions	SurvivorCare intervention (nurse-led survivorship care package) to usual post-treatment care, for patients with potentially cured CRC
Outcomes	• Psychological distress • Unmet needs • Quality of life

Jefford 2013 (*Continued*)

Notes	-

NCT00199654

Methods	This is a phase III open-labelled, multicentre, multidisciplinary, randomised study, comparing 2 arms of 188 participants (i.e. 376 total participants)
Participants	**Inclusion criteria** • Signed and dated written informed consent • Confirmed stage II or III (TNM) • Participant with curative surgery for colorectal adenocarcinoma • Total digestive endoscopy prior surgery or postsurgery • Age > 18 years old • Normal liver ultrasound and chest X-ray or thoraco-abdomino pelvic computed tomography (CT) scan • In fertile women, efficient contraception or postmenopausal participant (amenorrhoea for at least 1 year) **Exclusion criteria** • Serious concomitant pathology • Uncontrolled diabetes with a classical treatment (glycaemia > 1.4 g/l) • Other malignancy within the last 5 years (except for curatively treated basocellular carcinoma of the skin or in situ cervical carcinoma) • Uncontrolled infection • Women who are pregnant or lactating • Inability to understand informed consent • Psychological or geographic impossibility to follow up for 3 years
Interventions	• PET
Outcomes	• Evaluation of overall survival in the 2 groups • Evaluation of the rate of curative surgery • Comparison of the medical cost in the 2 detection strategies
Notes	Study start date: February 2004 Study completion date: April 2013 No study results are published yet.

Verberne 2015

Methods	Randomised, multicentred stepped wedge cluster
Participants	Participants treated surgically with curative intent for colorectal cancer
Interventions	Experimental arm: second monthly CEA (repeated at 1 month if increased by 20%), imaging performed after 2 consecutive rises
Outcomes	• Recurrences • Curable recurrences • Time to recurrence

Verberne 2015 (*Continued*)

Notes	Trial Registry Number 2182

CRC: colorectal cancer.
TNM: tumour, node, metastases.
CT: computerised tomography.
PET: positron emission tomography.
CEA: carcinoembryonic antigen.

Characteristics of ongoing studies *[ordered by study ID]*

COLOFOL

Trial name or title	A pragmatic study to assess the frequency of surveillance tests after curative resection in patients with stage II and III colorectal cancer - a randomised multicentre trial
Methods	RCT
Participants	• Radical surgery (R0 resection) for colorectal adenocarcinoma - with or without adjuvant treatment • Age < 75 years • "Clean colon" verified by perioperative barium enema or colonoscopy in the last 3 months post surgery • Stage II to III (T2, N1-2, M0, T3-4, any N, M0)(Edge 2010)
Interventions	Low-frequency follow-up group: CEA 1 month postoperatively then CEA; CT or MRI of the liver, or both; and X-ray/CT of the lungs 12 and 36 months after surgery High-frequency follow-up group: CEA 1 month postoperatively then CEA; CT or MRI of the liver or PET scans, or both; as well as X-ray or CT of the lungs at 6, 12, 18, 24, and 36 months
Outcomes	Primary outcomes will be total mortality and cancer-specific mortality at 5 years. Secondary outcome will be recurrence-free survival.
Starting date	22 September 2005
Contact information	Peer Wille-Jørgensen, MD, Dr Med Sci Department of Surgical Gastroenterology, KH:S - Bispebjerg Hospital DK-2400 Copenhagen Nordvest Denmark pwj01@bbh.hosp.dk
Notes	Study ID: NCT00225641 This study is active, but not recruiting.

NCT00995202

Trial name or title	Follow-up care with or without CEA assessments in patients who have undergone surgery for stage II or stage III colorectal cancer
Methods	RCT n = 1925
Participants	• Pathologically confirmed adenocarcinoma of the colon or rectum • Stage II or III disease • No distant metastatic disease • Has undergone curative resection for no residual tumour • Carcinoembryonic antigen (CEA) $\leq$ 1.5 x upper limit of normal after surgery • WHO performance status 0 to 1 • Not pregnant or nursing • Fertile participants must use effective contraception • No inflammatory bowel disease • No other malignancy within the past 5 years except basal cell carcinoma of the skin or carcinoma in situ of the cervix, or both • No genetic syndromes
Interventions	Other: diagnostic laboratory biomarker analysis Procedure: computed tomography Procedure: diagnostic colonoscopy Procedure: standard follow-up care Procedure: ultrasound imaging
Outcomes	• Overall survival • Disease-free survival • Curative resection rate in case of recurrence • 5-year overall survival rate • Cost-effectiveness study • Quality of life
Starting date	Study start date: September 2009 Estimated primary completion date: April 2018
Contact information	Come Lepage, MD Centre Hospitalier Universitaire de Dijon
Notes	This study is ongoing, but not recruiting participants.

NCT01628211

Trial name or title	Second look laparoscopy in colorectal cancer (HIPEC)
Methods	Randomised phase II
Participants	Participants who have had radical resection of mucinous colorectal cancer **Inclusion criteria** • Histologic diagnosis of colorectal adenocarcinoma

NCT01628211 (*Continued*)

	• Mucinous histotype • Stage I to III • Radical (R0) surgical resection of primary tumour • CT scan with contrast showing no evidence of disease recurrence 6 months after primary surgery • Age ≥ 18 ≤ 65 years • Performance status ECOG ≤1 • Normal hepatic, renal, and haematologic function • Adjuvant chemotherapy permitted • Signed informed consent **Exclusion criteria** • Residual disease after surgical resection of primary tumour • Distant metastasis • Active systemic infection • Chronic cardiovascular illness that would contraindicate abdominal dilatation with pneumoperitoneum • Concomitant or previous malignancy with 5 years of surgical resection of primary tumour (except for adequately treated non-melanoma skin cancer and in situ cervical cancer) • Pregnancy or lactation • Refusal or incapability of providing informed consent • Impossibility of complying with study schedules and follow-up
Interventions	Second-look laparoscopy, followed by peritonectomy, hyperthermic intraperitoneal chemotherapy (HIPEC) , or systemic chemotherapy in case of peritoneal carcinosis
Outcomes	**Primary outcome measures** • Overall survival (2 years) **Secondary outcome measures** • Number of participants with periotoneal carcinosis diagnosed at laparoscopy in the experimental arm (6 months) • Changes in quality of life (6 months) • Overall survival (5 years) • Worst grade adverse events per participant (7 months) • Number of participants with radiologic evidence of disease after initial surgery (6 months) • List of therapies and clinical outcomes of participants who had radiologic evidence of disease within 6 months after initial surgery
Starting date	Study start date: April 2012 Estimated study completion date: 2017
Contact information	Francesco Perrone, MD, PhD +39 081 5903571 francesco.perrone@usc-intnapoli.net Mariliana Piccirillo, MD +39 081 5903383 marilina.piccirolli@usc-intnapoli.net
Notes	This study is currently recruiting participants.

RCT: randomised controlled trial.
CEA: carcinoembryonic antigen.
CT: computerised tomography.
MRI: magnetic resonance imaging.
WHO: World Health Organization.
HIPEC: hyperthermic intraperitoneal chemotherapy.
ECOG: Eastern Cooperative Oncology Group.

DATA AND ANALYSES

Comparison 1. Intensive follow-up versus minimalist follow-up

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Overall survival	12		Hazard Ratio (Random, 95% CI)	0.90 [0.78, 1.02]
2 Colorectal cancer-specific survival	8		Hazard Ratio (Random, 95% CI)	0.93 [0.78, 1.12]
3 Relapse-free survival	13		Hazard Ratio (Random, 95% CI)	1.03 [0.90, 1.18]
4 Salvage surgery	13	5157	Risk Ratio (M-H, Random, 95% CI)	1.98 [1.53, 2.56]
5 Interval recurrences	7	3933	Risk Ratio (M-H, Random, 95% CI)	0.59 [0.41, 0.86]
6 Colonoscopy complications	1	1561	Risk Ratio (M-H, Fixed, 95% CI)	2.08 [0.11, 40.17]
7 OS SGA CEA versus NO CEA	4		Hazard Ratio (Random, 95% CI)	0.99 [0.81, 1.21]
7.1 CEA	3		Hazard Ratio (Random, 95% CI)	1.06 [0.78, 1.43]
7.2 No CEA	1		Hazard Ratio (Random, 95% CI)	0.90 [0.67, 1.21]
8 OS CT versus no CT	8		Hazard Ratio (Random, 95% CI)	0.92 [0.77, 1.09]
8.1 CT	5		Hazard Ratio (Random, 95% CI)	0.91 [0.74, 1.12]
8.2 No CT	3		Hazard Ratio (Random, 95% CI)	1.31 [0.40, 4.23]
9 OS CT versus < 2 or no CT	9		Hazard Ratio (Random, 95% CI)	0.87 [0.73, 1.05]
9.1 CT	6		Hazard Ratio (Random, 95% CI)	0.82 [0.65, 1.04]
9.2 < 2 or no CT	3		Hazard Ratio (Random, 95% CI)	1.04 [0.70, 1.54]
10 Overall survival SGA	12		Hazard Ratio (Random, 95% CI)	0.90 [0.78, 1.02]
10.1 Setting	2		Hazard Ratio (Random, 95% CI)	1.39 [0.38, 5.12]
10.2 More intensive versus less intensive	10		Hazard Ratio (Random, 95% CI)	0.89 [0.78, 1.02]
11 Overall survival SGA “dose” of follow-up	9		Hazard Ratio (Random, 95% CI)	0.86 [0.74, 1.00]
11.1 More visits and tests versus fewer visits and tests	6		Hazard Ratio (Random, 95% CI)	0.82 [0.68, 0.99]
11.2 VIsits and tests versus minimal or no follow-up	3		Hazard Ratio (Random, 95% CI)	0.90 [0.64, 1.26]

Analysis 1.1. Comparison 1 Intensive follow-up versus minimalist follow-up, Outcome 1 Overall survival.

Review: Follow-up strategies for patients treated for non-metastatic colorectal cancer

Comparison: 1 Intensive follow-up versus minimalist follow-up

Outcome: 1 Overall survival

Study or subgroup	log [Hazard Ratio] (SE)	Hazard Ratio IV,Random,95% CI	Weight	Hazard Ratio IV,Random,95% CI
Augestad 2013	1.24 (0.91)		0.6 %	3.46 [0.58, 20.56]
FACS 2014	0.16 (0.17)		14.7 %	1.17 [0.84, 1.64]
GILDA 1998	0 (0.14)		21.0 %	1.00 [0.76, 1.32]
Kjeldsen 1997	-0.1 (0.15)		18.6 %	0.90 [0.67, 1.21]
Mäkelä 1995	-0.16 (0.33)		4.1 %	0.85 [0.45, 1.63]
Ohlsson 1995	-0.38 (0.33)		4.1 %	0.68 [0.36, 1.31]
Pietra 1998	-0.56 (0.24)		7.7 %	0.57 [0.36, 0.91]
Rodr guez-Moranta 2006	-0.23 (0.29)		5.3 %	0.79 [0.45, 1.40]
Schoemaker 1998	-0.26 (0.2)		10.9 %	0.77 [0.52, 1.14]
Strand 2011	-0.16 (0.48)		2.0 %	0.85 [0.33, 2.18]
Treasure 2014	0.19 (0.36)		3.5 %	1.21 [0.60, 2.45]
Wang 2009	-0.27 (0.24)		7.7 %	0.76 [0.48, 1.22]
Total (95% CI)			**100.0 %**	**0.90 [0.78, 1.02]**

Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 11.44$, df = 11 (P = 0.41); $I^2 = 4\%$

Test for overall effect: Z = 1.64 (P = 0.10)

Test for subgroup differences: Not applicable

Analysis 1.2. Comparison 1 Intensive follow-up versus minimalist follow-up, Outcome 2 Colorectal cancer-specific survival.

Review: Follow-up strategies for patients treated for non-metastatic colorectal cancer

Comparison: 1 Intensive follow-up versus minimalist follow-up

Outcome: 2 Colorectal cancer-specific survival

Study or subgroup	log [Hazard Ratio] (SE)	Hazard Ratio IV,Random,95% CI	Weight	Hazard Ratio IV,Random,95% CI
Augestad 2013	1.24 (0.91)		1.1 %	3.46 [0.58, 20.56]
FACS 2014	0.12 (0.22)		18.1 %	1.13 [0.73, 1.74]
GILDA 1998	0.15 (0.3)		9.8 %	1.16 [0.65, 2.09]
Kjeldsen 1997	-0.02 (0.18)		27.1 %	0.98 [0.69, 1.39]
Ohlsson 1995	-0.31 (0.38)		6.1 %	0.73 [0.35, 1.54]
Pietra 1998	-0.44 (0.23)		16.6 %	0.64 [0.41, 1.01]
Rodr guez-Moranta 2006	-0.26 (0.35)		7.2 %	0.77 [0.39, 1.53]
Wang 2009	-0.015 (0.25)		14.1 %	0.99 [0.60, 1.61]
Total (95% CI)			**100.0 %**	**0.93 [0.78, 1.12]**

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 6.76$, df = 7 (P = 0.45); $I^2 = 0.0\%$

Test for overall effect: Z = 0.73 (P = 0.46)

Test for subgroup differences: Not applicable

Analysis 1.3. Comparison 1 Intensive follow-up versus minimalist follow-up, Outcome 3 Relapse-free survival.

Review: Follow-up strategies for patients treated for non-metastatic colorectal cancer

Comparison: 1 Intensive follow-up versus minimalist follow-up

Outcome: 3 Relapse-free survival

Study or subgroup	log [Hazard Ratio] (SE)	Hazard Ratio IV,Random,95% CI	Weight	Hazard Ratio IV,Random,95% CI
Augestad 2013	0.32 (0.54)		1.6 %	1.38 [0.48, 3.97]
FACS 2014	0.41 (0.18)		13.4 %	1.51 [1.06, 2.14]
GILDA 1998	0.2 (0.13)		23.6 %	1.22 [0.95, 1.58]
Kjeldsen 1997	-0.04 (0.16)		16.6 %	0.96 [0.70, 1.31]
Mäkelä 1995	0.14 (0.39)		3.1 %	1.15 [0.54, 2.47]
Ohlsson 1995	-0.06 (0.41)		2.8 %	0.94 [0.42, 2.10]
Pietra 1998	-0.02 (0.28)		5.9 %	0.98 [0.57, 1.70]
Rodr guez-Moranta 2006	0.09 (0.28)		5.9 %	1.09 [0.63, 1.89]
Schoemaker 1998	-0.3 (0.23)		8.6 %	0.74 [0.47, 1.16]
Secco 2002	-0.2 (0.22)		9.3 %	0.82 [0.53, 1.26]
Strand 2011	-0.2 (0.56)		1.5 %	0.82 [0.27, 2.45]
Treasure 2014	-0.34 (0.34)		4.1 %	0.71 [0.37, 1.39]
Wang 2009	-0.39 (0.36)		3.6 %	0.68 [0.33, 1.37]
Total (95% CI)			**100.0 %**	**1.03 [0.90, 1.18]**

Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 12.70$, df = 12 (P = 0.39); $I^2 = 5\%$

Test for overall effect: Z = 0.47 (P = 0.64)

Test for subgroup differences: Not applicable

Analysis 1.4. Comparison 1 Intensive follow-up versus minimalist follow-up, Outcome 4 Salvage surgery.

Review: Follow-up strategies for patients treated for non-metastatic colorectal cancer

Comparison: 1 Intensive follow-up versus minimalist follow-up

Outcome: 4 Salvage surgery

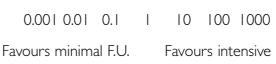
Study or subgroup	Intensive n/N	Minimal n/N	Risk Ratio M-H,Random,95% CI	Weight	Risk Ratio M-H,Random,95% CI
Augestad 2013	4/55	3/55		2.9 %	1.33 [0.31, 5.68]
FACS 2014	64/901	7/301		8.1 %	3.05 [1.42, 6.59]
GILDA 1998	57/622	46/620		18.0 %	1.24 [0.85, 1.79]
Kjeldsen 1997	21/290	7/307		7.1 %	3.18 [1.37, 7.36]
Mäkelä 1995	5/52	3/54		3.1 %	1.73 [0.44, 6.88]
Ohlsson 1995	5/53	3/54		3.1 %	1.70 [0.43, 6.75]
Pietra 1998	21/104	6/103		6.8 %	3.47 [1.46, 8.24]
Rodr guez-Moranta 2006	18/127	10/132		8.7 %	1.87 [0.90, 3.90]
Schoemaker 1998	6/167	5/158		4.2 %	1.14 [0.35, 3.65]
Secco 2002	31/192	13/145		11.0 %	1.80 [0.98, 3.32]
Sobhani 2008	15/65	2/65		2.9 %	7.50 [1.79, 31.49]
Treasure 2014	62/108	26/108		18.0 %	2.38 [1.64, 3.46]
Wang 2009	9/161	8/158		6.1 %	1.10 [0.44, 2.79]
Total (95% CI)	**2897**	**2260**		**100.0 %**	**1.98 [1.53, 2.56]**

Total events: 318 (Intensive), 139 (Minimal)

Heterogeneity: $Tau^2 = 0.06$; $Chi^2 = 17.30$, df = 12 (P = 0.14); $I^2 = 31\%$

Test for overall effect: Z = 5.18 (P < 0.00001)

Test for subgroup differences: Not applicable



Analysis 1.5. Comparison 1 Intensive follow-up versus minimalist follow-up, Outcome 5 Interval recurrences.

Review: Follow-up strategies for patients treated for non-metastatic colorectal cancer

Comparison: 1 Intensive follow-up versus minimalist follow-up

Outcome: 5 Interval recurrences

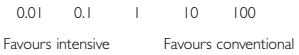
Study or subgroup	Intensive n/N	Conventional n/N	Risk Ratio M-H,Random,95% CI	Weight	Risk Ratio M-H,Random,95% CI
FACS 2014	41/901	28/301		18.3 %	0.49 [0.31, 0.78]
GILDA 1998	42/622	35/620		18.9 %	1.20 [0.77, 1.85]
Kjeldsen 1997	38/290	59/307		20.2 %	0.68 [0.47, 0.99]
Mäkelä 1995	6/52	8/54		9.2 %	0.78 [0.29, 2.09]
Secco 2002	33/192	62/145		20.5 %	0.40 [0.28, 0.58]
Sobhani 2008	3/65	7/65		6.2 %	0.43 [0.12, 1.59]
Wang 2009	3/161	11/158		6.6 %	0.27 [0.08, 0.94]
Total (95% CI)	**2283**	**1650**		**100.0 %**	**0.59 [0.41, 0.86]**

Total events: 166 (Intensive), 210 (Conventional)

Heterogeneity: $Tau^2 = 0.15$; $Chi^2 = 17.70$, df = 6 (P = 0.01); $I^2 = 66\%$

Test for overall effect: Z = 2.73 (P = 0.0063)

Test for subgroup differences: Not applicable



Analysis 1.6. Comparison 1 Intensive follow-up versus minimalist follow-up, Outcome 6 Colonoscopy complications.

Review: Follow-up strategies for patients treated for non-metastatic colorectal cancer

Comparison: 1 Intensive follow-up versus minimalist follow-up

Outcome: 6 Colonoscopy complications

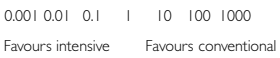
Study or subgroup	Intensive n/N	Conventional n/N	Risk Ratio M-H,Fixed,95% CI	Weight	Risk Ratio M-H,Fixed,95% CI
Wang 2009	3/1204	0/357		100.0 %	2.08 [0.11, 40.17]
Total (95% CI)	**1204**	**357**		**100.0 %**	**2.08 [0.11, 40.17]**

Total events: 3 (Intensive), 0 (Conventional)

Heterogeneity: not applicable

Test for overall effect: Z = 0.48 (P = 0.63)

Test for subgroup differences: Not applicable



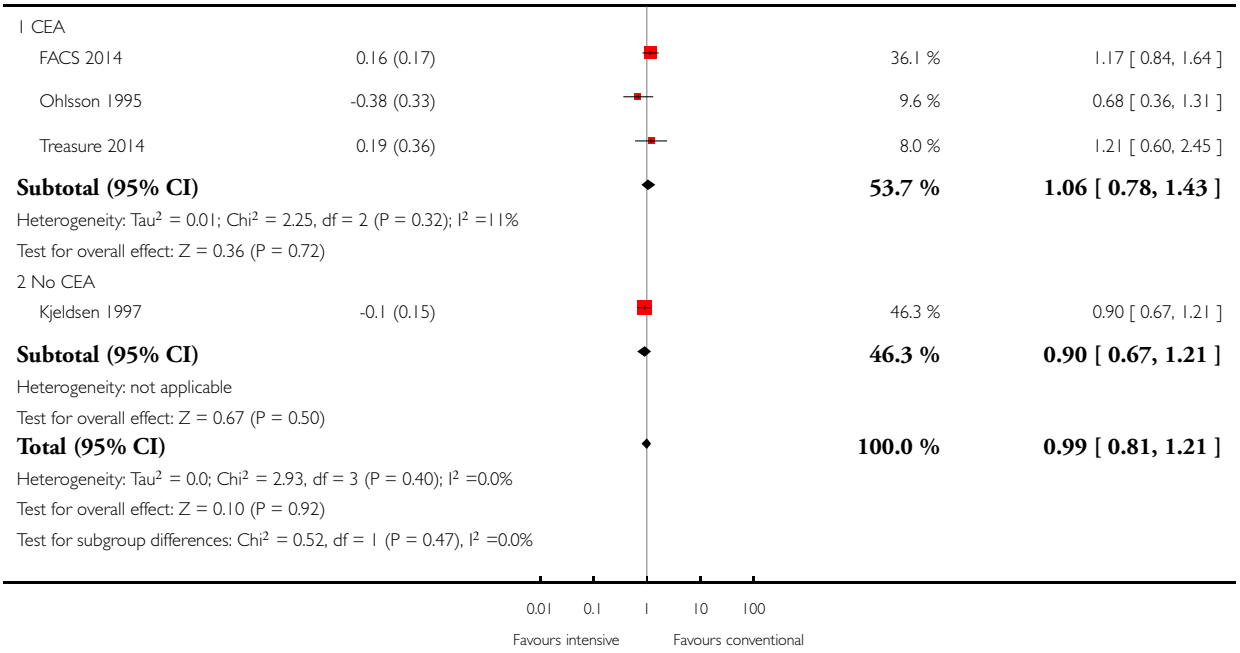
Analysis 1.7. Comparison 1 Intensive follow-up versus minimalist follow-up, Outcome 7 OS SGA CEA versus NO CEA.

Review: Follow-up strategies for patients treated for non-metastatic colorectal cancer

Comparison: 1 Intensive follow-up versus minimalist follow-up

Outcome: 7 OS SGA CEA versus NO CEA

Study or subgroup	log [Hazard Ratio] (SE)	Hazard Ratio IV,Random,95% CI	Weight	Hazard Ratio IV,Random,95% CI
1 CEA				
FACS 2014	0.16 (0.17)		36.1 %	1.17 [0.84, 1.64]
Ohlsson 1995	-0.38 (0.33)		9.6 %	0.68 [0.36, 1.31]
Treasure 2014	0.19 (0.36)		8.0 %	1.21 [0.60, 2.45]
Subtotal (95% CI)			**53.7 %**	**1.06 [0.78, 1.43]**
Heterogeneity: Tau^2 = 0.01; Chi^2 = 2.25, df = 2 (P = 0.32); I^2 =11%				
Test for overall effect: Z = 0.36 (P = 0.72)				
2 No CEA				
Kjeldsen 1997	-0.1 (0.15)		46.3 %	0.90 [0.67, 1.21]
Subtotal (95% CI)			**46.3 %**	**0.90 [0.67, 1.21]**
Heterogeneity: not applicable				
Test for overall effect: Z = 0.67 (P = 0.50)				
Total (95% CI)			**100.0 %**	**0.99 [0.81, 1.21]**
Heterogeneity: Tau^2 = 0.0; Chi^2 = 2.93, df = 3 (P = 0.40); I^2 =0.0%				
Test for overall effect: Z = 0.10 (P = 0.92)				
Test for subgroup differences: Chi^2 = 0.52, df = 1 (P = 0.47), I^2 =0.0%				

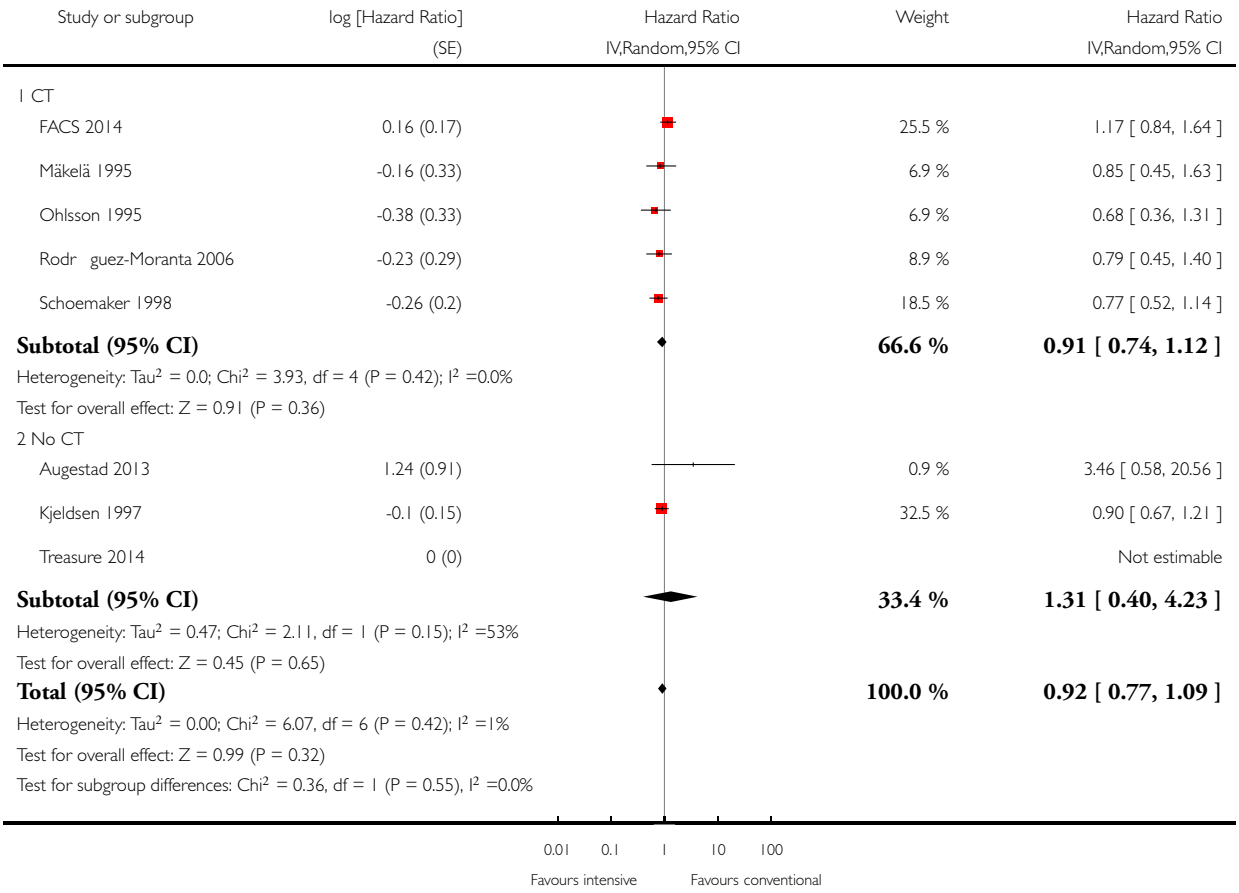


Analysis 1.8. Comparison 1 Intensive follow-up versus minimalist follow-up, Outcome 8 OS CT versus no CT.

Review: Follow-up strategies for patients treated for non-metastatic colorectal cancer

Comparison: 1 Intensive follow-up versus minimalist follow-up

Outcome: 8 OS CT versus no CT

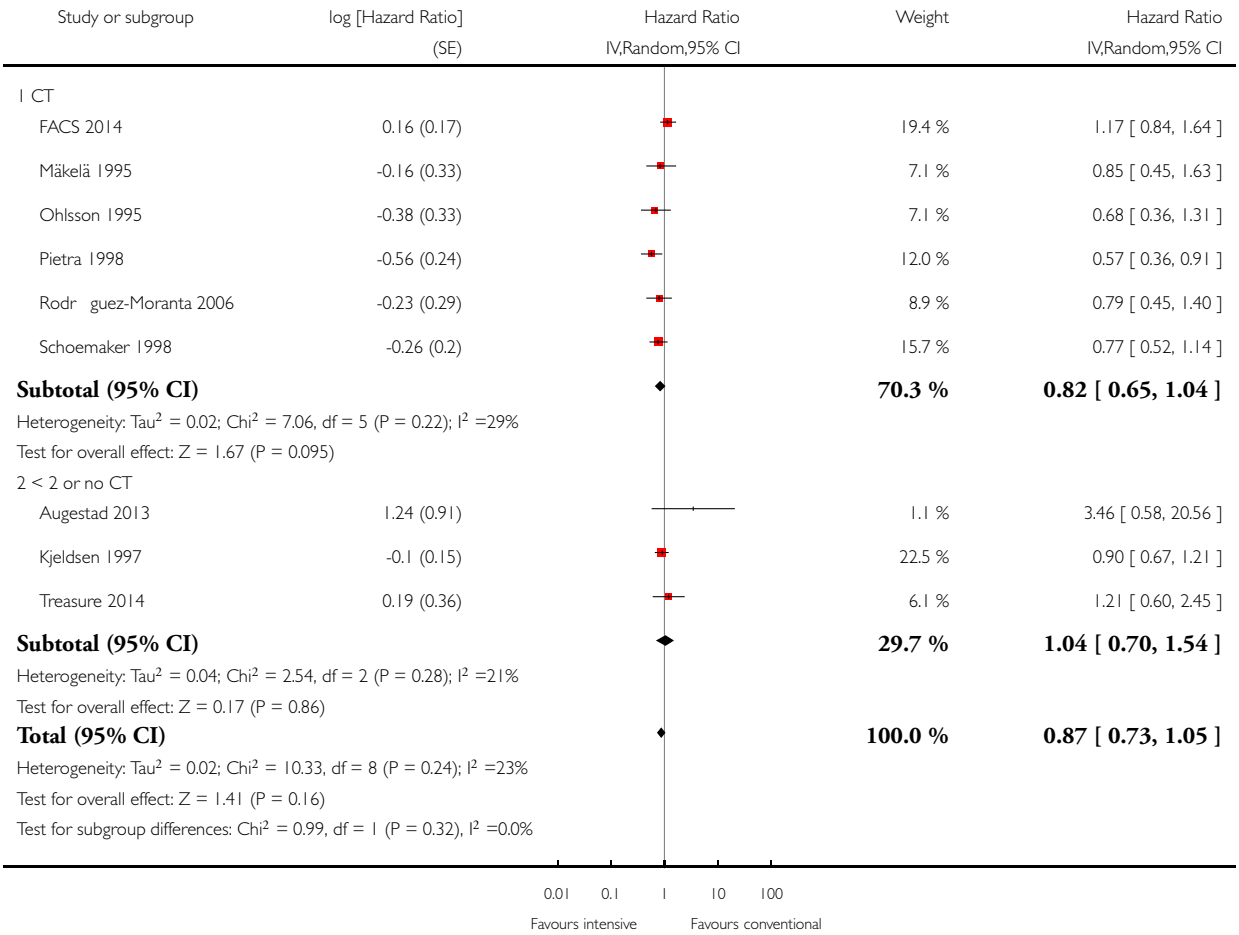


Analysis 1.9. Comparison 1 Intensive follow-up versus minimalist follow-up, Outcome 9 OS CT versus < 2 or no CT.

Review: Follow-up strategies for patients treated for non-metastatic colorectal cancer

Comparison: 1 Intensive follow-up versus minimalist follow-up

Outcome: 9 OS CT versus < 2 or no CT

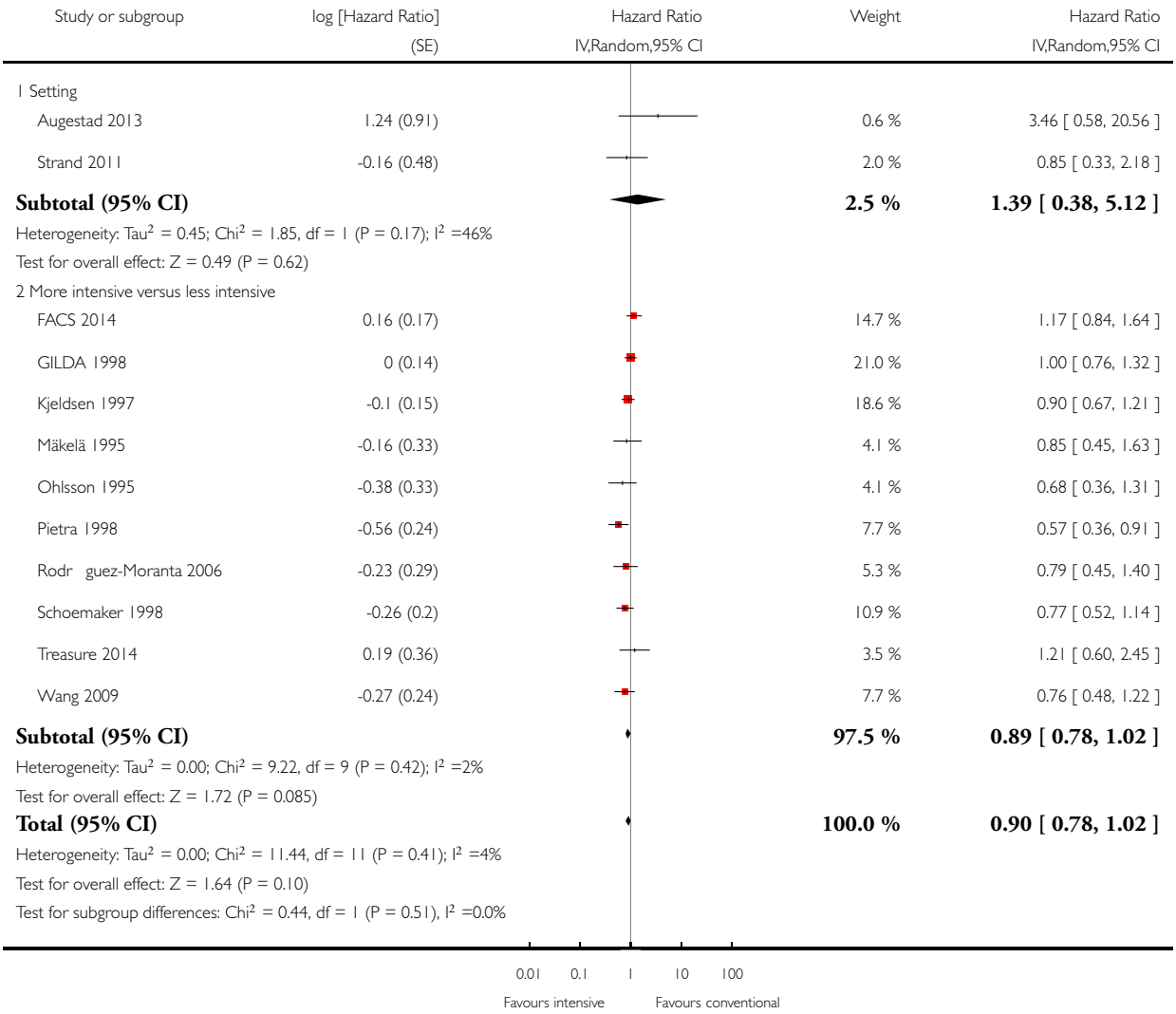


Analysis 1.10. Comparison 1 Intensive follow-up versus minimalist follow-up, Outcome 10 Overall survival SGA.

Review: Follow-up strategies for patients treated for non-metastatic colorectal cancer

Comparison: 1 Intensive follow-up versus minimalist follow-up

Outcome: 10 Overall survival SGA

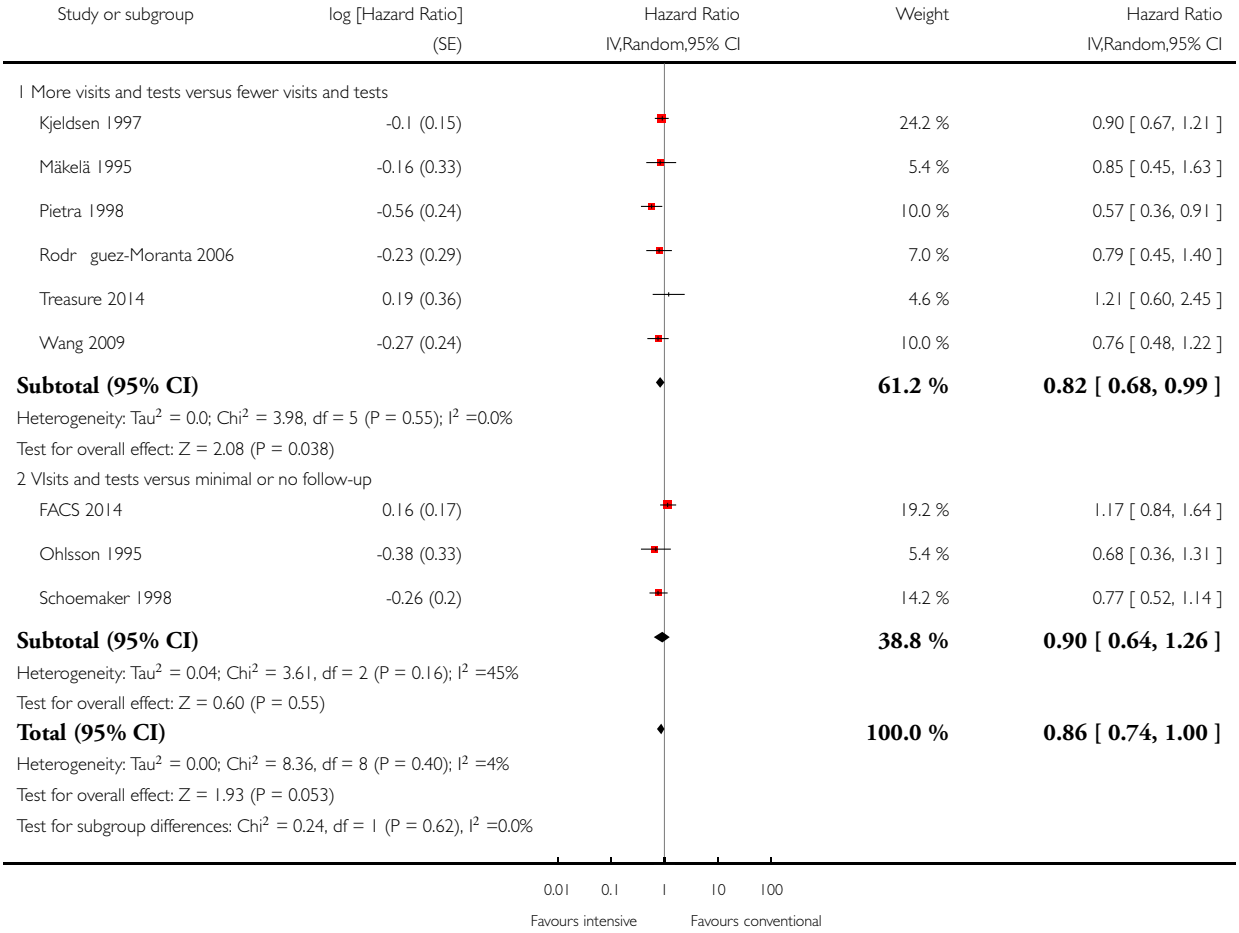


Analysis 1.11. Comparison 1 Intensive follow-up versus minimalist follow-up, Outcome 11 Overall survival SGA "dose" of follow-up.

Review: Follow-up strategies for patients treated for non-metastatic colorectal cancer

Comparison: 1 Intensive follow-up versus minimalist follow-up

Outcome: 11 Overall survival SGA "dose" of follow-up



ADDITIONAL TABLES

Table 1. Systematic reviews on topic

Study	Number of studies included	Search date	Outcomes for OS for comparison intensive versus less intensive follow-up for participants treated with curative intent for CRC
Pita-Fernández 2014	11	June 2014	HR 0.75 (95% CI 0.66 to 0.86)
Augestad 2014	Unclear	No search date given	No quantitative meta-analysis was presented.
Baca 2011	15	January 2000 to 2001	No quantitative meta-analysis was presented.
Tjandra 2007	8	June 2007	OR 0.74 (95% CI 0.59 to 0.93)
Renehan 2002	5	April 2001	RR 0.81 (95% CI 0.70 to 0.94)

OS: overall survival.
HR: hazard ratio.
CI: confidence interval.
OR: odds ratio.
RR: risk ratio.

APPENDICES

Appendix 1. CENTRAL search strategy

Cochrane Central
Database: EBM Reviews - Cochrane Central Register of Controlled Trials <April 2016>
Search Strategy:

---

1 (colorectal or colon$ or rectal or rectum or sigmoid).ti,ab,sh. (24231)
2 (cancer$ or neoplasm$ or or tumor or carcinoma or adenoma or adenocarcinoma).ti,ab,sh. (81443)
3 1 and 2 (10445)
4 random$.af. (713326)
5 double blind.mp. (184812)
6 single blind.mp. (24374)
7 or/4-6 (739346)
8 recurr$.ti,ab,sh. (31008)
9 metastas$.ti,ab,sh. (6552)
10 8 or 9 (36060)
11 3 and 7 and 10 (1697)
12 limit 11 to yr=2006-2016 (906)
13 (follow-up or follow up).ti,ab,sh. (88787)
14 longitudinal.ti,ab,sh. (7264)
15 survival.ti,ab,sh. (38629)

16 mortality.ti,ab,sh. (28096)
17 prognosis.ti,ab,sh. (16514)
18 quality of life.ti,ab,sh. (34855)
19 Treatment Outcome/ (97141)
20 (treatment adj3 outcome).ti,ab. (8454)
21 or/13-20 (232471)
22 12 and 21 (758)

Appendix 2. MEDLINE search strategy

MEDLINE
Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations <May 20, 2016>, Ovid MEDLINE(R) <1996 to Present with Daily Update>
Search Strategy:
---
1 exp colorectal neoplasms/ (116407)
2 randomized controlled trial.pt. (322284)
3 random allocation/ (52293)
4 random$.af. (854660)
5 double blind method/ (91088)
6 single blind method/ (19369)
7 controlled clinical trial.pt. (45443)
8 or/2-7 (905129)
9 recurrence/ (97196)
10 neoplasm recurrence, local/ (64834)
11 neoplasm metastasis/ (43469)
12 recurr$.ti,ab,sh. (376234)
13 or/9-12 (431795)
14 follow up studies/ (376887)
15 follow up.ti,ab,sh. (577775)
16 exp longitudinal studies/ (82818)
17 exp survival analysis/ (204612)
18 exp mortality/ (243407)
19 exp prognosis/ (1064148)
20 office visits/ (4839)
21 "Episode of Care"/ (1387)
22 exp population surveillance/ (48291)
23 Practice Patterns, Physicians'/ (42538)
24 exp treatment outcome/ (725188)
25 "Outcome Assessment (Health Care)"/ (51554)
26 quality of life.mp. or "Quality of Life"/ (214742)
27 or/14-26 (2025787)
28 1 and 8 and 13 (2176)
29 1 and 8 and 27 (5301)
30 28 or 29 (5809)
31 limit 30 to humans (5736)
32 limit 31 to yr="2006 -Current" (3795)

Appendix 3. EMBASE search stategy

Embase
Database: Embase <1980 to 2016 May 23>, Embase Classic <1947 to 1979>
Search Strategy:

---

1 colon tumor/ (24596)
2 colon cancer/ (54320)
3 colon carcinoma/ (21376)
4 colon adenocarcinoma/ (8481)
5 colorectal tumor/ (18557)
6 sigmoid carcinoma/ (926)
7 rectum carcinoma/ (12139)
8 rectum cancer/ (25073)
9 rectum tumor/ (16280)
10 rectum adenoma/ (1907)
11 colorectal carcinoma/ (19399)
12 colorectal cancer/ (98996)
13 or/1-12 (262185)
14 randomization/ (70575)
15 randomized controlled trial/ (404026)
16 double blind procedure/ (133057)
17 single blind procedure/ (22126)
18 random$.af. (1263540)
19 or/14-18 (1292538)
20 metastasis/ (258705)
21 cancer recurrence/ (96706)
22 tumor recurrence/ (45501)
23 recurrent disease/ (145450)
24 (recur$ or metastas$).ti,ab. (975672)
25 or/20-24 (1161930)
26 13 and 19 and 25 (5349)
27 (rat or rats or mouse or mice).ti,ab,sh. (3468071)
28 (monkey$ or rabbit$ or hamster$).ti,ab,sh. (639841)
29 (bovine or sheep).ti,ab,sh. (331074)
30 animal/ or experimental animal/ (1787198)
31 or/27-30 (5244594)
32 26 not 31 (5113)
33 longitudinal study/ (88575)
34 follow up/ (1057646)
35 (follow-up or follow up).ti,ab. (1064657)
36 prospective study/ (334785)
37 treatment outcome/ (720484)
38 cancer survival/ (205796)
39 quality of life/ (316649)
40 prognosis/ (516100)
41 mortality/ (662941)
42 morbidity/ (273023)
43 exp survival/ (803758)
44 or/33-43 (3736259)
45 13 and 19 and 44 (10019)
46 45 not 31 (9818)
47 46 not 32 (5790)

48 case report/ (2137256)
49 letter/ or letter.pt. (941176)
50 48 or 49 (2894625)
51 32 or 47 (10903)
52 51 not 50 (10718)
53 limit 52 to yr=2006-2016 (7843)

Appendix 4. CINAHL search strategy

CINAHL- EBSCOhost
Note: search has been rekeyed from the original printout to enhance readability

1	MH "Rectal Neoplasms+	1,802
2	MH Colonic Neoplasms+	4,298
3	MH Colorectal Neoplasms+	15,461
4	S1 OR S2 OR S3	15,461
5	"follow up"	80,974
6	MH Recurrence OR recur*	44,486
7	MH Neoplasm recurrence, local	5,399
8	MH Prospective Studies OR longitudinal	194,415
9	S5 OR S6 OR S7 OR S8	274,290
10	S4 AND S9	2,714
11	MW meta-analysis	17,460
12	TX meta analy* OR metaanaly*	27,334
13	TX Cochrane*	22,935
14	PT nursing interventions	1,379
15	MH literature review	3,584
16	MH literature searching	835
17	MH computerized literature searching	4,940
18	MH reference databases	1,726
19	TX review* OR overview*	2,887,898

(*Continued*)

20	TX pooled data OR pooled analy*	1,951
21	PT review	106,455
22	TX systematic* OR methodologic* OR quantitative OR research* OR literature OR studies OR trial* OR effective*	1,683,194
23	S21 AND S22	65,969
24	TX (synthesis* AND (literature OR studies OR data))	11,542
25	TX ((hand OR manual*) AND search*)	5,121
26	TX ((electronic* OR bibliography*) AND (database* OR data base*))	12,971
27	S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20	2,890,087
28	S23 OR S24 OR S25 OR s26	90,915
29	S27 or S28	2,891,877
30	S10 and S29	2,657
31	MW randomized	26,597
32	MW random	60,954
33	TX random*	177,131
34	S31 OR S32 OR S33	177,131
35	S10 AND S34	562
36	S30 OR S35	2,664
37	PT letter OR case study	325,844
38	S36 NOT S37	2,479
39	Limiters Published Date: 20060101-20161231	1,916
40		

Appendix 5. Science Citation Index search strategy

Science Citation Index and Conference Abstracts

# 15	3,292	#7 AND #12 *Indexes=SCI-EX-PANDED, CPCI-S, ESCI* *Timespan=2006-2016*			
# 14	11,791	#7 OR #13 *Indexes=SCI-EX-PANDED, CPCI-S, ESCI* *Timespan=2006-2016*			
# 13	11,050	#5 AND #12 *Indexes=SCI-EX-PANDED, CPCI-S, ESCI* *Timespan=2006-2016*			
# 12	2,049,980	#11 OR #10 OR #9 OR # 8 *Indexes=SCI-EX-PANDED, CPCI-S, ESCI* *Timespan=2006-2016*			
# 11	163,800	TS="quality of life" *Indexes=SCI-EX-PANDED, CPCI-S, ESCI* *Timespan=2006-2016*			
# 10	955,659	TS=(prognosis or outcome) *Indexes=SCI-EX-PANDED, CPCI-S, ESCI* *Timespan=2006-2016*			
# 9	877,979	TS=(survival or mortality or morbidity) *Indexes=SCI-EX-*			

(*Continued*)

		PANDED, CPCI-S, ESCI *Timespan=2006-2016*			
# 8	661,485	TS=(follow up or follow-up or longitudinal) *Indexes=SCI-EX-PANDED, CPCI-S, ESCI* *Timespan=2006-2016*			
# 7	4,033	#5 AND #6 *Indexes=SCI-EX-PANDED, CPCI-S, ESCI* *Timespan=2006-2016*			
# 6	404,566	TS=(recur* or metastas*) *Indexes=SCI-EX-PANDED, CPCI-S, ESCI* *Timespan=2006-2016*			
# 5	17,480	#3 AND #4 *Indexes=SCI-EX-PANDED, CPCI-S, ESCI* *Timespan=2006-2016*			
# 4	849,999	TS=(random* or double blind* or single blind*) *Indexes=SCI-EX-PANDED, CPCI-S, ESCI* *Timespan=2006-2016*			
# 3	166,220	#1 AND #2 *Indexes=SCI-EX-PANDED, CPCI-S, ESCI* *Timespan=2006-2016*			

(*Continued*)

# 2	1,263,394	TS=(cancer or tumor* or * or carcinoma* or adenoma* or adenocarcinoma*) *Indexes=SCI-EXPANDED, CPCI-S, ESCI* *Timespan=2006-2016*			

# 1	364,015	TS=(colorectal or colon* or rectal or rectum or sigmoid)			

Appendix 6. Criteria for judging risk of bias in the 'Risk of bias' assessment tool

RANDOM SEQUENCE GENERATION **Selection bias (biased allocation to interventions) due to inadequate generation of a randomised sequence**	
Criteria for a judgement of 'low risk' of bias	The investigators describe a random component in the sequence generation process such as: • referring to a random number table; • using a computer random number generator; • coin tossing; • shuffling cards or envelopes; • throwing dice; • drawing of lots; or • minimisation*. *Minimisation may be implemented without a random element, and this is considered to be equivalent to being random
Criteria for the judgement of 'high risk' of bias	The investigators describe a non-random component in the sequence generation process. Usually, the description would involve some systematic, non-random approach, for example: • sequence generated by odd or even date of birth; • sequence generated by some rule based on date (or day) of admission; or • sequence generated by some rule based on hospital or clinic record number. Other non-random approaches happen much less frequently than the systematic approaches mentioned above and tend to be obvious. They usually involve judgement or some method of non-

(*Continued*)

	random categorisation of participants, for example: • allocation by judgement of the clinician; • allocation by preference of the participant; • allocation based on the results of a laboratory test or a series of tests; or • allocation by availability of the intervention.
Criteria for the judgement of 'unclear risk' of bias	Insufficient information about the sequence generation process to permit judgement of 'low risk' or 'high risk'
ALLOCATION CONCEALMENT **Selection bias (biased allocation to interventions) due to inadequate concealment of allocations prior to assignment**	
Criteria for a judgement of 'low risk' of bias	Participants and investigators enrolling participants could not foresee assignment because 1 of the following, or an equivalent method, was used to conceal allocation: • central allocation (including telephone, web-based, and pharmacy-controlled randomisation); • sequentially numbered drug containers of identical appearance; or • sequentially numbered, opaque, sealed envelopes.
Criteria for the judgement of 'high risk' of bias	Participants or investigators enrolling participants could possibly foresee assignments and thus introduce selection bias, such as allocation based on the following: • using an open random allocation schedule (e.g. a list of random numbers); • assignment envelopes were used without appropriate safeguards (e.g. if envelopes were unsealed or nonopaque or not sequentially numbered); • alternation or rotation; • date of birth; • case record number; or • any other explicitly unconcealed procedure.
Criteria for the judgement of 'unclear risk' of bias	Insufficient information to permit judgement of 'low risk' or 'high risk'. This is usually the case if the method of concealment is not described or not described in sufficient detail to allow a definite judgement - for example, if the use of assignment envelopes is described, but it remains unclear whether envelopes were sequentially numbered, opaque, and sealed
BLINDING OF PARTICIPANTS AND PERSONNEL **Performance bias due to knowledge of the allocated interventions by participants and personnel during the study**	
Criteria for a judgement of 'low risk' of bias	Any 1 of the following: • no blinding or incomplete blinding, but the review authors judge that the outcome is not likely to be influenced by lack of blinding; or

(*Continued*)

	• blinding of participants and key study personnel ensured, and unlikely that the blinding could have been broken.
Criteria for the judgement of 'high risk' of bias	Any 1 of the following: • no blinding or incomplete blinding, and the outcome is likely to be influenced by lack of blinding; or • blinding of key study participants and personnel attempted, but likely that the blinding could have been broken, and the outcome is likely to be influenced by lack of blinding.
Criteria for the judgement of 'unclear risk' of bias	Any 1 of the following: • insufficient information to permit judgement of 'low risk' or 'high risk'; or • the study did not address this outcome.
BLINDING OF OUTCOME ASSESSMENT **Detection bias due to knowledge of the allocated interventions by outcome assessors**	
Criteria for a judgement of 'low risk' of bias	Any 1 of the following: • no blinding of outcome assessment, but the review authors judge that the outcome measurement is not likely to be influenced by lack of blinding; or • blinding of outcome assessment ensured, and unlikely that the blinding could have been broken.
Criteria for the judgement of 'high risk' of bias	Any 1 of the following: • no blinding of outcome assessment, and the outcome measurement is likely to be influenced by lack of blinding; or • blinding of outcome assessment, but likely that the blinding could have been broken, and the outcome measurement is likely to be influenced by lack of blinding.
Criteria for the judgement of 'unclear risk' of bias	Any 1 of the following: • insufficient information to permit judgement of 'low risk' or 'high risk'; or • the study did not address this outcome.
INCOMPLETE OUTCOME DATA **Attrition bias due to amount, nature, or handling of incomplete outcome data**	
Criteria for a judgement of 'low risk' of bias	Any 1 of the following: • no missing outcome data; • reasons for missing outcome data unlikely to be related to true outcome (for survival data, censoring unlikely to be introducing bias); • missing outcome data balanced in numbers across intervention groups, with similar reasons for missing data across groups; • for dichotomous outcome data, the proportion of missing

(*Continued*)

	outcomes compared with observed event risk not enough to have a clinically relevant impact on the intervention effect estimate; • for continuous outcome data, plausible effect size (difference in means or standardised difference in means) among missing outcomes not enough to have a clinically relevant impact on observed effect size; or • missing data have been imputed using appropriate methods.
Criteria for the judgement of 'high risk' of bias	Any 1 of the following: • reason for missing outcome data likely to be related to true outcome, with either imbalance in numbers or reasons for missing data across intervention groups; • for dichotomous outcome data, the proportion of missing outcomes compared with observed event risk enough to induce clinically relevant bias in intervention effect estimate; • for continuous outcome data, plausible effect size (difference in means or standardized difference in means) among missing outcomes enough to induce clinically relevant bias in observed effect size; • 'as-treated' analysis done with substantial departure of the intervention received from that assigned at randomisation; or • potentially inappropriate application of simple imputation.
Criteria for the judgement of 'unclear risk' of bias	Any 1 of the following: • insufficient reporting of attrition/exclusions to permit judgement of 'low risk' or 'high risk' (e.g. number randomised not stated, no reasons for missing data provided); or • the study did not address this outcome.
SELECTIVE REPORTING **Reporting bias due to selective outcome reporting**	
Criteria for a judgement of 'low risk' of bias	Any of the following: • the study protocol is available and all of the study's prespecified (primary and secondary) outcomes that are of interest in the review have been reported in the prespecified way; or • the study protocol is not available, but it is clear that the published reports include all expected outcomes, including those that were prespecified (convincing text of this nature may be uncommon).
Criteria for the judgement of 'high risk' of bias	Any 1 of the following: • not all of the study's prespecified primary outcomes have been reported; • 1 or more primary outcomes is reported using measurements, analysis methods, or subsets of the data (e.g. subscales) that were not prespecified;

(*Continued*)

	• 1 or more reported primary outcomes were not prespecified (unless clear justification for their reporting is provided, such as an unexpected adverse effect); • 1 or more outcomes of interest in the review are reported incompletely so that they cannot be entered in a meta-analysis; or • the study report fails to include results for a key outcome that would be expected to have been reported for such a study.
Criteria for the judgement of 'unclear risk' of bias	Insufficient information to permit judgement of 'low risk' or 'high risk'. It is likely that the majority of studies will fall into this category
OTHER BIAS **Bias due to problems not covered elsewhere in the table**	
Criteria for a judgement of 'low risk' of bias	The study appears to be free of other sources of bias.
Criteria for the judgement of 'high risk' of bias	There is at least 1 important risk of bias. For example, the study: • had a potential source of bias related to the specific study design used; • has been claimed to have been fraudulent; or • had some other problem.
Criteria for the judgement of 'unclear risk' of bias	There may be a risk of bias, but there is either: • insufficient information to assess whether an important risk of bias exists; or • insufficient rationale or evidence that an identified problem will introduce bias.

WHAT'S NEW

Last assessed as up-to-date: 20 May 2016.

Date	Event	Description
22 June 2016	New citation required and conclusions have changed	We have updated the review: new studies added and conclusions changed. We have updated inconsistencies between the text and abstract present in the first publication of the review

HISTORY

Protocol first published: Issue 2, 2000

Review first published: Issue 1, 2002

Date	Event	Description
23 July 2008	Amended	Converted to new review format
27 August 2006	New citation required and conclusions have changed	Substantive amendment

CONTRIBUTIONS OF AUTHORS

All authors read through all abstracts, appraised the potential papers, appraised the included trials, and extracted the data. The text of the review was a collaborative effort by all four authors. AS ran the searches, and BH checked that they were correct.

DECLARATIONS OF INTEREST

Mark Jeffery was an international member of the Follow-up After Colorectal Surgery (FACS) trial management committee.

For other authors in the byline (BEH, PNH, AMS), nothing to declare.

SOURCES OF SUPPORT

Internal sources

- Princess Alexandra Hospital Cancer Collaborative Group, Australia.

Supported AS (who ran search strategies for the review)

External sources

- No sources of support supplied

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

- We have simplified the wording of the Objectives. We have deleted the reference to secondary endpoints as they are itemised under Types of outcome measures.
- Types of participants: we have simplified cancer staging.
- Types of interventions: we have added the phrase “investigations including but not limited to” for clarity.
- Types of outcome measures > Secondary outcomes:
    - we have renamed ’Disease-specific survival’ as ’Colorectal cancer-specific survival’.
    - We changed ’Time to diagnosis of recurrence’ to ’Relapse-free survival’.

- We changed 'Incidence of surgery (with curative intent for recurrence)' to 'Salvage surgery (surgery performed with curative intent for relapse of colorectal cancer (CRC))'.
- We replaced 'Interval (between planned visits) recurrences' with 'Interval recurrences (relapse of CRC detected between follow-up visits)'.
- We clarified the quality of life (QoL) outcome to indicate that we permitted the use of trial-specific QoL instruments.
- Harms and costs of surveillance now specifically include that of investigations.
- We defined CRC (colorectal cancer) -specific survival and RFS (relapse-free survival).

- We moved the search strategy to Appendices.
- Our reporting of the search strategy is now consistent with MECIR (Methodological Expectations of Cochrane Intervention Review) guidelines.
- We updated our reporting of the processes relating to selection of studies and data extraction and management in accordance with the MECIR guidelines.
- We modified our assessment of risk of bias to ensure consistency with the current *Cochrane Handbook for Systematic Reviews of Interventions* recommendations.
- We modified our measures of treatment effect to incorporate time-to-event data where possible.
- We have added a 'Summary of findings' table.
- We now present time-to-event data with hazard ratios, rather than risk ratios.
- We have deleted the multiple post hoc subgroup analyses we performed in the earlier version of the review (clinic visits and tests versus no clinic visits and tests, more clinic visits versus fewer clinic visits, more tests versus fewer tests, community versus hospital, liver imaging versus no liver imaging, time to recurrence with intensive follow-up versus time to recurrence with less intensive follow-up).These were included in previous versions of the review, but we now recognise that large numbers of undirected sub-group analyses may lead to spurious explanations of heterogeneity (Higgins 2016). In response to a reviewer suggestion, we included post hoc subgroup analysis of follow-up "dose" and the use of different settings (specialist versus GP- or nurse-led follow-up).
- We used a random-effects model for meta-analysis based on reviewer input.
- In response to reviewer input, we performed sensitivity analysis by excluding one study (Ohlsson 1995), where the intensity of follow-up in the intensive arm was comparable with the intensity of follow-up in the control arm of other studies.

INDEX TERMS

Medical Subject Headings (MeSH)

Clinical Protocols; Colorectal Neoplasms [*mortality; *therapy]; Disease-Free Survival; Follow-Up Studies; Neoplasm Recurrence, Local; Quality of Life; Randomized Controlled Trials as Topic; Salvage Therapy

MeSH check words

Female; Humans; Male