

# Comparative efficacy and tolerability of adjuvant systemic treatments against resectable colon cancer: a network meta-analysis

Ji Cheng , Xiaoming Shuai, Jinbo Gao, Guobin Wang, Kaixiong Tao and Kailin Cai

*Ther Adv Med Oncol*

2020, Vol. 12: 1–21

DOI: 10.1177/  
1758835920974195

© The Author(s), 2020.  
Article reuse guidelines:  
[sagepub.com/journals-  
permissions](https://sagepub.com/journals-permissions)

## Abstract

**Background:** Currently, 6-month oxaliplatin-based chemotherapy has been recommended as the preferred adjuvant treatment against high-risk stage 2 and stage 3 colon cancer patients.

**Methods:** Record retrieval was conducted in PubMed, Web of Science, Cochrane Central Register of Controlled Trials, American Society of Clinical Oncology and European Society for Medical Oncology meeting libraries from inception to November 2019. Regarding survival and tolerability, randomized controlled trials comparing different adjuvant systemic regimens against high-risk stage 2 and stage 3 colon cancer were eligible. Disease-free survival was primary endpoint. Network calculation was based on a random-effects model, and relative ranking of each node was numerically indicated by *p* score.

**Results:** A total of 30 trials were included, corresponding to 54,109 patients. Regarding disease-free survival, none of the analyzed regimens displayed significant superiority against common comparator 6-month capecitabine plus oxaliplatin (XELOX), while 12-month [network hazard ratio (HR) 0.81 (0.60–1.10); 0.79 (0.57–1.10)] and 3-month XELOX [0.95 (0.86–1.04); 0.93 (0.83–1.05)] were top-ranking regimens showing non-inferiority among overall and stage 3 patients. Moreover, by pairwise meta-analysis, 3-month XELOX demonstrated significant superiority against 6-month XELOX among low-risk stage 3 patients [pairwise HR 0.78 (0.63–0.97)]. Concerning adverse events, 3-month oxaliplatin-based chemotherapy was significantly better than the 6-month counterpart with respect to peripheral sensory neuropathy, thrombocytopenia and fatigue. The 12-month capecitabine monotherapy failed to display non-inferiority among other major adverse events.

**Conclusions:** The 3-month XELOX treatment could be an alternative option of the 6-month regimen among low-risk stage 3 patients. Among high-risk stage 3 patients, 6-month oxaliplatin-based regimens still seem more competitive. In addition, clinical application of 12-month capecitabine monotherapy should be cautious, despite its top rankings, especially among non-Asian countries.

**Keywords:** 3-month capecitabine plus oxaliplatin, adjuvant treatment, network meta-analysis, resectable colon cancer, systematic review

Received: 25 December 2019; revised manuscript accepted: 23 October 2020.

## Introduction

Colorectal cancer is currently the fourth most common and fifth most lethal malignancy worldwide. It is estimated that nearly 1,400,000 cases occur annually, while almost 700,000 patients die because of it each year.<sup>1</sup> In the United States, although the overall incidence of colorectal cancer

is reported to decrease during the past decade due to earlier diagnosis and better cancer prevention,<sup>2,3</sup> more efforts are still required to enhance the survival probability among those cancer sufferers.

Adjuvant systemic chemotherapy has become the standard of care against stage 2 and stage 3 colon

Correspondences to:

**Ji Cheng**  
Department of  
Gastrointestinal Surgery,  
Union Hospital, Tongji  
Medical College, Huazhong  
University of Science  
and Technology, No.1277  
Jiefang Avenue, Wuhan  
430022, China  
[jicheng1@hust.edu.cn](mailto:jicheng1@hust.edu.cn)

**Kailin Cai**  
Department of  
Gastrointestinal Surgery,  
Union Hospital, Tongji  
Medical College, Huazhong  
University of Science and  
Technology, No. 1277  
Jiefang Avenue, Wuhan  
430022, China  
[caikailin@hust.edu.cn](mailto:caikailin@hust.edu.cn)

**Xiaoming Shuai**  
**Jinbo Gao**  
**Guobin Wang**  
Department of  
Gastrointestinal Surgery,  
Huazhong University of  
Science and Technology,  
Wuhan, China

**Kaixiong Tao**  
Department of  
Gastrointestinal Surgery,  
Union Hospital, Tongji  
Medical College, Huazhong  
University of Science and  
Technology, Wuhan, China  
Ji Cheng is the leading  
author

cancer following curative surgeries. Both National Comprehensive Cancer Network (NCCN; 2019.V2) and Chinese Society of Clinical Oncology (2019.V1) guidelines recommend oxaliplatin-based regimens in the adjuvant setting against stage 3 and high-risk stage 2 operable colon cancer, while fluoropyrimidine monotherapy (capecitabine or 5-FU/leucovorin) is optional for low-risk stage 2 patients.<sup>3,4</sup> The latest European Society for Medical Oncology (ESMO) colon cancer guideline (2013) suggests that oxaliplatin-based regimens should be regarded as the preferred options among stage 3 patients, while no specific regimen is recommended for high-risk stage 2 cases.<sup>5</sup> In the Japanese Society for Cancer of the Colon and Rectum guideline for colorectal cancer (2019), oxaliplatin-based regimens have also been recognized as the preferred regimens against stage 3 cases, while capecitabine, S-1 and UFT monotherapy are also considered as effective alternatives. However, fluoropyrimidine monotherapy is not indicated for stage 2 cases due to lacking evidence from randomized controlled trials (RCTs) based on the Japanese population.<sup>6</sup> In recent years, duration of adjuvant treatments has become the research hotspot in this field. Five largescale RCTs including ACHIEVE, HORG-IDEA, IDEA, SCOT and TOSCA (results of CALGB/SWOG 80702 had not been formally published in journal or meeting abstract) studied the relative efficacy and tolerability of 3-month *versus* 6-month oxaliplatin-based regimens,<sup>1,7-10</sup> which generally concluded that the option of shorter or longer treatment depended on regimen types and patient characteristics. The 3-month capecitabine plus oxaliplatin (XELOX) seemed to be non-inferior to 6-month regimen, while 3-month FOLFOX failed to show non-inferiority to its 6-month counterpart. Meanwhile, longer duration was more beneficial among high-risk stage 3 patients. However, the proportion of FOLFOX/XELOX regimen across five trials was not quite comparable, was the authors' conclusions.<sup>1,7-10</sup> Therefore, all these results added complexity to regimen selection in the adjuvant colorectal cancer setting.

Currently, there is still scarcity of comprehensive hierarchical evidence to compare and rank all possible regimens simultaneously, which could offer more statistically straightforward and accurate outcomes than pairwise comparisons. Network meta-analysis could provide indirect calculations between regimens that lack direct comparisons.<sup>11</sup> Hence, in consideration of the

rapidly growing types of chemotherapeutic strategies, as well as methodological imperfections regarding pairwise RCTs and meta-analyses, we decided to perform the first systematic review and network meta-analysis in this field.

## Methods

### *Registration and guidelines*

The protocol of our systematic review and network meta-analysis had been listed in PROSPERO [CRD42020147304]. The design, conduct and writing of this systematic review and network meta-analysis complied with the requirements from Preferred Reporting Items for Systematic Review and Meta-Analysis Checklist for Network Meta-analysis and Cochrane Handbook 5.1. Each step was performed by two researchers of our group. Any disagreement was resolved by the third researcher.

### *Search strategy*

Electronic databases including PubMed, Web of Science and Cochrane Central Register of Controlled Trials were thoroughly examined. Additionally, we also searched major databases for meeting abstracts, including the American Society of Clinical Oncology (ASCO) and ESMO Meeting Library. The searching process started from 1 July until 10 November 2019, covering possible indexes published from inception to November 2019. Both abstract and main text of the retrieved records were rigorously checked in order to guarantee the accuracy of selection. The full electronic search strategy is presented in the Supplemental Materials.

### *Selection criteria*

Studies that met all following criteria were therefore included (Participants, Intervention, Comparator, Outcome and Study design [PICOS] framework):

- (1) Participants: patients should be diagnosed with previously untreated high-risk stage 2 and stage 3 resectable colon cancer without pathological selection. For trials studying targeted therapies, subgroup data of certain pathological or genetic status was permitted; however, overall results of unselected population should also be provided. Upper rectal cancer cases were also allowed since they shared similar biological features and therapeutic options with

colon cancer patients. Patients with synchronous malignancies other than colon cancer were not permitted.

- (2) Intervention: adjuvant systemic treatments should be given after curative surgeries, including intravenous or oral chemotherapeutic and targeted medications. Since there were several outdated drugs that were used against colon cancer but were no longer utilized currently in the clinical setting (such as mitomycin-C, methotrexate, vincristine, semustine and edrecolomab), we only included chemotherapeutic and targeted drugs that were currently approved and recommended for use against colon cancer by major countries, including capecitabine, oxaliplatin, irinotecan, 5-FU/leucovorin, S-1, UFT, bevacizumab, cetuximab and raltitrexed. Comparisons between different regimens deriving from any of these drugs in the adjuvant setting were deemed eligible. Moreover, comparisons between different durations of treatment by the same chemotherapeutic regimen were also qualified. Therefore, trials containing hyperthermic intraperitoneal chemotherapy, intraarterial chemotherapy, preoperative or postoperative radiotherapy were regarded as ineligible. Also, since adjuvant systemic treatments had been widely accepted as standard of care for high-risk stage 2 and stage 3 colon cancer, trials featuring comparisons between chemotherapy and observation only were also not included.
- (3) Comparator: ‘XELOX (6M)’ (6-month capecitabine plus oxaliplatin regimen), ‘FP (6M)’ (6-month fluoropyrimidine plus platinum regimen) were common comparator nodes of network meta-analysis under different scenarios.
- (4) Outcome: time-to-event disease-free survival (DFS) data (hazard ratio or Kaplan–Meier curves) were mandatory, while results of overall survival (OS) and adverse events were dispensable.
- (5) Study design: phase II and phase III RCTs reported from inception to November 2019 without language limitations.

Studies were excluded for the following reasons:

- (1) Besides chemotherapeutic or targeted medications, auxiliary therapeutics were also contained and comparatively studied,

including non-steroidal anti-inflammatory drugs (NSAIDs), nutritional supportive methods (vitamins), unspecified herbal medicine (lentinan), general immunomodulators (interferons, polysaccharide K, polyadenylic–polyuridylic acid and *Bacillus Calmette–Guerin*) or levamisole (eTable 1).

- (2) Cross-over design of RCTs.

#### *Risk-of-bias assessment*

The quality of each eligible trial was assessed by The Cochrane Risk-of-Bias Tool. The entire scale consisted of seven categories, including random sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective reporting and other sources of bias. According to Cochrane Handbook 5.1, each category could be scored as low risk, unclear risk or high risk of bias once met certain criteria. If the majority of items were judged as low risk of bias, then the entire methodological design of network meta-analysis was regarded as low risk of bias, and *vice versa*. Here, trials were regarded to be low quality if four or more categories were evaluated as high risk of bias.

#### *Data extraction*

Pre-designed forms were utilized to collect and organize the original data. Baseline characteristics, efficacy and tolerability data were extracted from main text, tables, survival curves or supplemental material, which had been cross-checked by two different researchers in our group before quantitative synthesis.

#### *Endpoints and nodes*

The primary endpoint was DFS, while secondary endpoints included OS and adverse events. The definitions of DFS were mainly consistent across different trials (Supplemental Material). In terms of adverse events, we analyzed 12 common types of treatment-related adverse events including leucopenia, neutropenia, anaemia, thrombocytopenia, nausea/vomiting, anorexia, diarrhoea, fatigue, hand–foot syndrome, peripheral sensory neuropathy, alanine transaminase (ALT)/aspartate transaminase (AST) and creatinine. We only counted grade 3 or higher (National Cancer Institute Common Terminology Criteria for Adverse Events) adverse events due to their

clinical significance. Criteria for adverse events judgement were also generally consistent across different trials (Supplemental Material).

The major principle for node classification was to combine homogenic arms together so that sample sizes and advantages of direct randomization could be enlarged. Key indicators to ensure homogeneity were clinical and methodological features, which jointly contributed to statistical homogeneity across the trials. Since we only included RCTs into our pooled analysis, methodological heterogeneity was low among included trials. Therefore, clinical features were critical for maintaining homogeneity inside each node, such as treatment regimens and pathological stages. Moreover, since DFS was the primary endpoint, baseline DFS rate (3 years and 5 years) was crucial for preliminarily judging statistical homogeneity, which also reflected clinical homogeneity across different trials within the same node. Hence, taken together, we classified nodes by different treatment regimens since it was the main focus of our meta-analysis and also acted as the major clinical heterogenic factor inside the network. Nevertheless, if baseline survival rates of different studies inside the same node were still not consistent, this might hint other underlying clinical heterogeneity besides treatment regimens, such as clinical stages and lymphadenectomy statuses, which would be further analyzed *via* sensitivity and subgroup analysis. On the other hand, in order to form an intact network for statistical calculation and also minimize an unnecessary number of nodes to enhance statistical power, we also integrated some regimens that were slightly different in terms of treatment schedules into one node, as long as their baseline DFS rates were comparable. To be specific, the majority of nodes in our meta-analysis were made according to their original treatment schedules, such as node 'XELOX (6M)' corresponding to 6-month capecitabine plus oxaliplatin regimen. Although different studies utilized slightly different regimens of FOLFIRI, there was only one 'FOLFIRI' node inside our network. Among all eligible studies, regimens of 5-FU plus leucovorin had several types of variations; therefore, node 'LV5FU2 (6M)', 'FU/FA-RP (Roswell Park regimen)' and 'FU/FA-MC (Mayo Clinic regimen)' were created to fit different schedules. For chemotherapeutic drugs plus bevacizumab or cetuximab, although the actual chemotherapeutic regimens were not completely identical across included trials, we still integrated them into two nodes 'F/

bevacizumab' and 'F/Cetuximab' (F here represented fluoropyrimidines) to facilitate network calculation, since their baseline survival rates were quite comparable. In addition, there were two types of node classification systems within our network meta-analysis, namely Node-1 and Node-2. The only difference between these two types was that Node-1 separated all fluoropyrimidine-plus-platinum regimens into specific regimens, such as mFOLFOX6, FOLFOX4, SOX and XELOX, while Node-2 combined them together so that comparisons between the 3-month schedule *versus* the 6-month schedule were much easier. As abovementioned, although we tried hard to restrict heterogeneity inside each node, there might still be certain degrees of heterogeneity that warranted further sensitivity or subgroup analysis. Treatment schedules of all included trials were listed in the Supplemental Material.

#### Statistical analysis

Hazard ratio (HR) and its 95% confidential interval (95% CI) were used as the effect size for DFS and OS. Risk ratio (RR) and its 95% CI were applied as the effect size for adverse events. If survival data were not directly provided, we estimated the values from Kaplan–Meier curves by methods described elsewhere.<sup>12,13</sup>

It was well known that network meta-analysis could offer a hierarchical ranking among multiple arms despite lacking direct comparisons.<sup>14,15</sup> This vital advantage was based on two key assumptions of network meta-analysis that were known as transitivity and consistency, respectively.<sup>15,16</sup>

When pairwise comparisons of A *versus* C and B *versus* C were separately provided, transitivity of network meta-analysis further validated the statistical comparison between A and B. Nevertheless, it required comparable baseline characteristics as the prerequisite condition for minimizing selection bias and therefore justifying subsequent connections between indirect arms.<sup>17</sup> Because all eligible studies were randomized trials without significant heterogeneity on methodological design, clinical features were crucial to determine baseline heterogeneity, as well as network transitivity. We carefully compared key clinical features among different arms inside each node and then removed those with significant heterogeneity by performing sensitivity analysis. Besides possible clinical and methodological disparities, we also evaluated statistical heterogeneity

inside our network calculation.  $I^2$  was used as the main indicator for statistical heterogeneity, with its value <25%, 25–50% and >50% suggesting low, moderate and high heterogeneity, respectively. Moreover,  $Q$  static of heterogeneity also helped to assess statistical heterogeneity.

On the other side, consistency, another main assumption for network meta-analysis, referred to statistically consistent results between direct and indirect calculations concerning the same comparison. Significant differences between direct and indirect results could suggest inconsistency across network meta-analysis, as well as unsuitability for transitivity. Therefore, we utilized several approaches to evaluate network consistency, including the loop-specific method and the  $Q$  static. The loop-specific method could analyze mutual variance between direct and indirect results *via* closed loops. Inconsistency factor (IF) was the quantitative indicator for the loop-specific method, which hinted inconsistency once its 95% CI excluded zero.<sup>15</sup> Furthermore, the  $Q$  static of inconsistency was another indicator to estimate consistency across the network. Both consistency and homogeneity were fundamental requirements before producing reliable results by network meta-analysis. When inconsistency or significant heterogeneity was detected, data from the most inconsistent or heterogeneous comparisons were removed to examine whether the results remained stable.

Network plot as well as funnel plot were applied to demonstrate network structure and detect publication bias, respectively. The more symmetrical the funnel plot was, the less publication bias pooled results would have. We performed random-effects network calculation based on a frequentist model, with either HR or RR as the effect size. Based on the non-inferior margin of previous literature on a similar topic,<sup>18</sup> we set 1.12 for the HR, as well as 1.25 for the RR to be the non-inferior margin in our network calculation. In addition, we also utilized  $p$  score to rank all regimens based on their network estimates. The closer the  $p$  score approached 1, the better the regimen could be. However, if one regimen ranked in top place, however crossed a non-inferior margin as well, it still could not be fully recommended and trusted. The final conclusion was made by considering both the network ranking and non-inferiority of each regimen. Sensitivity analysis was performed to detect the stability of pooled outcomes by deleting studies with significant clinical

heterogeneity. Subgroup analysis by different pathological stages was also conducted to validate potential heterogenic factors, as well as provide more clinically meaningful evidence. Network meta-analysis was conducted on R software 3.4.3, assisted by STATA 14.0 in terms of graphical functions.

#### *Role of the funding source*

The sponsors had no role in study design, data collection, data analysis, data interpretation or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

## Results

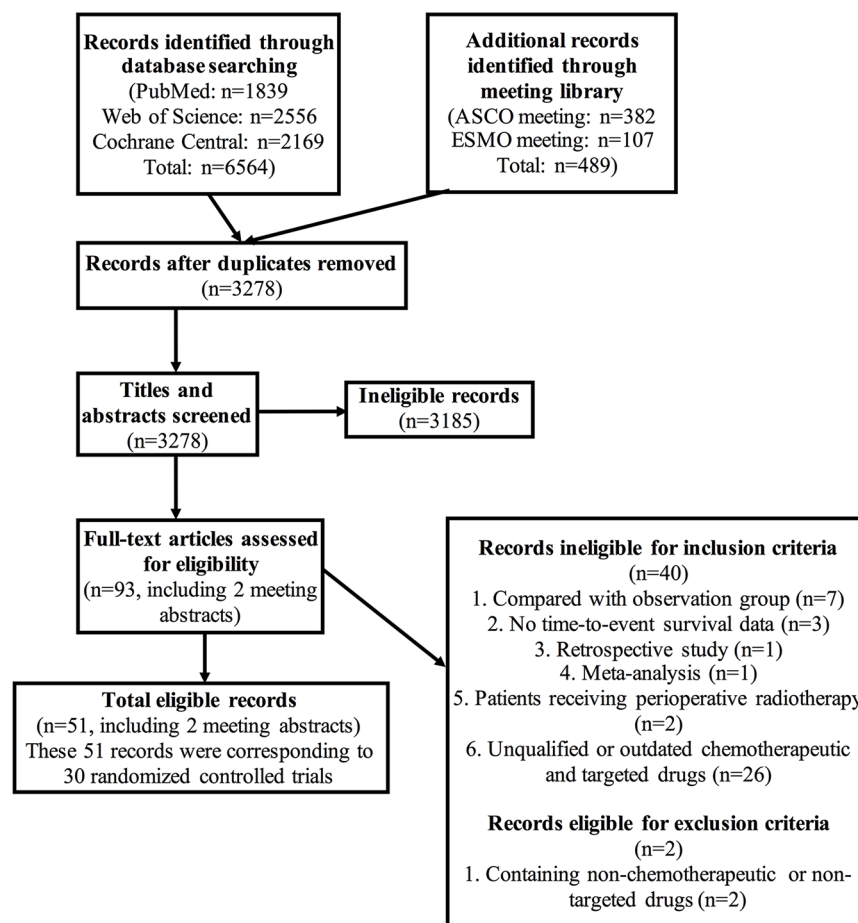
### *Baseline features*

After screening 7053 preliminary records, 51 records were included into our systematic review and network meta-analysis, corresponding to 30 RCTs. Selection flowchart and reasons of ineligibility by full-text assessment are described in Figure 1 and eTable 1, respectively. Due to limitation on number of references, we included citations of all eligible trials in the Supplemental Material.

All 30 trials were phase III RCTs, and the majority of them had formal registration identifiers. Of these, 22 trials were conducted among the Western population while only 8 trials were launched amid Asian countries, which were all completed by Japanese institutions. Total sample sizes of eligible trials were 54,109, ranging from 169 to 6088, individually. All included trials were relatively comparable in terms of median age (around 60-years old) and sex ratio of enrolled patients (male dominant). Only 15 trials recruited stage 3 colon cancer patients, while 15 trials studied both high-risk stage 2 and stage 3 colon cancer cases. The distribution of tumour location and performance status of included patients were also consistent across different trials. Therefore, overall, the baseline clinical features of all eligible trials were comparable, while the impact of different stages would be further analyzed by subgroup analysis (Tables 1 and 2).

### *Risk of bias*

Generally, the entire systematic review had a low risk of bias, since more than half the indicators



**Figure 1.** Selection flowchart.

ASCO, American Society of Clinical Oncology; ESMO, European Society for Medical Oncology.

scored as low risk of bias (56%), while unclear risk (24%) or high risk of bias (20%) took up smaller proportions (Figure 2). Individually, none of the included trials was in high risk of bias for methodological design (eTable 2).

Of note, since the majority of trials were rigorously randomized as well as centrally allocated, 70% and 87% of included trials were scored as low risk of bias in terms of random sequence generation and allocation concealment, respectively, while no high risk of bias was reported in these two key domains. Due to open-label design and impossibility for treatment masking with greatly differently administered arms, all the included trials (100%) were scored as high risk of bias in terms of blinding of participants and personnel. The majority of trials did not report relevant information regarding blinding of outcome assessment, especially whether independent reviewers were introduced into the evaluation of DFS; therefore, most

of them were scored as unclear risk of bias (87%). Since efficacy and tolerability of the majority of trials were based on intent to treat and safety analysis, set respectively, most trials reported enough endpoints; 93% and 67% of the eligible trials had low risk of bias regarding incomplete outcome data and selective reporting, respectively. Additionally, since many eligible trials featured balanced clinical characteristics, 70% of all qualified trials were scored as low risk of bias with respect to other sources of bias (Figure 2).

#### *Primary endpoint: disease-free survival*

*Network geometry.* 30 RCTs were merged into quantitative analysis, corresponding to 19 network nodes by type 1 node classification (Figure 3 and Tables 1 and 2).

*Transitivity.* As was mentioned in the Methods section, we rearranged all included arms by

**Table 1.** Baseline features of all eligible trials (part 1).

Study	Trial (identifier)	Region	Phase	Enrolment	Treatment	Node-1	Node-2	Sample size	Median age	Sex (M/F)	Stage (2/3)	Location (R/L)	ECOG (0/1)
Paschke 2019	FOGT 4 (BfArM4019926)	Germany	3	NA	FOLFIRI (6M)	FOLFIRI	FOLFIRI	136	64.8	81/55	19/117	59/76	All 0-1
Souglaklos 2019	HORG-IDEA (NCT01308086)	Greece	3	2009.4-2015.10	FU/FA (6M) FOLF0X4 (3M) XELOX (3M) FOLF0X4 (6M) XELOX (6M)	FU/FA-RP FOLF0X4 (3M) XELOX (3M) FOLF0X4 (6M) XELOX (6M)	FP (3M) FP (6M) FP (6M) FP (6M)	195 362 196 362	67.0 67.0 65.0 66.0	109/86 204/158 104/92 205/157	50/145 156/206 51/145 156/206	80/115 158/204 86/110 170/192	152/43 300/62 171/25 309/53
Takahashi 2019	ACTS-CC 02 (JapicCTI-101073)	Japan	3	2010.4-2014.10	SOX (6M)	SOX (6M)	FP (6M)	460	65.0	253/207	All stage 3	171/139	429/31
Tomita 2019	JFMC37-0801 (UMIN00001367)	Japan	3	2008.9-2009.12	UFT/LV (6M) Capecitabine (12M) Capecitabine (6M)	UFT/LV (6M) Capecitabine (12M) Capecitabine (6M)	UFT/LV (6M) Capecitabine (12M) Capecitabine (6M)	472 650 654	66.0 65.0 65.0	252/220 343/307 352/302	All stage 3	263/252 262/258	All 0-1
Yoshino 2019	ACHIEVE (UMIN00008543)	Japan	3	2012.8-2014.6	mFOLFOX6 (3M) XELOX (3M) mFOLFOX6 (6M) XELOX (6M)	mFOLFOX6 (3M) XELOX (3M) mFOLFOX6 (6M) XELOX (6M)	FP (3M) XELOX (3M) FP (6M) XELOX (6M)	163 487 159 482	69.0 65.0 67.0 65.0	77/86 252/235 81/78 239/243	All stage 3	NA	156/7 467/20 155/4 467/15
André 2018	IDEA (NCT00958737)	France	3	2009.5-2014.5	mFOLFOX6 (3M) XELOX (3M) mFOLFOX6 (6M) XELOX (6M)	mFOLFOX6 (3M) XELOX (3M) mFOLFOX6 (6M) XELOX (6M)	FP (3M) XELOX (3M) FP (6M) XELOX (6M)	895 107 914 94	64.8 63.8 65.1 60.3	506/389 57/50 516/398 65/29	All stage 3	335/505 42/64 341/527 28/65	660/224 79/25 671/230 69/23
Hamaguchi 2018	JCOG0910 (UMIN00003272)	Japan	3	2010.3-2013.8	S-1 (6M) Capecitabine (6M)	S-1 (6M) Capecitabine (6M)	S-1 (6M) Capecitabine (6M)	782 782	66.0 66.0	387/395 418/364	All stage 3	288/245 290/242	758/24 755/27

(Continued)

Table 1. (Continued)

Study	Trial (Identifier)	Region	Phase	Enrolment	Treatment	Node-1	Node-2	Sample size	Median age	Sex (M/F)	Stage (2/3)	Location (R/L)	ECOG (0/1)
Iveson 2018	SCOT (ISRCTN59757862)	Western	3	2008.3–2013.11	mFOLFOX6 (3M)	mFOLFOX6 (3M)	FP (3M)	993	65.0	1843/1201	551/2493	NA	2190/854
Kusumoto 2018	ACTS-CC (NCT00660894)	Japan	3	2008.4–2009.6	XELOX (3M)	XELOX (3M)	XELOX (3M)	2051					
					mFOLFOX6 (6M)	mFOLFOX6 (6M)	FP (6M)	988	65.0	1844/1200	549/2499	2144/900	
					XELOX (6M)	XELOX (6M)	XELOX (6M)	2056					
Sobrero 2018	TOSCA (NCT00646607)	Italy	3	2007.6–2013.3	S-1 (6M)	S-1 (6M)	S-1 (6M)	758	66.0	411/347	All stage 3	324/278	722/36
					UFT/LV (6M)	UFT/LV (6M)	UFT/LV (6M)	760	65.5	403/357	268/314	727/33	
					FOLFOX4 (3M)	FOLFOX4 (3M)	FP (3M)	1848	63.4	1035/807	641/1207	630/734	1749/90
Kerr 2016	QUASAR 2 (ISRCTN45133151)	Western	3	2005.4–2010.10	XELOX (3M)	XELOX (3M)	XELOX (3M)						
					FOLFOX4 (6M)	FOLFOX4 (6M)	FP (6M)	1867	63.1	1027/837	648/1219	594/745	1760/103
					XELOX (6M)	XELOX (6M)	XELOX (6M)						
André 2015	MOSAIC (NCT00275210)	Western	3	1998.10–2001.1	Capecitabine/bevacizumab (12M)	F/bevacizumab (12M)	F/bevacizumab (12M)	973	65.0	555/418	371/602	NA	All 0–1
					Capecitabine (6M)	Capecitabine (6M)	Capecitabine (6M)	968	65.0	554/414	373/595		
					FOLFOX4 (6M)	FOLFOX4 (6M)	FP (6M)	1123	61.0	630/493	451/672	394/605	968*
Pectasides 2015	ANZCTR12610000509066	Greece	3	2005.11–2008.1	LV5FU2 (6M)	LV5FU2 (6M)	LV5FU2 (6M)	1123	60.0	588/535	448/675	374/630	984*
					mFOLFOX6 (6M)	mFOLFOX6 (6M)	FP (6M)	197	62.4	111/86	61/130	58/86	177/16
					XELOX (6M)	XELOX (6M)	FP (6M)	211	63.7	117/94	68/135	64/87	195/13
Sadahiro 2015	JFMC33-0502 (UMIN-CTR C000000245)	Japan	3	2005.10–2007.9	UFT/LV (18M)	UFT/LV (18M)	UFT/LV (18M)	537	64.0	264/273	75/462	218/211	517/20
					UFT/LV (6M)	UFT/LV (6M)	UFT/LV (6M)	534	64.0	294/240	68/466	199/221	503/31
Schmoll 2015	XELOXA (NCT00069121)	Western	3	2003.4–2004.10	XELOX (6M)	XELOX (6M)	FP (6M)	944	61.0	513/431	All stage 3	NA	701/235
					FU/FA (6M)	FU/FA-PP	FU/FA-PP	942	62.0	500/442	727/210		

(Continued)



Table 1. (Continued)

Study	Trial (identifier)	Region	Phase	Enrolment	Treatment	Node-1	Node-2	Sample size	Median age	Sex (M/F)	Stage (2/3)	Location (R/L)	ECOG (0/1)
Huang 2014	NCCTG N0147 (NCT00079274)	USA	3	2004.2– 2009.11	FOLFIRI/ cetuximab (6M)	F/cetuximab (6M)	F/cetuximab (6M)	40	59.0	22/18	All stage 3	24/16	97/49
					FOLFIRI (6M)	FOLFIRI	FOLFIRI	106	57.0	56/50		61/43	
					mFOLFOX6/ cetuximab (6M)	F/cetuximab (6M)	F/cetuximab (6M)	1297	58.0	681/616		NA	NA
					mFOLFOX6 (6M)	mFOLFOX6 (6M)	FP (6M)	1283	58.0	678/605			
Shimada 2014	JCOG0205 (NCT00190515)	Japan	3	2003.2– 2006.11	UFT/LV (6M)	UFT/LV (6M)	UFT/LV (6M)	551	61.0	302/249	All stage 3	NA	522/29
Taieb 2014	PETACC-8 (EudraCT2005-003463-23)	Western	3	2005.12– 2009.11	FOLFAX4/ cetuximab (6M)	F/cetuximab (6M)	F/cetuximab (6M)	1159	60.0	676/483	All stage 3	445/699	913/200
					FOLFAX4 (6M)	FOLFAX4 (6M)	FP (6M)	1185	60.0	664/521		445/720	935/196
Allegra 2013	NSABP C-08 (NCT00096278)	USA	3	2004.9– 2006.10	mFOLFOX6/ Bevacizumab (12M)	F/bevacizumab (12M)	F/ bevacizumab (12M)	1335	NA	666/669	334/1001	NA	1075/259
					mFOLFOX6 (6M)	mFOLFOX6 (6M)	FP (6M)	1338		665/673	332/1006		1089/249
Shichinohe 2013	HGCSG-CAD (NCT00209742)	Japan	3	2005.4– 2012.12	UFT/LV (6M)	UFT/LV (6M)	UFT/LV (6M)	84	NA	NA	All stage 3	NA	NA
					UFT/LV (18M)	UFT/LV (18M)	UFT/LV (18M)	85					
de Gramont 2012	AVANT (NCT00112918)	Western	3	2004.12– 2007.6	XELOX/ Bevacizumab (12M)	F/bevacizumab (12M)	F/ bevacizumab (12M)	1145	58.0	625/520	187/952	NA	978/165
					FOLFAX4/ bevacizumab (12M)	F/bevacizumab (12M)	F/ bevacizumab (12M)	1155	58.0	587/568	194/960		987/166
					FOLFAX4 (6M)	FOLFAX4 (6M)	FP (6M)	1151	58.0	656/495	192/955		994/156
Twelves 2012	X-ACT	Western	3	1998.11– 2001.11	Capecitabine (6M)	Capecitabine (6M)	Capecitabine (6M)	1004	62.0	542/462	All stage 3	NA	853/151
					FU/FA (6M)	FU/FA-MC	FU/FA-MC	983	63.0	531/452			836/147

(Continued)

Table 1. (Continued)

Study	Trial (identifier)	Region	Phase	Enrolment	Treatment	Node-1	Node-2	Sample size	Median age	Sex (M/F)	Stage (2/3)	Location (R/L)	ECOG (0/1)
Papadimitriou 2011	HeCOG (ACTRN12610000148077)	Greece	3	1999.1–2004.9	FOLFIRI (9M)	FOLFIRI	FOLFIRI	441	65.0	235/206	214/227	155/254	All 0–2
Yothers 2011	NSABP C-07 (NCT00004931)	USA	3	2000.2–2002.11	FU/FA (8M)	FU/FA-RP	FU/FA-RP	432	65.0	199/233	211/221	162/238	NA
Van Cutsem 2009	PETACC-3 (NCT00026273)	Western	3	2000.1–2002.4	FOLFIRI (6M)	FOLFIRI	FOLFIRI	1485	60.0	826/659	435/1044	NA	1228/250
Ychou 2009	FFCD9802	France	3	1998.11–2002.9	LV5FU2 (6M)	LV5FU2 (6M)	LV5FU2 (6M)	1497	60.0	834/663	445/1050	62/73	1222/272
Popov 2008	PETACC-1 (ISRCTN2194324)	Western	3	1998.2–1999.7	Raltitrexed (6M)	Raltitrexed (6M)	Raltitrexed (6M)	952	62.6	518/434	All stage 3	62/73	132/63
André 2007	GERCOR C96.1	Western	3	1996.9–1999.11	FU/FA (6M)	FU/FA-RP	FU/FA-RP	453	59.7	240/213	195/254	118/172	103/82
Saltz 2007	CALGB 89803	Western	3	1999.4–2001.4	FU/FA (8M)	FU/FA-RP	FU/FA-RP	629	NA	348/281	All stage 3	296/248	463/145
Lembersky 2006	NSABP C-06	USA	3	1997.2–1999.3	FOLFIRI (7M)	FOLFIRI	FOLFIRI	635	NA	354/281	283/270	326/177	459/155
					UFT/LV (6M)	UFT/LV (6M)	UFT/LV (6M)	781	NA	418/363	365/416	326/177	NA
					FU/FA (6M)	FU/FA-RP	FU/FA-RP	770	NA	397/373	357/413	333/169	NA

\*The numbers suggesting integrated numbers of both ECOG-0 and ECOG-1. For trials studying targeted drugs such as cetuximab and bevacizumab, results of efficacy were based on the unselected population to maintain histopathological homogeneity across all eligible trials. For 'region', some trials were completed by several Western countries instead of a sole nation, therefore 'Western' was used under this circumstance. Details of the rationale to name and constitute each node within our meta-analysis were depicted in the Methods section. The HR was the result of the upper arm versus lower arm in each trial, except for those that were specifically labelled, such as 'm3 versus m6'. 6M, 6-month regimen; 12M, 12-month regimen; DFS, disease-free survival; ECOG, European Cooperative Oncology Group; F3/F6, FOLFFOX4 (3M/6M); FA-MC, Mayo Clinic regimen; FA-RP, Roswell Park regimen; FB, FOLFFOX4/bevacizumab; HR, hazard ratio; m3/m6, mFOLFFOX6 (3M/6M); M/F, male/female; NA, not available; OS, overall survival; R/L, right/left; X3/X6, XELOX (3M/6M); XB, XELOX/bevacizumab; XELOX, capecitabine plus oxaliplatin.

**Table 2.** Baseline features of all eligible trials (part 2).

Study author(s)	Treatment	Node-1	Node-2	Sample size	3-year DFS	5-year DFS	3-year OS	5-year OS	DFS HR (Node-1)	DFS HR (Node-2)	OS HR (Node-1)	OS HR (Node-2)
Paschke 2019	FOLFIRI (6M)	FOLFIRI	FOLFIRI	136	NA	63.3%	85.5%	72.7%	0.92 [95% CI, 0.53–1.61]	0.92 [95% CI, 0.53–1.61]	0.90 [95% CI, 0.60–1.30]	0.90 [95% CI, 0.60–1.30]
	FU/FA (6M)	FU/FA-RP	FU/FA-RP	133	NA	65.1%	81.1%	69.9%				
Souglakos 2019	FOLFOX4 (3M)	FOLFOX4 (3M)	FP (3M)	195	73.1%	NA	NA	NA	F3 versus F6: 1.29 [95% CI, 1.08–1.51]	1.05 [95% CI, 0.61–1.55]	NA	NA
	XELOX (3M)	XELOX (3M)		362	77.9%				X3 versus X6: 0.99 [95% CI, 0.73–1.34]			
	FOLFOX4 (6M)	FOLFOX4 (6M)	FP (6M)	196	76.9%							
	XELOX (6M)	XELOX (6M)		362	78.1%							
Takahashi 2019	SOX (6M)	SOX (6M)	FP (6M)	460	62.7%	NA	NA	NA	0.90 [95% CI, 0.74–1.09]	0.90 [95% CI, 0.74–1.09]	NA	NA
	UFT/LV (6M)	UFT/LV (6M)	UFT/LV (6M)	472	60.6%							
Tomita 2019	Capecitabine (12M)	Capecitabine (12M)	Capecitabine (12M)	650	75.3%	68.7%	94.8%	87.6%	0.86 [95% CI, 0.71–1.04]	0.86 [95% CI, 0.71–1.04]	0.73 [95% CI, 0.55–0.96]	0.73 [95% CI, 0.55–0.96]
	Capecitabine (6M)	Capecitabine (6M)	Capecitabine (6M)	654	70.0%	65.3%	91.2%	83.2%				
Yoshino 2019	mFOLFOX6 (3M)	mFOLFOX6 (3M)	FP (3M)	163	73.9%	67.1%	NA	NA	m3 versus m6: 1.07 [95% CI, 0.71–1.60]	0.95 [95% CI, 0.76–1.20]	NA	NA
	XELOX (3M)	XELOX (3M)		487	81.4%	67.6%						
	mFOLFOX6 (6M)	mFOLFOX6 (6M)	FP (6M)	159	72.3%	70.0%			X3 versus X6: 0.90 [95% CI, 0.68–1.20]			
	XELOX (6M)	XELOX (6M)		482	79.7%	76.4%						
André 2018	mFOLFOX6 (3M)	mFOLFOX6 (3M)	FP (3M)	895	72.0%	65.3%	91.0%	81.1%	m3 versus m6: 1.27 [95% CI, 1.07–1.51]	1.24 [95% CI, 1.05–1.46]	m3 versus m6: 1.16 [95% CI, 0.90–1.49]	1.15 [95% CI, 0.91–1.46]
	XELOX (3M)	XELOX (3M)		107	72.0%	63.6%	92.0%	83.4%				
	mFOLFOX6 (6M)	mFOLFOX6 (6M)	FP (6M)	914	76.0%	71.7%	93.0%	84.9%	X3 versus X6: 0.97 [95% CI, 0.59–1.59]		X3 versus X6: 1.08 [95% CI, 0.49–2.37]	
	XELOX (6M)	XELOX (6M)		94	71.0%	65.3%	89.0%	87.7%				

(Continued)

Table 2. (Continued)

Study author(s)	Treatment	Node-1	Node-2	Sample size	3-year DFS	5-year DFS	3-year OS	5-year OS	DFS HR (Node-1)	DFS HR (Node-2)	OS HR (Node-1)	OS HR (Node-2)
Hameguchi 2018	S-1 (6M)	S-1 (6M)	S-1 (6M)	782	78.3%	73.3%	95.4%	90.6%	1.22 (95% CI, 1.00-1.50)	1.22 (95% CI, 1.00-1.50)	1.18 (95% CI, 0.83-1.68)	1.18 (95% CI, 0.83-1.68)
	Capecitabine (6M)	Capecitabine (6M)	Capecitabine (6M)	782	81.7%	77.5%	96.3%	92.7%				
Iveson 2018	mFOLFOX6 (3M)	mFOLFOX6 (3M)	FP (3M)	993	76.3%	70.8%	90.0%	81.1%	m3 versus m6: 1.16 (95% CI, 0.94-1.39)	1.01 (95% CI, 0.91-1.11)	NA	0.99 (95% CI, 0.96-1.14)
	XELOX (3M)	XELOX (3M)		2051	76.9%	72.2%						
	mFOLFOX6 (6M)	mFOLFOX6 (6M)	FP (6M)	988	79.2%	73.8%	89.6%	81.7%	X3 versus X6: 0.94 (95% CI, 0.84-1.07)			
	XELOX (6M)	XELOX (6M)		2056	76.1%	69.3%						
Kusumoto 2018	S-1 (6M)	S-1 (6M)	S-1 (6M)	758	75.5%	70.2%	93.6%	86.0%	0.88 (95% CI, 0.74-1.06)	0.88 (95% CI, 0.74-1.06)	0.92 (95% CI, 0.72-1.17)	0.92 (95% CI, 0.72-1.17)
	UFT/LV (6M)	UFT/LV (6M)	UFT/LV (6M)	760	72.5%	66.9%	92.7%	84.4%				
Sobrero 2018	FOLFOX4 (3M)	FOLFOX4 (3M)	FP (3M)	1848	NA	NA	NA	NA	F3 versus F6: 1.23 (95% CI, 1.03-1.46)	1.14 (95% CI, 0.99-1.32)	NA	NA
	XELOX (3M)	XELOX (3M)										
	FOLFOX4 (6M)	FOLFOX4 (6M)	FP (6M)	1867					X3 versus X6: 0.98 (95% CI, 0.77-1.26)			
	XELOX (6M)	XELOX (6M)										
Kerr 2016	Capecitabine/bevacizumab (12M)	F/bevacizumab (12M)	F/bevacizumab (12M)	973	75.4%	NA	87.5%	NA	1.06 (95% CI, 0.89-1.25)	1.06 (95% CI, 0.89-1.25)	1.11 (95% CI, 0.90-1.36)	1.11 (95% CI, 0.90-1.36)
	Capecitabine (6M)	Capecitabine (6M)	Capecitabine (6M)	968	78.4%	89.4%						
André 2015	FOLFOX4 (6M)	FOLFOX4 (6M)	FP (6M)	1123	78.6%	73.2%	88.1%	81.2%	0.82 (95% CI, 0.71-0.95)	0.82 (95% CI, 0.71-0.95)	0.85 (95% CI, 0.73-0.99)	0.85 (95% CI, 0.73-0.99)
	LV5FU2 (6M)	LV5FU2 (6M)	LV5FU2 (6M)	1123	73.4%	67.5%	86.5%	79.1%				
Pectasides 2015	mFOLFOX6 (6M)	mFOLFOX6 (6M)	FP (6M)	197	79.8%	73.4%	87.2%	80.3%	0.91 (95% CI, 0.58-1.44)	0.91 (95% CI, 0.58-1.44)	1.05 (95% CI, 0.68-1.60)	1.05 (95% CI, 0.68-1.60)
	XELOX (6M)	XELOX (6M)	FP (6M)	211	79.5%	70.5%	86.9%	80.2%				
Sadahiyo 2015	UFT/LV (18M)	UFT/LV (18M)	UFT/LV (18M)	537	73.6%	68.9%	91.9%	84.5%	1.00 (95% CI, 0.80-1.24)	1.00 (95% CI, 0.80-1.24)	1.05 (95% CI, 0.78-1.42)	1.05 (95% CI, 0.78-1.42)
	UFT/LV (6M)	UFT/LV (6M)	UFT/LV (6M)	534	73.2%	68.8%	91.6%	84.9%				

(Continued)

**Table 2.** (Continued)

Study author(s)	Treatment	Node-1	Node-2	Sample size	3-year DFS	5-year DFS	3-year OS	5-year OS	DFS HR (Node-1)	DFS HR (Node-2)	OS HR (Node-1)	OS HR (Node-2)
Schmoll 2015	XELOX (6M)	XELOX (6M)	FP (6M)	944	71.0%	66.1%	86.0%	77.6%	0.80 [95% CI, 0.69-0.93]	0.80 [95% CI, 0.69-0.93]	0.83 [95% CI, 0.70-0.99]	0.83 [95% CI, 0.70-0.99]
	FU/FA (6M)	FU/FA-RP	FU/FA-RP	942	67.0%	59.8%	84.0%	74.2%				
Huang 2014	FOLFIRI/cetuximab (6M)	F/cetuximab (6M)	F/cetuximab (6M)	40	87.0%	76.1%	92.0%	86.9%	0.53 [95% CI, 0.26-1.10]	0.53 [95% CI, 0.26-1.10]	0.45 [95% CI, 0.20-1.20]	0.45 [95% CI, 0.20-1.20]
	FOLFIRI (6M)	FOLFIRI	FOLFIRI	106	67.0%	60.8%	84.0%	73.7%				
	mFOLFOX6/cetuximab (6M)	F/cetuximab (6M)	F/cetuximab (6M)	1297	71.5%	NA	85.6%	NA	1.17 [95% CI, 0.99-1.39]	1.17 [95% CI, 0.99-1.39]	1.25 [95% CI, 0.99-1.59]	1.25 [95% CI, 0.99-1.59]
	mFOLFOX6 (6M)	mFOLFOX6 (6M)	FP (6M)	1283	74.6%		87.3%					
Shimada 2014	UFT/LV (6M)	UFT/LV (6M)	UFT/LV (6M)	551	77.8%	73.6%	93.9%	87.5%	1.02 [95% CI, 0.82-1.27]	1.02 [95% CI, 0.82-1.27]	1.06 [95% CI, 0.77-1.44]	1.06 [95% CI, 0.77-1.44]
	FU/FA (6M)	FU/FA-RP	FU/FA-RP	550	79.3%	74.3%	94.5%	88.4%				
Taieb 2014	FOLFOX4/Cetuximab (6M)	F/cetuximab (6M)	F/cetuximab (6M)	1159	73.7%	69.5%	88.0%	81.2%	1.06 [95% CI, 0.90-1.24]	1.06 [95% CI, 0.90-1.24]	1.08 [95% CI, 0.86-1.36]	1.08 [95% CI, 0.86-1.36]
	FOLFOX4 (6M)	FOLFOX4 (6M)	FP (6M)	1185	75.8%	69.6%	89.7%	82.3%				
Allegra 2013	mFOLFOX6/bevacizumab (12M)	F/bevacizumab (12M)	F/bevacizumab (12M)	1335	77.9%	70.2%	90.1%	82.5%	0.93 [95% CI, 0.81-1.08]	0.93 [95% CI, 0.81-1.08]	0.95 [95% CI, 0.79-1.13]	0.95 [95% CI, 0.79-1.13]
	mFOLFOX6 (6M)	mFOLFOX6 (6M)	FP (6M)	1338	75.1%	70.2%	89.1%	80.7%				
Shichinohe 2013	UFT/LV (6M)	UFT/LV (6M)	UFT/LV (6M)	84	73.8%	NA	95.2%	NA	1.15 [95% CI, 0.62-2.13]	1.15 [95% CI, 0.62-2.13]	0.57 [95% CI, 0.17-1.95]	0.57 [95% CI, 0.17-1.95]
	UFT/LV (18M)	UFT/LV (18M)	UFT/LV (18M)	85	77.6%		91.8%					
de Gramont 2012	XELOX/bevacizumab (12M)	F/bevacizumab (12M)	F/bevacizumab (12M)	1145	75.0%	70.4%	88.9%	82.0%	XB versus F: 1.07 [95% CI, 0.90-1.28]	XB versus F: 1.07 [95% CI, 0.90-1.28]	XB versus F: 1.15 [95% CI, 0.93-1.42]	XB versus F: 1.15 [95% CI, 0.93-1.42]
	FOLFOX4/Bevacizumab (12M)	F/bevacizumab (12M)	F/bevacizumab (12M)	1155	73.0%	65.4%	87.6%	81.0%	FB versus F: 1.17 [95% CI, 0.98-1.39]	FB versus F: 1.17 [95% CI, 0.98-1.39]	FB versus F: 1.27 [95% CI, 1.03-1.57]	FB versus F: 1.27 [95% CI, 1.03-1.57]
	FOLFOX4 (6M)	FOLFOX4 (6M)	FP (6M)	1151	76.0%	70.6%	90.3%	85.0%				
Twelves 2012	Capecitabine (6M)	Capecitabine (6M)	Capecitabine (6M)	1004	64.2%	60.8%	81.3%	71.4%	0.88 [95% CI, 0.77-1.01]	0.88 [95% CI, 0.77-1.01]	0.86 [95% CI, 0.74-1.01]	0.86 [95% CI, 0.74-1.01]
	FU/FA (6M)	FU/FA-MC	FU/FA-MC	983	60.6%	56.7%	77.6%	68.4%				

(Continued)

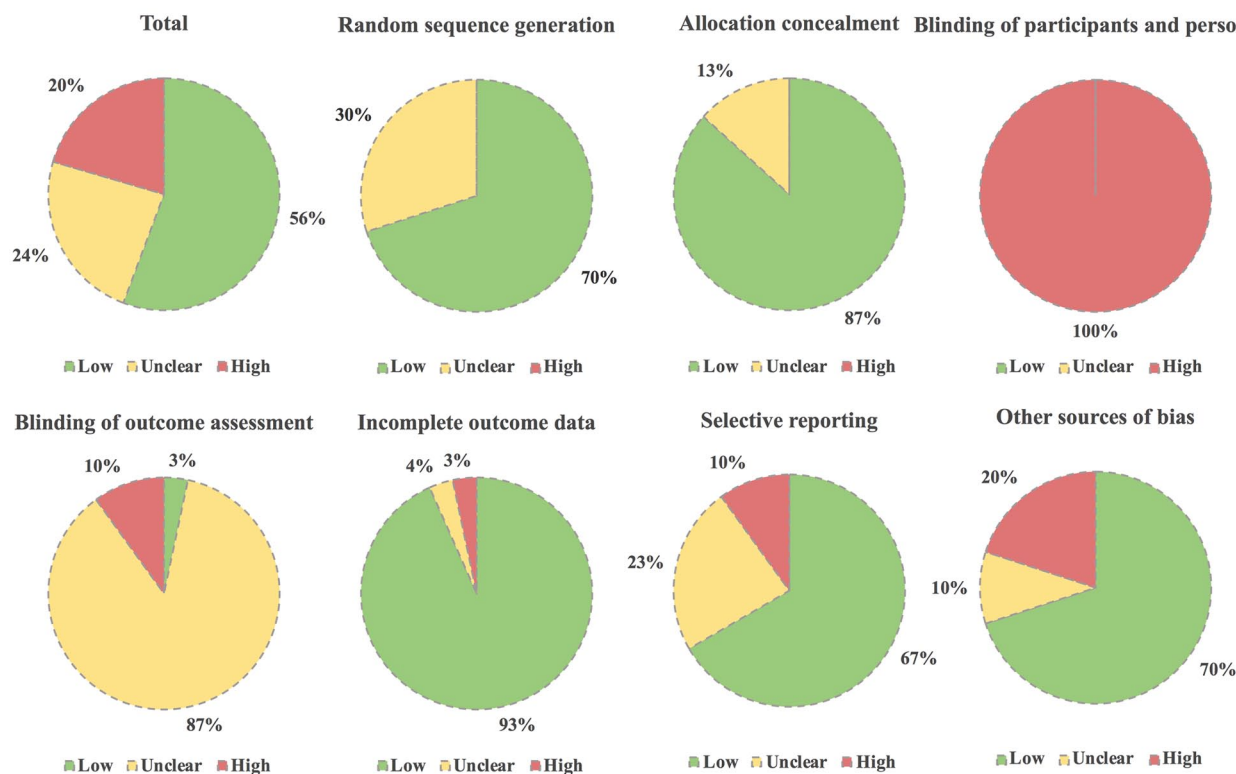
Table 2. (Continued)

Study author(s)	Treatment	Node-1	Node-2	Sample size	3-year DFS	5-year DFS	3-year OS	5-year OS	DFS HR (Node-1)	DFS HR (Node-2)	OS HR (Node-1)	OS HR (Node-2)
Papadimitriou 2011	FOLFIRI (9M)	FOLFIRI	FOLFIRI	441	78.0%	70.0%	88.0%	78.0%	0.94 (95% CI, 0.74-1.19)	0.94 (95% CI, 0.74-1.19)	0.90 (95% CI, 0.69-1.17)	0.90 (95% CI, 0.69-1.17)
	FU/FA (8M)	FU/FA-RP	FU/FA-RP	432	76.0%	68.0%	86.0%	76.0%				
Yothers 2011	FLOX (6M)	mFOLFOX6 (6M)	FP (6M)	1200	75.9%	69.4%	87.3%	80.2%	0.82 (95% CI, 0.72-0.93)	0.82 (95% CI, 0.72-0.93)	0.88 (95% CI, 0.75-1.02)	0.88 (95% CI, 0.75-1.02)
	FU/FA (6M)	FU/FA-RP	FU/FA-RP	1207	71.5%	64.2%	86.3%	78.4%				
Van Cutsem 2009	FOLFIRI (6M)	FOLFIRI	FOLFIRI	1485	69.3%	63.8%	86.4%	78.1%	0.89 (95% CI, 0.79-1.00)	0.89 (95% CI, 0.79-1.00)	NA	NA
	LV5FU2 (6M)	LV5FU2 (6M)	LV5FU2 (6M)	1497	67.1%	61.0%	85.0%	76.4%				
Ychou 2009	FOLFIRI (6M)	FOLFIRI	FOLFIRI	199	51.0%	45.7%	74.7%	61.0%	1.12 (95% CI, 0.85-1.47)	1.12 (95% CI, 0.85-1.47)	1.20 (95% CI, 0.87-1.67)	1.20 (95% CI, 0.87-1.67)
	LV5FU2 (6M)	LV5FU2 (6M)	LV5FU2 (6M)	198	60.0%	51.5%	78.9%	67.0%				
Popov 2008	Raltitrexed (6M)	Raltitrexed (6M)	Raltitrexed (6M)	952	NA	46.7%	71.7%	61.9%	1.14 (95% CI, 0.99-1.32)	1.14 (95% CI, 0.99-1.32)	1.04 (95% CI, 0.87-1.24)	1.04 (95% CI, 0.87-1.24)
	FU/FA (6M)	FU/FA-MC	FU/FA-MC	969	50.9%	50.9%	73.1%	62.3%				
André 2007	LV5FU2 (6M)	LV5FU2 (6M)	LV5FU2 (6M)	452	72.7%	67.1%	86.0%	79.7%	1.01 (95% CI, 0.78-1.31)	1.01 (95% CI, 0.78-1.31)	1.02 (95% CI, 0.75-1.40)	1.02 (95% CI, 0.75-1.40)
	FU/FA (6M)	FU/FA-RP	FU/FA-RP	453	73.9%	67.4%	88.0%	79.7%				
Saltz 2007	FU/FA (8M)	FU/FA-RP	FU/FA-RP	629	69.0%	61.0%	81.0%	71.0%	0.91 (95% CI, 0.75-1.11)	0.91 (95% CI, 0.75-1.11)	0.94 (95% CI, 0.75-1.19)	0.94 (95% CI, 0.75-1.19)
	FOLFIRI (7M)	FOLFIRI	FOLFIRI	635	66.0%	59.0%	80.0%	68.0%				
Lembersky 2006	UFT/LV (6M)	UFT/LV (6M)	UFT/LV (6M)	781	74.5%	67.0%	88.3%	78.5%	1.00 (95% CI, 0.85-1.19)	1.00 (95% CI, 0.85-1.19)	1.01 (95% CI, 0.83-1.25)	1.01 (95% CI, 0.83-1.25)
	FU/FA (6M)	FU/FA-RP	FU/FA-RP	770	74.5%	68.2%	85.6%	78.7%				

\*The numbers suggesting integrated numbers of both ECOG-0 and ECOG-1.

For trials studying targeted drugs such as cetuximab and bevacizumab, results of efficacy were based on unselected population to maintain histopathological homogeneity across all eligible trials. For 'region', some trials were completed by several Western countries instead of a sole nation, therefore 'Western' was used under this circumstance. Details of the rationality to name and constitute each node within our meta-analysis were depicted in the Methods section. The HR was the result of upper arm versus lower arm in each trial, except for those that were specifically labelled, such as 'm3 versus m6'.

6M, 6-month; 12M, 12-month; CI, confidence interval; DFS, disease-free survival; ECOG, European Cooperative Oncology Group; F3/F6, FOLFOX4 (3M/6M); FA, FA-MC, Mayo Clinic regimen; FA-RP, Roswell Park regimen; FLOX, FB, FOLFOX4/bevacizumab; FOLFIRI, FOLFIRI; FOLF, FOLF, hazard ratio; m3/m6, mFOLFOX6 (3M/6M); LV, LV; M/F, male/female; NA, not available; OS, overall survival; R/L, right/left; SOX, UFT, X3/X6, XELOX (3M/6M); XB, XELOX/bevacizumab; XELOX, capecitabine plus oxaliplatin.



**Figure 2.** Risk-of-bias assessment of eligible trials.

different nodes to evaluate the homogeneity inside each node (eTable 3), especially their baseline DFS rate. For node 'FU/FA-MC', 'FOLFOX4 (3M)', 'FOLFOX4 (6M)', 'SOX (6M)', 'capecitabine (12M)', 'mFOLFOX6 (3M)', 'mFOLFOX6 (6M)', 'S-1 (6M)', 'UFT/LV (18M)', 'F/bevacizumab (12M)' and 'Raltitrexed (6M)', baseline survival rates were relatively comparable within each node, thus legitimizing transitivity across the network. For node 'FOLFIRI', 'FU/FA-RP', 'XELOX (3M)', 'XELOX (6M)', 'UFT/LV (6M)', 'capecitabine (6M)', 'LV5FU2 (6M)' and 'F/cetuximab (6M)', each node had one or two trials featuring slightly incomparable baseline survival rates with other trials in the same node, and those trials would be removed in the sensitivity analysis subsequently (eTable 3). Therefore, homogeneity inside each node of our network meta-analysis was guaranteed, assuming there was transitivity.

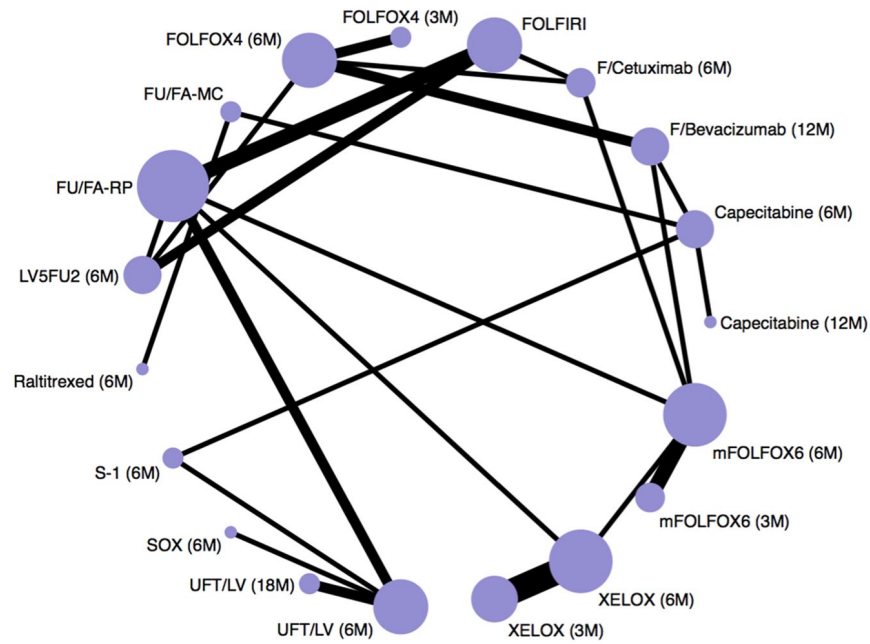
**Consistency and heterogeneity.** Five closed loops were found inside our network meta-analysis. The 95% CI of IF of all closed loops contained zero, suggesting there was no inconsistency between direct and indirect results (eTable 4).  $Q$  static for assessing inconsistency ( $Q$  inconsistency) also

implied there was no inconsistency within the network ( $Q$  inconsistency:  $p=0.262$ ). In terms of statistical heterogeneity, both  $I^2$  static ( $I^2=0\%$ ) and  $Q$  static ( $Q$  heterogeneity:  $p=0.969$ ) hinted there was no significant heterogeneity across eligible trials.

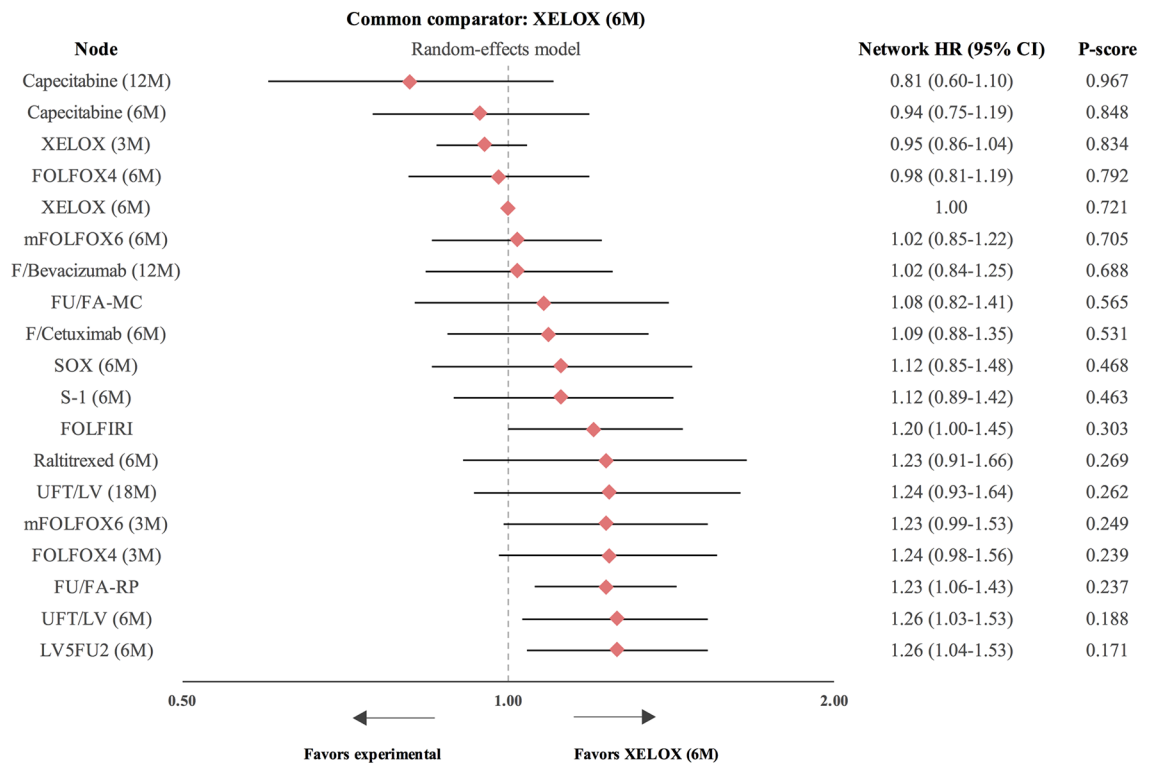
**Publication bias.** There was no publication bias amid all included trials due to symmetrical distribution of effect sizes by funnel plot (eFigure 1).

**Network calculation.** By Node-1 classification, 'capecitabine (12M)' [network HR 0.81 (0.60–1.10),  $p$  score = 0.967] was the highest-ranking regimen that displayed non-inferiority against common comparator 'XELOX (6M)' together with 'XELOX (3M)' [network HR 0.95 (0.86–1.04),  $p$  score = 0.834; Figures 4 and 5].

**Sensitivity analysis.** First, by Node-2 classification, which integrated all fluoropyrimidine-plus-platinum regimens, 'capecitabine (12M)' topped the entire ranking and was non-inferior to common comparator 'FP (6M)' [network HR 0.80 (0.61–1.05),  $p$  score = 0.981; eFigure 2]. The network remained in low heterogeneity and high consistency inside despite changing node classifications (data not shown). Besides, 'FP (6M)'

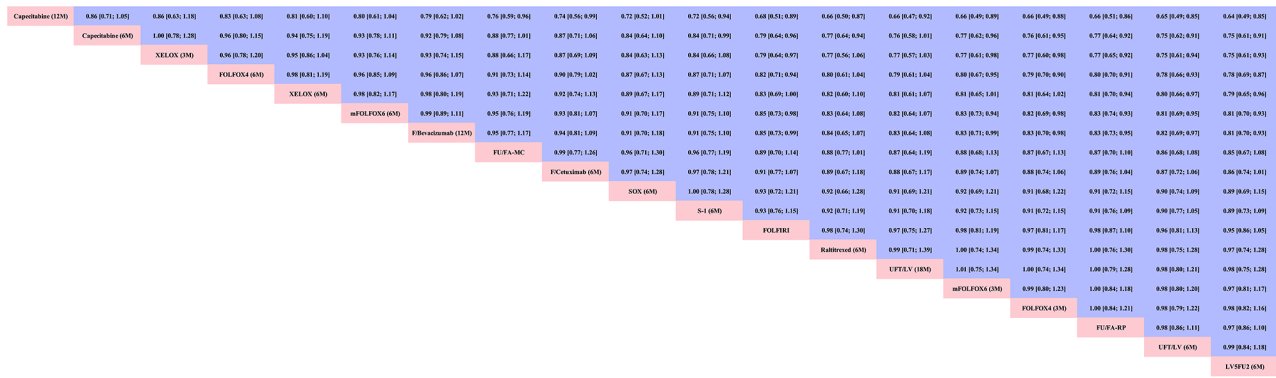


**Figure 3.** Network structure plot of disease-free survival. The size of node implicated the number of studies within each node, while the width of the line was proportional to the amount of mutual direct comparisons. 6M, 6-month regimen; 12M, 12-month regimen; FA-MC, Mayo Clinic regimen; FA-RP, Roswell Park regimen; XELOX, capecitabine plus oxaliplatin.



**Figure 4.** Network forest plot of disease-free survival. 6M, 6-month regimen; 12M, 12-month regimen; CI, confidence interval; FA, ; FA-MC, Mayo Clinic regimen; FA-RP, Roswell Park regimen; FOLFIRI, ; FOLFOX, ; FU, ; HR, hazard ratio; mFOLFOX6, ; LV, ; SOX, ; UFT, ; XELOX, capecitabine plus oxaliplatin.





**Figure 5.** Network league table of disease-free survival. Treatments were hierarchically ranked according to their *p* score. The higher position a node was located at, the better survival outcome it could have. Values situated at the intersection of a specific column and row were the network-effect sizes (HR and 95% CI) of row-defining node versus column-defining node. 6M, 6-month regimen; 12M, 12-month regimen; FA, ; FA-MC, Mayo Clinic regimen; FA-RP, Roswell Park regimen; FOLFIRI, ; FOLFIRI, ; FU, ; mFOLFOX6, ; HR, hazard ratio; LV, ; SOX, ; UFT, ; XELOX, capecitabine plus oxaliplatin.

demonstrated borderline superiority against ‘FP (3M)’ [network HR 1.08 (1.00–1.16)]. Second, by deleting trials that displayed incomparable baseline survival rates with other counterparts in the same node, as well as trials that contained smaller sample sizes (less than 200; eTable 3), ‘XELOX (3M)’ was the only node displaying non-inferiority against ‘XELOX (6M)’ (data not shown). Third, since the definitions of DFS were not always consistent among all eligible studies, we additionally deleted trials defining DFS similar to recurrence-free survival, which only counted tumour recurrences but not secondary malignancies as events (Supplemental Material). As a result, ‘XELOX (3M)’ was still the only non-inferior regimen compared with ‘XELOX (6M)’ in the hierarchy (data not shown). All these suggested that network outcomes for DFS were stable and solid.

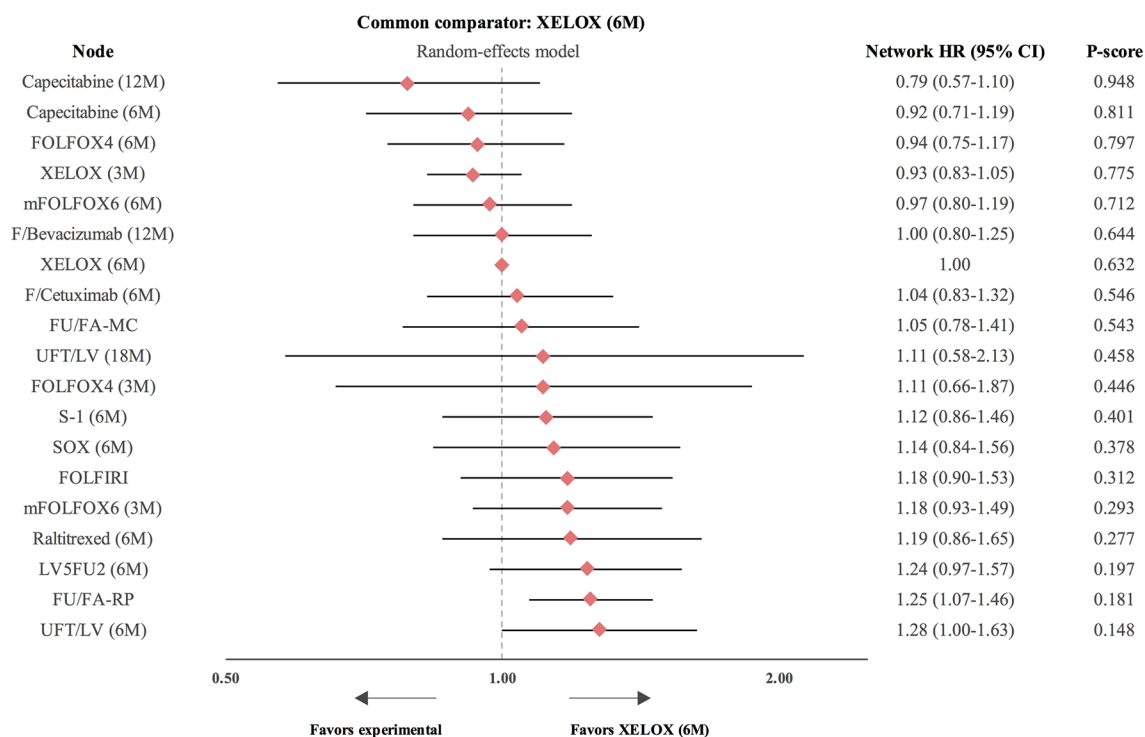
*Network subgroup analysis.* Via Node-1 classification, we only calculated subgroup data for stage 3 due to insufficient data of high-risk stage 2 cases (eTables 5 and 6). Here, ‘capecitabine (12M)’ [network HR 0.79 (0.57–1.10), *p* score = 0.948] still ranked as the top node and demonstrated non-inferiority to ‘XELOX (6M)’, together with ‘XELOX (3M)’ [network HR 0.93 (0.83–1.05), *p* score = 0.775; Figure 6]. By Node-2 classification, ‘capecitabine (12M)’ topped the entire hierarchy among stage 3 patients and displayed non-inferiority against ‘FP (6M)’ [network HR 0.80 (0.61–1.06), *p* score = 0.965; eFigure 3]. On the other hand, ‘FP (3M)’ failed to show non-inferiority against ‘FP (6M)’ among stage 3 and high-risk stage 2 patients [stage 3: network HR

1.07 (0.99–1.16), *p* score = 0.551; high-risk stage 2: network HR 1.14 (0.95–1.38), *p* score = 0.207; eFigure 4].

*Pairwise subgroup analysis.* Based on the above-mentioned network calculation results, we did more specific pairwise meta-analyses to eliminate certain heterogenic factors that might bias comparisons between 3-month and 6-month regimens. Here, only the five major largescale RCTs, including ACHIEVE, HORG-IDEA, IDEA, SCOT and TOSCA were included (results of CALGB/SWOG 80702 had not been formally published in a journal or meeting abstract). Among low-risk stage 3 patients, the ‘XELOX (3M)’ regimen was significantly better than the ‘XELOX (6M)’ regimen [pairwise HR 0.78 (0.63–0.97), *p* = 0.02], while both ‘mFOLFOX6 (3M)’ [pairwise HR 1.16 (0.95–1.42), *p* = 0.15] and ‘FP (3M)’ [pairwise HR 1.03 (0.92–1.16), *p* = 0.60] could not demonstrate non-inferiority against their 6-month counterparts. Within high-risk stage 3 patients, ‘XELOX (3M)’ [pairwise HR 1.05 (0.90–1.23), *p* = 0.50] failed to show non-inferiority against its longer-duration counterpart, while ‘mFOLFOX6 (3M)’ [pairwise HR 1.31 (1.11–1.55), *p* = 0.002] and ‘FP (3M)’ [pairwise HR 1.14 (1.01–1.29), *p* = 0.03] were significantly worse than ‘mFOLFOX6 (6M)’ and ‘FP (6M)’, respectively.

*Secondary endpoint: overall survival*

25 trials were included in the OS calculation (Tables 1 and 2). Regardless of Node-1 or Node-2



**Figure 6.** Network forest plot of disease-free survival among stage 3 patients. 6M, 6-month regimen; 12M, 12-month regimen; CI, confidence interval; FA, ; FA-MC, Mayo Clinic regimen; FA-RP, Roswell Park regimen; FOLFIRI, ; FOLFOX, ; FU, ; HR, hazard ratio; mFOLFOX6, ; LV, ; SOX, ; UFT, ; XELOX, capecitabine plus oxaliplatin.

classification, ‘capecitabine (12M)’ was the best node among all analyzed counterparts, displaying non-inferiority against common comparator ‘XELOX (6M)’ [Node-1: network HR 0.71 (0.48–1.06), *p* score = 0.976; Node-2: network HR 0.72 (0.51–1.02), *p* score = 0.990]. Moreover, ‘FP (3M)’ also demonstrated non-inferiority against ‘FP (6M)’ [network HR 1.01 (0.93–1.08), *p* score = 0.771; eFigures 5 and 6]. Overall, inconsistency and heterogeneity remained at a very low level (data not shown).

**Secondary endpoint: adverse events**

Details of safety profile are displayed in eTable 7. Node-2 classification was used here to present network results, since not all included types of adverse events provided separate data for Node-1 classification. ‘FP (3M)’ was significantly better than its 6-month counterpart with respect to peripheral sensory neuropathy [network RR 0.31 (0.23–0.42)], thrombocytopenia [network RR 0.68 (0.47–0.98)] and fatigue [network RR 0.56 (0.32–0.95)], while it being non-inferior to the 6-month regimen regarding neutropenia [network

RR 0.74 (0.45–1.21)], leucopenia [network RR 0.81 (0.57–1.13)] and diarrhoea [network RR 0.79 (0.61–1.02)]. For anaemia [network RR 1.31 (0.37–4.62)], anorexia [network RR 1.14 (0.70–1.85)] and nausea/vomiting [network RR 1.09 (0.81–1.47)], ‘FP (3M)’ did not display non-inferiority against ‘FP (6M)’. The 12-month capecitabine monotherapy only exhibited superiority against 6-month oxaliplatin-based chemotherapy in terms of leucopenia [network RR 0.02 (0.00–0.94)] and thrombocytopenia [network RR 0.01 (0.00–0.66)], but failed to display non-inferiority among other major adverse events.

**Discussion**

Adjuvant systemic treatments for resectable colon cancer have drawn a lot of academic attention during the past decade. Currently, XELOX and FOLFOX regimens have been widely accepted as the standard options, especially for high-risk stage 2 and stage 3 patients.<sup>3,4,6</sup> However, the evidence is mainly based on pairwise RCTs, and sometimes it is difficult to make accurate comparisons among so many regimens, especially since novel

medications are constantly introduced to the market. Therefore, network meta-analysis is a necessity in this situation.

In terms of DFS, 12-month capecitabine monotherapy topped the hierarchy and showed non-inferiority against the 6-month XELOX regimen, together with the 3-month XELOX regimen, which ranked in the third place; however, with the most condensed interval-of-effect size. Similar results were obtained in terms of subgroup analysis among stage 3 patients. The more specific pairwise subgroup analysis suggested that 3-month XELOX was better than its 6-month counterpart among low-risk stage 3 patients, while none of the 3-month regimens displayed non-inferiority against 6-month treatments among high-risk stage 3 patients, and the 6-month mFOLFOX6 regimen was even significantly better than its 3-month counterpart. However, if we applied Node-2 classification by integrating all fluoropyrimidine-plus-platinum regimens together, 12-month capecitabine monotherapy rather than 3-month oxaliplatin-based chemotherapy became non-inferior to 6-month oxaliplatin-based chemotherapy among stage 3 patients, while no matter among low-risk or high-risk stage 3 patients, 3-month oxaliplatin-based chemotherapy failed to reach non-inferiority against its 6-month counterpart. This implies that types of fluoropyrimidine and schedules might have impact on survival benefits of 3-month treatment. For high-risk stage 2 patients, 6-month oxaliplatin-based chemotherapy was still the optimal option, since none of the included regimens seemed to be at least non-inferior to it. Nevertheless, more original trials are warranted in the future because network calculation of this part of subgroup analysis is only based on Node-2 classification, due to inadequate data of individual arms, which could possibly be biased by different types of fluoropyrimidine and schedule.

Regarding OS, 12-month capecitabine monotherapy topped the hierarchy, exhibiting non-inferiority against 6-month XELOX, while 3-month XELOX failed to do so. However, *via* Node-2 classification, both 12-month capecitabine monotherapy as well as 3-month oxaliplatin-based chemotherapy displayed non-inferiority against 6-month oxaliplatin-based chemotherapy. This might be caused mainly by the fact that most trials investigating 3-month *versus* 6-month regimens took DFS as the primary endpoint and did not report OS data, which resulted in the wide-range network-effect size of the 3-month XELOX regimen and crossed the non-inferiority margin. Therefore, for 12-month

capecitabine monotherapy and 3-month fluoropyrimidine-plus-platinum regimens, more studies are needed to further investigate their OS benefits before making reliable conclusions. Regarding adverse events, although we could only make network analysis based on Node-2 classification, 3-month oxaliplatin-based chemotherapy was at least non-inferior to its 6-month counterpart among the most of common adverse events, especially peripheral sensory neuropathy, thrombocytopenia and fatigue, which the 3-month regimen was significantly better for. This result is also anticipated and easily understood, since shortened periods of chemotherapeutic treatments cause fewer detrimental effects on recipients. Nevertheless, 12-month capecitabine monotherapy only exhibited superiority against 6-month oxaliplatin-based chemotherapy in terms of leucopenia and thrombocytopenia, while failing to display non-inferiority among other major adverse events. This may probably hint that long-haul chemotherapy, despite of capecitabine monotherapy, will still worsen tolerability among treatment recipients. However, since there was only one trial reporting 12-month capecitabine monotherapy so far, we should also take the possible underpower of statistical calculation into account while making judgement on its real safety effects.

Current NCCN guideline on colon cancer suggests that the 3-month XELOX regimen could be used among low-risk stage 3 patients due to its non-inferiority against its 6-month counterpart, while 6-month oxaliplatin-based chemotherapy is still a more reliable choice regarding high-risk stage 3 patients. It also supports the application of 6-month oxaliplatin-based chemotherapy among high-risk stage 2 patients based on current evidence.<sup>3</sup> Meanwhile, reported by Sobrero and colleagues in the 2020 ASCO annual meeting, the latest pooled analysis of six IDEA trials also suggests that 3-month oxaliplatin-based regimens are non-inferior to 6-month regimens, especially among low-risk stage 3 patients. Although our network meta-analysis failed to make more groundbreaking discoveries when compared with current guidelines, this was still the first systematic review and network meta-analysis in this field, which might provide useful hints for design of largescale RCTs in the future. The confirmation by our meta-analysis might further support the use of corresponding regimens in the future, which therefore should be recognized as the major significance and novelty of our work. Meanwhile, somewhat surprisingly, 12-month capecitabine monotherapy

also displayed non-inferiority against current standard treatments, despite lacking Western data on its suitability, as well as its possibly higher toxicity and worse compliance. The ranking of 12-month capecitabine monotherapy is the product of indirect network calculation that could also be regarded as a possible topic in future design of randomized trials.

Although our systematic review and network meta-analysis were rigorously designed and conducted, there were still some limitations. First, although all eligible trials were proven to be clinically comparable without significant heterogeneity, and sensitivity analysis had also been conducted to ensure the homogeneity of baseline survival rates in the same node, impact by underlying heterogeneity could not be fully eliminated, such as different regions, races and extents of lymphadenectomy. Therefore, future updates, especially individual patient data network meta-analyses, are welcomed. Second, we still need more trials (including CALGB/SWOG 80702 trial) to enhance statistical power, as well as provide more subgroup analyses for better clinical interpretations, such as subgroup data among low-risk and high-risk stage 3 patients, respectively.

Taken together, with its at least non-inferior survival benefit and even better safety profile, 3-month XELOX treatment could be an alternative option of traditional 6-month regimen among low-risk stage 3 patients. Among high-risk stage 3 patients, 6-month oxaliplatin-based regimens still seem more competitive. For high-risk stage 2 cases, we still recommend 6-month oxaliplatin-based regimens until more compelling evidence emerges. In addition, due to inadequate statistical power and possibly higher toxicity, clinical application of 12-month capecitabine monotherapy should still be undertaken with caution, despite its top ranking, especially among non-Asian countries.

### Acknowledgement

We thank all members in our department for offering clinical and methodological suggestions during the entire performance of our meta-analysis.

### Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: this meta-analysis was funded by National Natural Science Foundation of China (81902487) to Ji Cheng and National Natural Science Foundation of China (81874184) to Kaixiong Tao.


### Conflict of interest statement

The authors declare that there is no conflict of interest.

### Author contributions

Study design: Ji Cheng, Guobin Wang and Kaixiong Tao. Manuscript writing and revision: Ji Cheng, Kaixiong Tao and Kailin Cai. Literature retrieval: Ji Cheng and Jinbo Gao. Discretion of eligibility: Ji Cheng and Xiaoming Shuai. Quality assessment: Ji Cheng and Jinbo Gao. Data extraction: Ji Cheng and Xiaoming Shuai. Statistical analysis: Ji Cheng and Kaixiong Tao.

### ORCID iD

Ji Cheng  <https://orcid.org/0000-0002-7673-9157>

### Supplemental material

Supplemental material for this article is available online.

### References

1. Iveson TJ, Kerr RS, Saunders MP, *et al.* 3 versus 6 months of adjuvant oxaliplatin-fluoropyrimidine combination therapy for colorectal cancer (SCOT): an international, randomised, phase 3, non-inferiority trial. *Lancet Oncol* 2018; 19: 562–578.
2. Siegel RL, Miller KD and Jemal A. Cancer statistics, 2019. *CA Cancer J Clin* 2019; 69: 7–34.
3. National Comprehensive Cancer Network. Clinical practice guidelines in oncology. Colon cancer, version 2. 2019, <https://www.nccn.org> (2019, accessed May 15 2019).
4. Committee CCC. Chinese guidelines on the management of colorectal cancer (2019.V1). 2019.
5. Labianca R, Nordlinger B, Beretta GD, *et al.* Early colon cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2013; 24(Suppl. 6): vi64–vi72.
6. Hashiguchi Y, Muro K, Saito Y, *et al.* Japanese Society for Cancer of the Colon and Rectum (JSCCR) guidelines 2019 for the treatment of colorectal cancer. *Int J Clin Oncol*. Epub ahead of print 15 June 2019. DOI: 10.1007/s10147-019-01485-z.
7. Souglakos J, Boukovinas I, Kakolyris S, *et al.* Three versus six months adjuvant FOLFOX or CAPOX for high risk stage II and stage III colon cancer patients: the efficacy results of Hellenic

- Oncology Research Group (HORG) participation to the International Duration Evaluation of Adjuvant chemotherapy (IDEA) project. *Ann Oncol*. Epub ahead of print June 2019. DOI: 10.1093/annonc/mdz193.
8. Andre T, Vernerey D, Mineur L, *et al.* Three versus 6 months of oxaliplatin-based adjuvant chemotherapy for patients with stage III colon cancer: disease-free survival results from a randomized, open-label, International Duration Evaluation of Adjuvant (IDEA) France, phase III trial. *J Clin Oncol* 2018; 36: 1469–1477.
  9. Sobrero A, Lonardi S, Rosati G, *et al.* FOLFOX or CAPOX in stage II to III colon cancer: efficacy results of the Italian three or six colon adjuvant trial. *J Clin Oncol* 2018; 36: 1478–1485.
  10. Yoshino T, Yamanaka T, Oki E, *et al.* Efficacy and long-term peripheral sensory neuropathy of 3 vs 6 months of oxaliplatin-based adjuvant chemotherapy for colon cancer: the ACHIEVE phase 3 randomized clinical trial. *JAMA Oncol*. Epub ahead of print 12 September 2019. DOI: 10.1001/jamaoncol.2019.2572.
  11. Cheng J, Cai M, Shuai X, *et al.* Multimodal treatments for resectable gastric cancer: a systematic review and network meta-analysis. *Eur J Surg Oncol*. Epub ahead of print 8 June 2019. DOI: 10.1016/j.ejso.2019.06.010.
  12. Tierney JF, Stewart LA, Ghersi D, *et al.* Practical methods for incorporating summary time-to-event data into meta-analysis. *Trials* 2007; 8: 16.
  13. Cheng J, Cai M, Shuai X, *et al.* Systemic therapy for previously treated advanced gastric cancer: a systematic review and network meta-analysis. *Crit Rev Oncol Hematol* 2019; 143: 27–45.
  14. Cipriani A, Higgins JP, Geddes JR, *et al.* Conceptual and technical challenges in network meta-analysis. *Ann Intern Med* 2013; 159: 130–137.
  15. Cheng J, Cai M, Shuai X, *et al.* Comparative efficacy and tolerability of antiemetic prophylaxis for adult highly emetogenic chemotherapy: a network meta-analysis of 143 randomized controlled trials. *Int J Cancer* 2018; 142: 1067–1076.
  16. Cheng J, Cai M, Shuai X, *et al.* Multimodal treatments for resectable esophagogastric junction cancer: a systematic review and network meta-analysis. *Ther Adv Med Oncol* 2019; 11: 1758835919838963.
  17. Palmer SC, Mavridis D, Navarese E, *et al.* Comparative efficacy and safety of blood pressure-lowering agents in adults with diabetes and kidney disease: a network meta-analysis. *Lancet* 2015; 385: 2047–2056.
  18. Grothey A, Sobrero AF, Shields AF, *et al.* Duration of adjuvant chemotherapy for stage III colon cancer. *N Engl J Med* 2018; 378: 1177–1188.

Visit SAGE journals online  
[journals.sagepub.com/  
 home/tam](http://journals.sagepub.com/home/tam)

 SAGE journals