

Prognosis comparison between wait and watch and surgical strategy on rectal cancer patients after treatment with neoadjuvant chemoradiotherapy: a meta-analysis

Kai Pang , Quan Rao, Shengqi Qin, Lan Jin, Hongwei Yao and Zhongtao Zhang

Abstract

Background: After achieving a clinical complete response through neoadjuvant chemoradiotherapy, a nonoperative management approach for rectal cancer patients known as Wait and Watch (W&W) has gained increasing attention. However, the W&W strategy has been related to higher local recurrence and ambiguous long-term survival. This meta-analysis compared key prognosis indicators between W&W and surgical treatment in an effort to clarify some long-standing points of confusion.

Methods: Pubmed, Web of Science, EMBase, Cochrane Library were searched for relevant researches comparing W&W with surgery treatment, with a time criteria set from 1 January 2002 to 4 July 2019. Endpoints were 2-year local regrowth/recurrence, 2-year distant metastasis (plus local regrowth/recurrence), 3- and 5-year disease-free survival (DFS), and overall survival (OS).

Results: In total, nine studies with 801 patients were enrolled, of which 348 were managed by W&W and 453 by surgery. Surgery patients were further divided into a pathological complete response (pCR) group (all included patients achieved pCR) and a surgery group (consisting of both pCR and non-pCR patients without deliberate screening). Compared with the surgery group, W&W patients have higher 3- and 5-year OS, and are not inferior on 2-year local regrowth (LR), 2-year distant metastasis (DM)/DM+LR, and 3- and 5-year DFS. On the other hand, compared with the pCR group, the W&W group is inferior on 2-year LR, 3- and 5-year DFS, and 5-year OS, and not inferior on 2-year DM/DM+LR and 3-year OS.

Conclusions: In contrast with patients undergoing surgical treatment, the W&W group has higher 3- and 5-year OS, and is not inferior on other major prognostic indicators, which, however, is based on the fact that the tumor stage in the W&W group is generally earlier. *Versus* surgically treated patients who acquired pCR, W&W group is inferior on all major prognostic indicators except 2-year DM/DM+LR and 3-year OS. Additionally, by comparison of cCR definitions across different studies, we conclude that implementation of the strictest cCR criteria is critical for W&W patients to acquire maximum prognostic benefit.

Keywords: distant metastasis, local recurrence, rectal cancer, surgery, survival, watch and wait

Received: 23 July 2019; revised manuscript accepted: 6 November 2019.

Introduction

Neoadjuvant chemoradiotherapy (NCRT) has become part of the standard treatment for locally advanced rectal cancer (LARC), which can

potentially lead to shrinkage of tumor volume, downgrade of tumor stage, increase of R0 resection rate and anus preservation rate, as well as decrease of local recurrence, ultimately helping

Ther Adv Gastroenterol

2019, Vol. 12: 1–12

DOI: 10.1177/
1756284819892477

© The Author(s), 2019.
Article reuse guidelines:
[sagepub.com/journals-
permissions](https://sagepub.com/journals-permissions)

Correspondence to:
Zhongtao Zhang
Department of General
Surgery, Beijing
Friendship Hospital,
Capital Medical University;
Beijing Key Laboratory
of Cancer Invasion and
Metastasis Research and
National Clinical Research
Center for Digestive
Diseases, 95[#] Yong-an
Road, Xi-Cheng District,
Beijing 100050, China
Zhongtao.zq@139.com

Kai Pang
Quan Rao
Shengqi Qin
Lan Jin
Hongwei Yao
Department of General
Surgery, Beijing
Friendship Hospital,
Capital Medical University,
Beijing Key Laboratory
of Cancer Invasion and
Metastasis Research and
National Clinical Research
Center for Digestive
Diseases, Beijing, China

some patients to achieve clinical complete response (cCR) or even pathological complete response (pCR). NCRT combined with total mesorectal excision (TME) is currently the gold standard treatment for LARC.¹ However, radical resection surgeries are sometimes accompanied with complications that severely influence patients' quality of life, like dysfunction of urination and sex.² After NCRT, around 20–30% of patients can achieve cCR, among whom postoperative pathological evidence indicates 5–44% achieved pCR. And patients that achieved pCR obtained significant benefits regarding local control and long-term survival.^{3,4} In 2002, Nakagawa and colleagues put forward the idea of nonsurgical treatment for patients with cCR after NCRT.⁵ In 2004, Habr-Gama and colleagues reported a clinical research on a wait-and-watch (W&W) strategy for the first time.⁶ Before long, Appelt and colleagues discovered that cCR rate can be increased by high-dose chemoradiotherapy.⁷ Later, Habr-Gama and colleagues reported that, for patients with regrowth after adopting a W&W strategy, timely surgeries can still effectively control local regrowth (LR) of tumors.⁸

In recent years, whether W&W strategy can be widely applied among cCR patients so as to avoid surgical trauma has become a focus of debate among physicians. In this meta-analysis, we compared key prognosis indicators between W&W strategy and surgical treatment in an effort to clarify some issues that have long been debated among colorectal surgeons. The comparison focused on the following indicators: 2-year LR/recurrence (LR), 2-year distant metastasis (DM) or DM+LR, and 3- and 5-year disease-free survival (DFS), and 3 and 5-year overall survival (OS).

Materials and methods

Registration information

This meta-analysis is registered on PROSPERO (<https://www.crd.york.ac.uk/prospero>). The protocol of this research can be accessed on PROSPERO website with the registration number CRD42019141601.

Literature search strategy

Literature search was conducted in the following databases: PubMed, Web of Science, EMBASE, Cochrane Library. Time period is set from

2002.1.1 to 2019.7.4. Search strategy is as follows: RECTAL and CARCINOMA or CANCER or NEOPLASM and WAIT and WATCH or SEE or WATCHFUL WAITING or NON-OPERATIVE and CHEMORADIOTHERAPY. Language was restricted to English only.

Inclusion and exclusion criteria

Inclusion criteria: subjects of studies must be rectal cancer patients receiving long-course NCRT; there must be comparison between a surgery group and a W&W group; research must contain sufficient data on relevant indicators.

Exclusion criteria: research only on W&W patients, without comparison with surgery patients; studies not containing sufficient data on desired indicators; quality of study measured as Low.

Data extraction

The following items were extracted from literature: first author; year of publication; TNM stage of tumors; sample size; chemoradiotherapy plan; type of research; LR; DM or DM+LR; 3- and 5-year DFS; 3- and 5-year OS. Data were extracted independently by two separate researchers. If opinions were inconsistent, a third senior researcher's advice was sought.

Evaluation of research quality

All research works included in this study are non-RCTs (randomized controlled trials). Newcastle-Ottawa Scale (NOS) was applied for quality evaluation.⁹ The evaluation focused on three aspects of each article: selection of study objects, comparability between groups and assessment of result/exposure. NOS is measured with 9 as full score, 1–3 as Low quality, 4–6 as Mediocre quality, 7–9 as High quality.

Statistical processing

Meta-analysis was conducted with STATA software (version 15.0, StataCorp LP, College Station, TX, USA). Risk ratio (RR) with 95% confidence interval (CI) was calculated for dichotomous data. Chi-square test and I^2 test were adopted to evaluate heterogeneity between studies, where $p < 0.10$ indicated significant heterogeneity. A random-effects model was used if

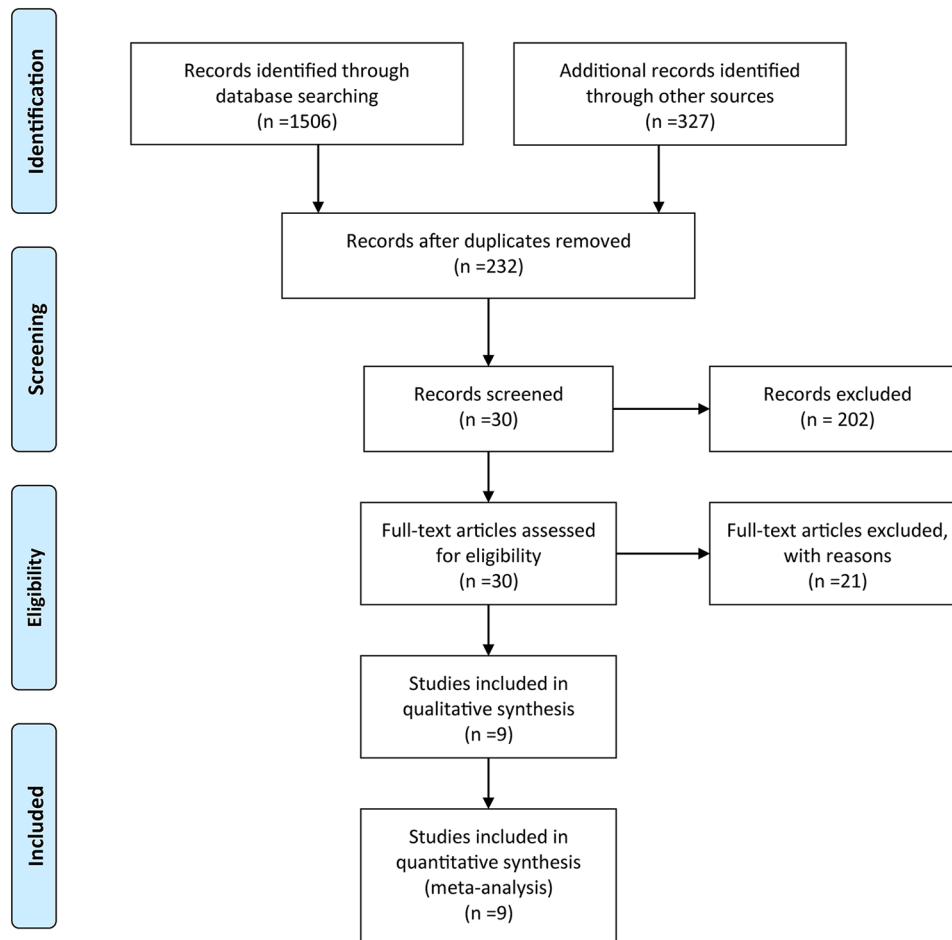


Figure 1. Flowchart of the search strategy.

the test of heterogeneity was significant; otherwise, a fixed-effects model was adopted instead. Egger's test was applied to assess the publication bias, where $p < 0.05$ was considered statistically significant.

Results

Search results

Based on the keywords and filtering criteria above, a total of nine nonrandom controlled trials were included (two retrospective cohort studies and seven prospective cohort studies).^{6,10-18} The selection workflow is illustrated in Figure 1. The included patients total 801, of whom 453 were managed by surgery (hereby designated as *Surgically treated* group). The *Surgically treated* group was further divided into two subgroups: the *pCR* group consists solely of patients whose

postoperation pathology reports indicate *pCR*, formed by combining selected *pCR* cohorts from respective studies; the *Surgery* group consists of both *pCR* and non-*pCR* patients, formed by combining natural cohorts from respective studies without deliberate screening. And the remaining 348 patients were managed by W&W strategy (hereby designated as W&W group, of which 235 patients were compared with the *Surgery* group, and the other 113 patients were compared with the *pCR* group). Basic characteristics of included research are shown in Table 1. The prognostic indicators of interest were collected and are shown in Tables 2 and 3.

Meta-analysis result

Two-year LR. Three articles reported 2-year LR for comparison between the W&W and surgery groups. Pooled analysis demonstrated that the

Table 1. Basic characteristics of included studies.

Study	Study design	NCRT regimen		Patients (n)			Pretreatment TNM \geq III			NOS score
		Chemotherapy	Radiotherapy	W&W	Surgery	pCR	W&W	Surgery	pCR	
Renehan ¹⁸	Prospective	5-FU-based	45 Gy	109	109	–	–	–	–	7
Lee ¹⁵	Prospective	5-FU-based	50.4 Gy	8	28	–	3	15	–	6
Lai ¹⁶	Retrospective	5-FU-based	45–50.4 Gy	18	26	–	7	18	–	6
Li ¹⁴	Prospective	Capecitabine	50 Gy; 25 Gy	30	92	–	16	53	–	6
Habr-Gama ⁶	Prospective	5-FU+LV	50.4 Gy	71	22	–	16	6	–	6
Smith ¹³	Retrospective	5-FU+ Capecitabine	Unspecified	18	–	30	7	–	12	7
Araujo ¹⁷	Prospective	5-FU+LV, Capecitabine	45–50.4 Gy	42	–	69	–	–	–	6
Smith ¹¹	Prospective	5-FU+ Capecitabine	(50.4 \pm 2.75) Gy	32	–	57	18	–	31	7
Mass ⁴	Prospective	Capecitabine	50.4 Gy	21	–	20	13	–	17	7

NCRT, neoadjuvant chemoradiotherapy; W&W, wait and watch.

Table 2. Prognostic indicators of included studies (W&W group versus Surgery group).

Study	2-year LR (%)		2-year DM/ DM+LR (%)		3-year DFS (%)		3-year OS (%)		5-year DFS (%)		5-year OS (%)	
	W&W	Surgery	W&W	Surgery	W&W	Surgery	W&W	Surgery	W&W	Surgery	W&W	Surgery
Renehan ¹⁸	–	–	–	–	88.0	78.0	96.0	87.0	–	–	–	–
Lee ¹⁵	–	–	–	–	75	85.0	–	–	–	–	–	–
Lai ¹⁶	11.1	0	0	3.85	–	–	100	96.2	–	–	100	92.3
Li ¹⁴	3.3	1.1	0	0	93.3	96.7	100	100	90.0	94.3	100	95.6
Habr-Gama ⁶	0	0	1.4	13.6	98.6	86.4	100	90.9	95.8	86.4	100	90.9

DFS, disease-free survival; DM, distant metastasis; LR, local regrowth/recurrence; OS, overall survival.

2-year LR rate in the W&W group and that in the Surgery group were not significantly different (RR = 3.535, 95%CI = 0.448–27.923, $p > 0.05$) (Figure 2a). No heterogeneity was observed among the studies ($p = 0.877$, $I^2 = 0.0\%$) and a fixed-effect model was adopted for analysis. Egger's test for bias of publication was not conducted due to insufficient studies.

Four articles reported 2-year LR for comparison between the W&W and pCR groups. Pooled

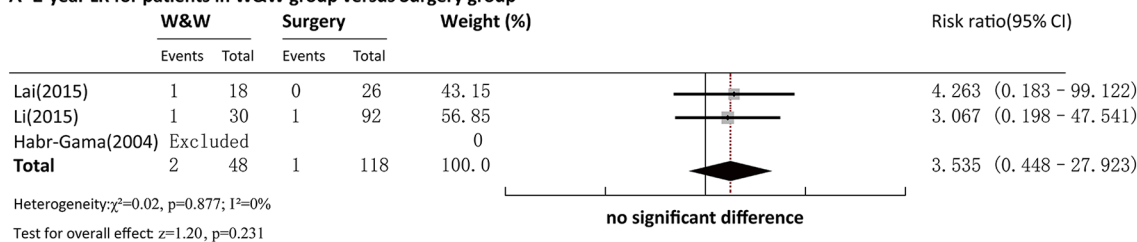
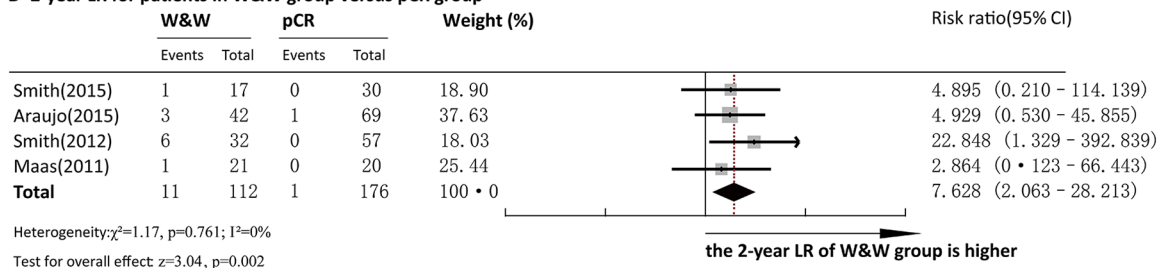
analysis demonstrated that the 2-year LR rate in the W&W group was significantly higher than that in the pCR group (RR = 6.422, 95%CI = 1.619–25.474, $p < 0.01$) (Figure 2b). No heterogeneity was observed among the studies ($p = 0.779$, $I^2 = 0.0\%$), and a fixed-effect model was adopted for analysis. Egger's test for bias of publication was not conducted due to insufficient studies.

Two-year DM or DM+LR. Three articles reported 2-year DM or DM+LR for comparison between

Table 3. Prognostic indicators of included studies (W&W group versus pCR group).

Study	2-year LR (%)		2-year DM/DM+LR (%)		3-year DFS (%)		3-year OS (%)		5-year DFS (%)		5-year OS (%)	
	W&W	pCR	W&W	pCR	W&W	pCR	W&W	pCR	W&W	pCR	W&W	pCR
Smith ¹³	5.6	0	5.6	3.3	88.9	96.7	100	96.7	88.9	96.7	100	96.7
Araujo ¹⁷	7.1	1.4	9.5	7.2	73.8	89.9	–	–	60.9	82.8	71.4	89.9
Smith ¹¹	18.8	0	6.3	1.8	–	–	–	–	–	–	–	–
Mass ⁴	4.8	0	0	0	95.2	95	100	95	–	–	–	–

DFS, disease-free survival; DM, distant metastasis; LR, local regrowth/recurrence; OS, overall survival.

A 2-year LR for patients in W&W group versus Surgery group**B 2-year LR for patients in W&W group versus pCR group****Figure 2.** Forest plot of 2-year LR rate. (a) Comparison between the W&W and Surgery groups. (b) Comparison between the W&W and pCR groups.

LR, local regrowth/recurrence; W&W, wait and watch.

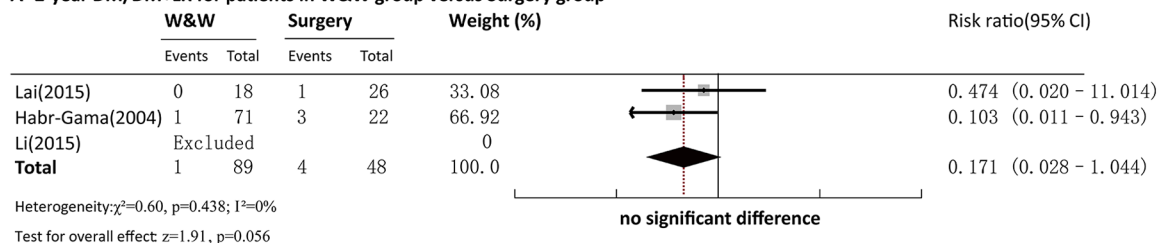
the W&W and Surgery groups. Pooled analysis demonstrated that the 2-year DM or DM+LR rate in the W&W group and that in the Surgery group were not significantly different ($RR = 0.171$, $95\%CI = 0.028-1.044$, $p > 0.05$) (Figure 3a). No heterogeneity was observed among the studies ($p = 0.438$, $I^2 = 0.0\%$), and a fixed-effect model was adopted for analysis. Egger's test for bias of publication was not conducted due to insufficient studies.

Four articles reported 2-year DM or DM+LR for comparison between the W&W and pCR groups. Pooled analysis demonstrated that the 2-year DM or DM+LR rate in W&W group and that in pCR

group were not significantly different ($RR = 1.656$, $95\%CI = 0.593-4.624$, $p > 0.05$) (Figure 3b). No heterogeneity was observed among the studies ($p = 0.765$, $I^2 = 0.0\%$), and a fixed-effect model was adopted for analysis. By Egger's test, there was no bias of publication ($p = 0.48$).

Three-year DFS. Four articles reported 3-year DFS for comparison between the W&W and Surgery groups. Pooled analysis demonstrated that the 3-year DFS in W&W group and that in Surgery group were not significantly different ($RR = 0.731$, $95\%CI = 0.265-2.015$, $p > 0.05$) (Figure 4a). Significant heterogeneity was observed among the studies ($p = 0.097$, $I^2 = 52.5\%$), and a

A 2-year DM/DM+LR for patients in W&W group versus Surgery group



B 2-year DM/DM+LR for patients in W&W group versus pCR group

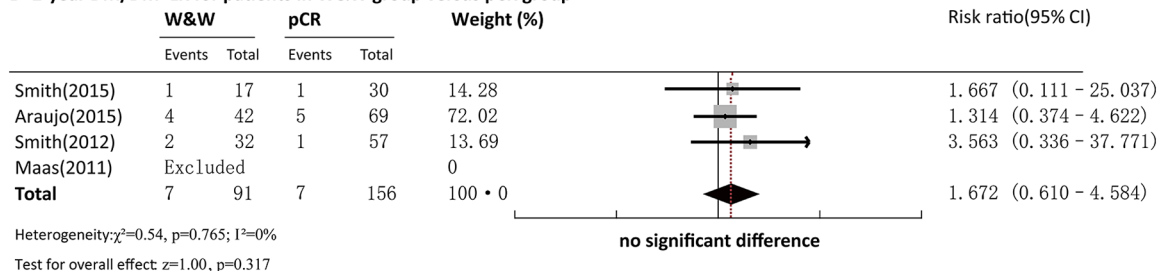
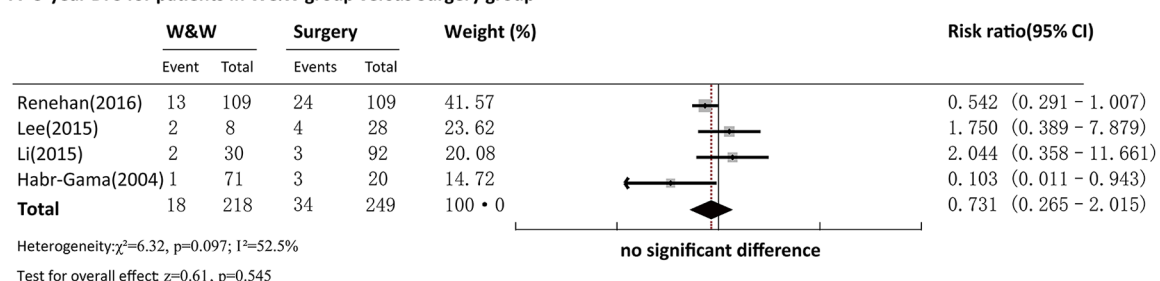


Figure 3. Forest plot of 2-year DM/DM + LR rate. (a) Comparison between the W&W and Surgery groups. (b) Comparison between the W&W and pCR groups. DM, distant metastasis; LR, local regrowth/recurrence; W&W, wait and watch.

A 3-year DFS for patients in W&W group versus Surgery group



B 3-year DFS for patients in W&W group versus pCR group

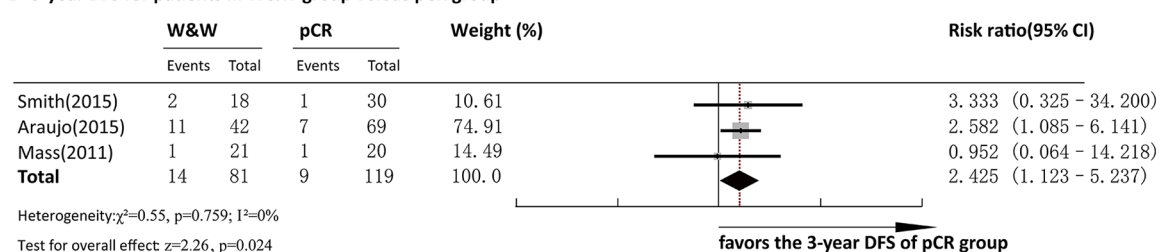


Figure 4. Forest plot of 3-year DFS. (a) Comparison between the W&W and Surgery groups. (b) Comparison between the W&W and pCR groups. DFS, disease-free survival; W&W, wait and watch.

random-effect model was adopted for analysis. Egger’s test showed no bias of publication ($p = 0.776$).

Three articles reported 3-year DFS for comparison between the W&W and pCR groups. Pooled

analysis demonstrated that the 3-year DFS in the pCR group was significantly higher than that in the W&W group (RR = 2.462, 95%CI = 1.131–5.358, $p < 0.05$) (Figure 4b). No heterogeneity was observed among the studies ($p = 0.749$, $I^2 = 0.0\%$), and a fixed-effect model was adopted

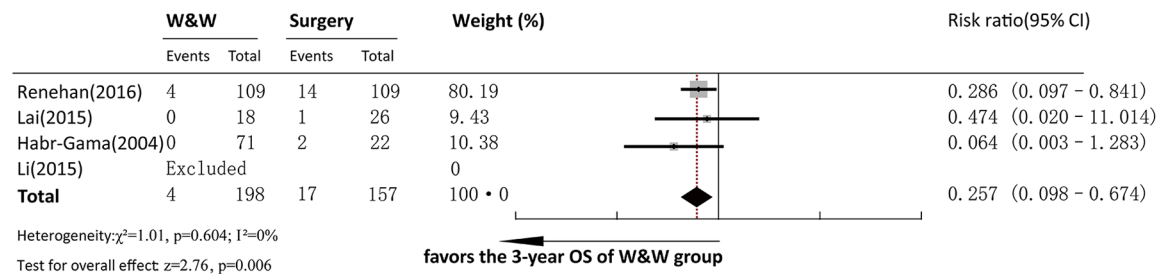
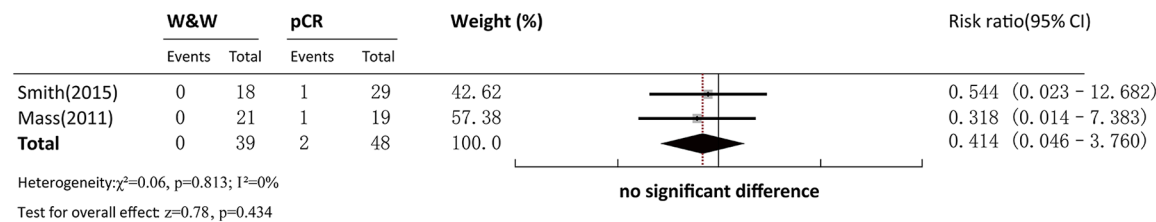
A 3-year OS for patients in W&W group versus Surgery group**B 3-year OS for patients in W&W group versus pCR group**

Figure 5. Forest plot of 3-year OS. (a) Comparison between the W&W and Surgery groups. (b) Comparison between the W&W and pCR groups. OS, overall survival; W&W, wait and watch.

for analysis. Egger's test showed no bias of publication ($p = 0.622$).

Three-year OS. Four articles reported 3-year OS for comparison between the W&W and Surgery groups. Pooled analysis demonstrated that the 3-year OS in the Surgery group was significantly lower than that in the W&W group (RR = 0.257, 95%CI = 0.098–0.674, $p < 0.05$) (Figure 5a). No heterogeneity was observed among the studies ($p = 0.604$, $I^2 = 0\%$), and a fixed-effect model was adopted for analysis. Egger's test for bias of publication was not conducted due to insufficient studies.

Two articles reported 3-year OS for comparison between the W&W and pCR groups. Pooled analysis demonstrated that the 3-year OS in the pCR group and that in the W&W group were not significantly different (RR = 0.427, 95%CI = 0.046–3.951, $p > 0.05$) (Figure 5b). No heterogeneity was observed among the studies ($p = 0.759$, $I^2 = 0.0\%$), and a fixed-effect model was adopted for analysis. Egger's test for bias of publication was not conducted due to insufficient studies.

Five-year DFS. Two articles reported 5-year DFS for comparison between the W&W and Surgery groups. Pooled analysis demonstrated that the

5-year DFS in the W&W group and that in the Surgery group were not significantly different (RR = 0.781, 95%CI = 0.136–4.467, $p > 0.05$) (Figure 6a). Significant heterogeneity was observed between the studies ($p = 0.089$, $I^2 = 65.5\%$), and a random-effect model was adopted for analysis. Egger's test for bias of publication was not conducted due to insufficient studies.

Two articles reported 5-year DFS for comparison between the W&W and pCR groups. Pooled analysis demonstrated that the 5-year DFS in the pCR group was significantly higher than that in the W&W group (RR = 2.076, 95%CI = 1.106–3.897, $p < 0.05$) (Figure 6b). No heterogeneity was observed between the studies ($p = 0.414$, $I^2 = 0.0\%$), and a fixed-effect model was adopted for analysis. Egger's test for bias of publication was not conducted due to insufficient studies.

Five-year OS. Three articles reported 5-year OS for comparison between the W&W and surgery groups. Pooled analysis demonstrated that the 5-year OS in the Surgery group was significantly lower than that in the W&W group (RR = 0.195, 95%CI = 0.039–0.974, $p < 0.05$) (Figure 7a). No heterogeneity was observed among the studies ($p = 0.696$, $I^2 = 0\%$), and a fixed-effect model was adopted for analysis.

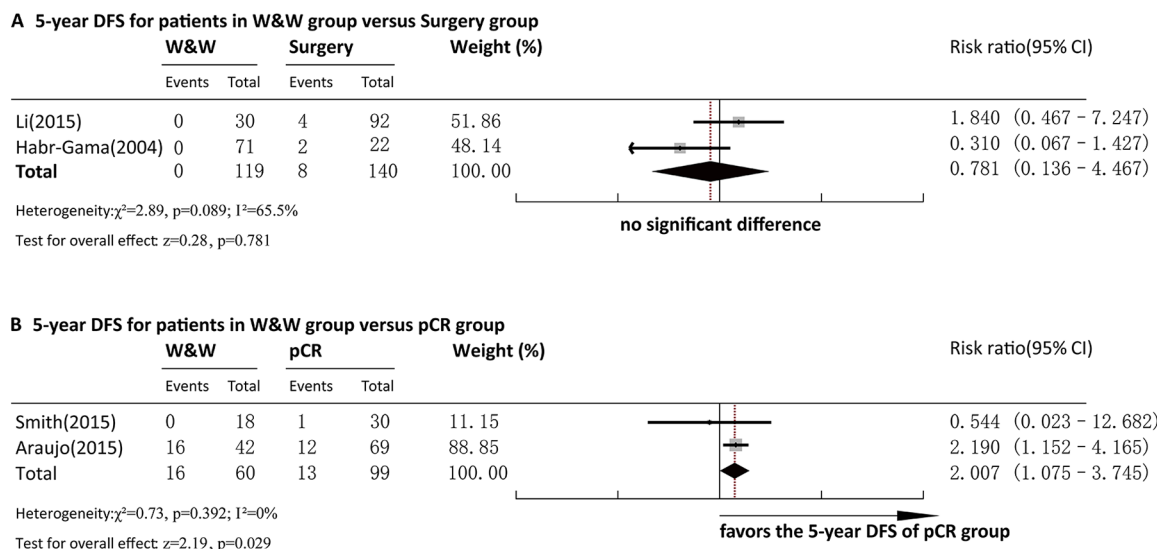


Figure 6. Forest plot of 5-year DFS. (a) Comparison between the W&W and Surgery groups. (b) Comparison between the W&W and pCR groups.
 DFS, disease-free survival; W&W, wait and watch.

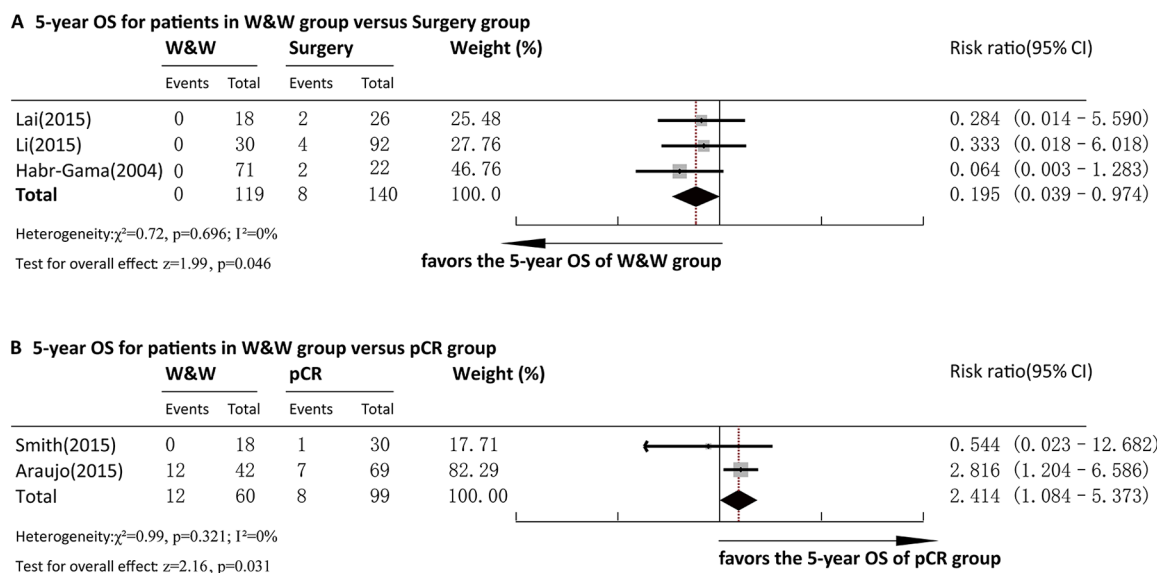


Figure 7. Forest plot of 5-year OS. (a) Comparison between the W&W and Surgery groups. (b) Comparison between the W&W and pCR groups.
 W&W, wait and watch; OS, overall survival.

Egger’s test for bias of publication was not conducted due to insufficient studies.

Two articles reported 5-year OS for comparison between the W&W and pCR groups. Pooled analysis demonstrated that the 5-year OS in the pCR group was significantly higher than that in

the W&W group (RR = 2.528, 95%CI = 1.113–5.741, $p < 0.05$) (Figure 7b). No heterogeneity was observed between the studies ($p = 0.339$, $I^2 = 0.0\%$), and a fixed-effect model was adopted for analysis. Egger’s test for bias of publication was not conducted due to insufficient studies.

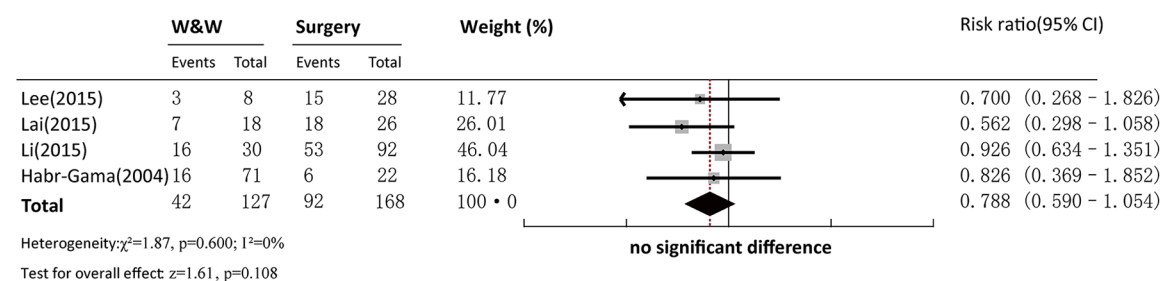
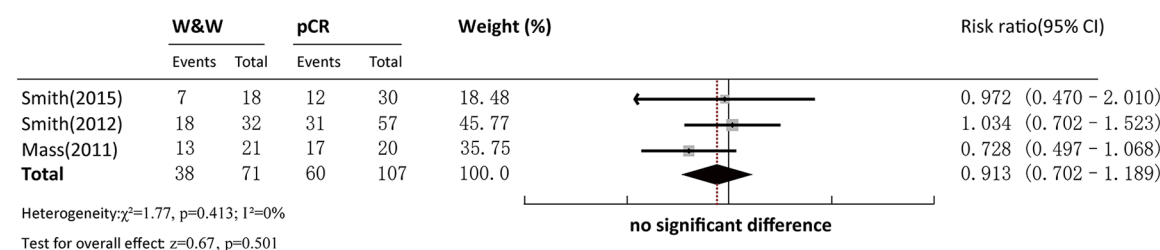
A Percentage of patients with stage III/IV disease in W&W group versus Surgery group**B Percentage of patients with stage III/IV disease in W&W group versus pCR group**

Figure 8. Forest plot of percentage of patients with stage III/IV disease. (a) Comparison between the W&W and Surgery groups. (b) Comparison between the W&W and pCR groups. OS, overall survival; W&W, wait and watch.

Percentage of patients with stage III/IV disease. Four articles reported distribution of patients' tumor stage for comparison between the W&W and Surgery groups. Pooled analysis demonstrated that the percentage of patients with stage III/IV disease was not significantly different between the W&W group and the Surgery group (RR = 0.788, 95%CI = 0.590–1.054, $p > 0.05$) (Figure 8a). No heterogeneity was observed among the studies ($p = 0.600$, $I^2 = 0\%$), and a fixed-effect model was adopted for analysis. Egger's test showed no bias of publication ($p = 0.343$).

Three articles reported the distribution of patients' tumor stage for comparison between the W&W and pCR groups. Pooled analysis demonstrated that the percentage of patients with stage III/IV disease was not significantly different between the W&W group and the pCR group (RR = 0.913, 95%CI = 0.702–1.189, $p > 0.05$) (Figure 8b). No heterogeneity was observed between the studies ($p = 0.413$, $I^2 = 0.0\%$), and a fixed-effect model was adopted for analysis. Egger's test showed no bias of publication ($p = 0.319$).

Discussion

How to maximally reduce surgical trauma has always been one of the top priorities for surgeons.

Despite the routine treatment of surgical resection for most rectal cancers of relatively earlier stage, Habr-Gama and colleagues were the first to carry out a nonsurgical strategy,⁶ which is now referred to as W&W. She believed that, for rectal cancer patients having achieved cCR after NCRT, some may avoid surgery altogether. And during the course of nonsurgical management, intensive follow-up protocols have to be implemented to make sure that LR/regrowth is duly dealt with.

The value of the W&W strategy has been recognized recently, but its use in clinical practice brings frustration as well as excitement. One major problem concerning the broader application of the W&W strategy is that, although it is considered as recommended for patients who have achieved cCR after NCRT, criteria differ between studies as to what the definition of cCR should be. That is to say, the outcome of cCR depends on the extent of precision of the selection process. Basically, the current methods for defining cCR include digital rectal examination, CT, MRI, EUS, proctoscopy, proctoscopic rebiopsy, and serum CEA level. However, not every center is capable of routinely implementing the strictest criteria, and concluding a diagnosis of cCR after confirming negative on all the diagnostic methods mentioned above.

References	Comparison	Digital rectal	Radiology			Proctoscopy	Proctoscopic	Serum CEA
		examination	CT	MRI	EUS		re-biopsy	level
Renehan et al.(2016)	W&W vs Surgery	√	√ (unspecified)			√	Unspecified	×
Lee et al.(2015)		×	×	√	×	×	×	×
Lai et al.(2015)		√	√ (unspecified)			√	Unspecified	×
Li et al.(2015)		√	√ (unspecified)			√	Unspecified	×
Habr-Gama et al.(2004)		√	√	×	×	√	√	√
Smith et al.(2015)	W&W vs pCR	Unspecified	Unspecified	Unspecified	Unspecified	Unspecified	Unspecified	Unspecified
Araujo et al.(2015)		√	×	√	×	√	Unspecified	×
Smith et al.(2012)		√	×	×	×	√	×	×
Mass et al.(2011)		√	×	√	×	√	√	×

Figure 9. Definitions of cCR in different studies.
cCR, clinical complete response; EUS, Endorectal ultrasound.

From Figure 9, we can see that, in all five studies comparing W&W with Surgery, the research of Habr-Gama and colleagues has the strictest cCR definition, and the 2-year LR, 3-year DFS, 3-year OS, 5-year DFS, and 5-year OS of her W&W group (Table 2) are all the most desirable when compared with their respective counterparts of the W&W groups in the other four studies. On the other hand, the research of Lee and colleagues has the laxest cCR definition (Figure 9), and the 3-year DFS (the only indicator in their study) of their W&W group (Table 2) is the worst when compared with its counterparts of the W&W groups in the other four studies. In the four studies comparing W&W with pCR, excluding the research of Smith and colleagues for its ambiguous cCR definition (Figure 9),¹³ the research of Mass and colleagues has the strictest cCR definition, and all the indicators (2-year LR, 2-year DM/DM+LR, 3-year DFS, 3-year OS as listed in Table 3) of their W&W group are the most desirable when compared with their respective counterparts in the W&W groups in the other two studies. On the other hand, the research of Smith and colleagues has the laxest cCR definition (Figure 9),¹¹ and the 2-year LR of their W&W group (Table 3) are the worst compared with their counterparts in the W&W groups in the other 2 studies, even though the other indicator (2-year DM/DM+LR) of their W&W group (Table 3) are only second worst.

Thus, as indicated above, different definitions of cCR have obvious and direct influence on the prognosis of the W&W patients, as well as on the relative prognostic advantages of W&W strategy over surgical treatment under equivalent conditions. And, implementing the strictest definition

of cCR can definitely maximize the prognostic benefits for W&W patients. In this case, we suggest the strictest definition of cCR being negative results on all the following examinations: digital rectal examination, radiology (CT or MRI or EUS), proctoscopy, proctoscopic rebiopsy, and serum CEA level. Considering the practical situations in most countries/regions, it is thus recommended that application of the W&W strategy be restricted to central hospitals with sufficient equipment and personnel, as well as a strict and standardized registry and follow-up system.

Despite rigorous selection, cCR does not necessarily correspond to pCR. In other words, for patients diagnosed with cCR after NCRT, the pathological result of a rebiopsy or surgical specimen does not always indicate a pCR. And for this group of patients, if radical surgery were not applied, recurrence is almost inevitable. Interestingly, in a retrospective point of view, the residue lesions of as many as 7% of patients validated as pCR by surgery would have been mistaken for cancerous ulcers before surgery.¹⁴ Furthermore, there is still another major obstacle concerning the W&W strategy, which is its inability to effectively address cancerous cells possibly remaining in lymph nodes, as well as in perirectal tumor deposits.

From the results of our pooled analysis, we can summarize that, regarding 2-year LR, 2-year DM/DM+LR, 3-year DFS, and 5-year DFS (Figures 2a, 3a, 4a, 6a), the prognosis of the W&W group is not inferior to that of the Surgery group, whereas, regarding 3-year (Figure 5a) and 5-year (Figure 7a) OS, the prognosis of the W&W group is even superior to that of the Surgery

group. However, as shown in Figure 8a, in every individual study, the percentages of patients with stage III/IV disease in the W&W groups are all lower than that in the Surgery groups. In other words, despite the fact that the pooled analysis in Figure 8a failed to reach significance, the TNM stages of patients included in the W&W group are predominantly earlier than those in the Surgery group (Figure 8a). This can, to a certain extent at least, explain W&W's seeming advantage on 3-year and 5-year OS, and, at the same time, undermine the credibility of W&W's noninferiority on other prognosis indicators, as mentioned above. On the other hand, the pCR group has significant advantage over the W&W group on 2-year LR, 3-year DFS, 5-year DFS, and 5-year OS (Figure 2b, 4b, 6b, 7b). On 2-year DM/DM+LR and 3-year OS (Figure 3b, 5b), the pCR group is not inferior to the W&W group.

Regarding this novel treatment strategy, we believe that its most significant advantage is not improving patients' long-term and short-term prognosis. Instead, it avoids the trauma and potentially critical complications of surgery, which is especially relevant for older patients. According to Smith and colleagues, patients aged over 80 can benefit significantly from a W&W strategy.¹⁹ Still, the most urgent questions faced with this management strategy are how to more accurately assess cCR, as well as how to optimize a more standardized assessment/follow-up protocol.

In conclusion, compared with patients undergoing surgery treatment, the W&W group had higher 3-year and 5-year OS and was not inferior on other major prognostic indicators, which, however, is based on the fact that the staging of tumors in W&W groups is generally earlier. And, compared with surgically treated patients who acquire pCR, the W&W group was inferior on all major prognostic indicators except 2-year DM/DM+LR and 3-year OS. Additionally, by comparison of cCR definitions across different studies, we conclude that implementation of the strictest cCR criteria is critical for W&W patients to acquire maximum prognostic benefit.

Acknowledgments

Our deepest gratitude goes to the Managing Editor, Alessandro Baliani, the Associate Editor, Muhammad Wasif Saif, the Peer Review Supervisor, Kanika Kambo, and the reviewers for

their careful work and thoughtful suggestions that have helped improve this paper substantially.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and publication of this article: This research was funded by Beijing Natural Science Foundation (No. 7172062).

Conflict of interest statement

The authors declare that there is no conflict of interest.


Ethics approval statement

Ethics approval was not required for this study.

Informed consent statement

Informed consent was not required for this study.

ORCID iD

Kai Pang  <https://orcid.org/0000-0002-5991-1728>

Data availability

Underlying research materials related to this paper (data, samples or models) can be accessed via e-mail to the corresponding author.

References

1. Fujita S, Akasu T, Mizusawa J, *et al.* Postoperative morbidity and mortality after mesorectal excision with and without lateral lymph node dissection for clinical stage or stage III lower rectal cancer (JCOG0212): results from a multicenter, randomized controlled, non-inferiority trial. *Lancet Oncol* 2012; 13: 616–621.
2. Harrison JD, Solomon MJ, Young JM, *et al.* Patient and physician preferences for surgical and adjuvant treatment options for rectal cancer. *Arch Surg* 2008; 143: 389–394.
3. Glynne-Jones R and Hughes R. Complete response after chemoradiotherapy in rectal cancer (watch-and-wait): have we cracked the code? *Clin Oncol* 2016; 28: 152–160.
4. Mass M, Nelemans PJ, Valentini V, *et al.* Long-term outcome in patients with a pathological complete response after chemoradiation for rectal cancer: a pooled analysis of individual patient data. *Lancet Oncol* 2010; 11: 835–844.
5. Nakagawa WT, Rossi BM, de O, Ferreira F, *et al.* Chemoradiation instead of surgery to treat mid

- and low rectal tumors: is it safe? *Ann Surg Oncol* 2002; 9: 568–573.
6. Habr-Gama A, Perez RO, Nadalin W, *et al.* Operative versus nonoperative treatment for stage 0 distal rectal cancer following chemoradiation therapy: long-term results. *Ann Surg* 2004; 240: 711–717.
 7. Appelt AL, Pløen J, Harling H, *et al.* High-dose chemoradiotherapy and watchful waiting for distal rectal cancer: a prospective observational study. *Lancet Oncol* 2015; 16: 919–927.
 8. Habr-Gama A, Gama-Rodrigues J, São Julião GP, *et al.* Local recurrence after complete clinical response and watch and wait in rectal cancer after neoadjuvant chemoradiation: impact of salvage therapy on local disease control. *Int J Radiat Oncol Biol Phys* 2014; 88: 822–828.
 9. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *Eur J Epidemiol* 2010; 25: 603–605.
 10. Maas M, Beets-Tan RG, Lambregts DM, *et al.* Wait-and-see policy for clinical complete responders after chemoradiation for rectal cancer. *J Clin Oncol* 2011; 29: 4633–4640.
 11. Smith JD, Ruby JA, Goodman KA, *et al.* Nonoperative management of rectal cancer with complete clinical response after neoadjuvant therapy. *Ann Surg* 2012; 256: 965–972.
 12. Ayloor Seshadri R, Kondaveeti SS, Jayanand SB, *et al.* Complete clinical response to neoadjuvant chemoradiation in rectal cancers: can surgery be avoided? *Hepatogastroenterology* 2013; 60:410–414.
 13. Smith RK, Fry RD, Mahmoud NN, *et al.* Surveillance after neoadjuvant therapy in advanced rectal cancer with complete clinical response can have comparable outcomes to total mesorectal excision. *Int J Colorectal Dis* 2015; 30: 769–774.
 14. Li J, Liu H, Yin J, *et al.* Wait-and-see or radical surgery for rectal cancer patients with a clinical complete response after neoadjuvant chemoradiotherapy: a cohort study. *Oncotarget* 2015; 6: 4235–4261.
 15. Lee SY, Kim CH, Kim YJ, *et al.* Oncologic outcomes according to treatment strategy in radiologic complete responders after neoadjuvant chemoradiation for rectal cancer. *Oncology* 2015; 89: 311–318.
 16. Lai CL, Lai MJ, Wu CC, *et al.* Rectal cancer with complete clinical response after neoadjuvant chemoradiotherapy, surgery, or “watch and wait”. *Int J Colorectal Dis* 2016; 31: 413–419.
 17. Araujo RO, Valadão M, Borges D, *et al.* Nonoperative management of rectal cancer after chemoradiation opposed to resection after complete clinical response, a comparative study. *Eur J Surg Oncol* 2015; 41: 1456–1463.
 18. Renehan AG, Malcomson L, Emsley R, *et al.* Watch-and-wait approach versus surgical resection after chemoradiotherapy for patients with rectal cancer (the OnCoRe project): a propensity-score matched cohort analysis. *Lancet Oncol* 2016; 17: 174–183.
 19. Smith FM, Rao C, Oliva Perez R, *et al.* Avoiding radical surgery improves early survival in elderly patients with rectal cancer, demonstrating complete clinical response after neoadjuvant therapy: results of a decision-analytic model. *Dis Colon Rectum* 2015; 58: 159–171.