Supplementary Online Content

Fokas E, Schlenska-Lange A, Polat B, et al; German Rectal Cancer Study Group. Chemoradiotherapy plus induction or consolidation chemotherapy as total neoadjuvant therapy for patients with locally advanced rectal cancer: long-term results of the CAO/ARO/AIO-12 randomized clinical trial. *JAMA Oncol*. Published online November 18, 2021. doi:10.1001/jamaoncol.2021.5445

This supplementary material has been provided by the authors to give readers additional information about their work.

eMethods

Randomization and masking

Patients were randomly assigned to one of the two groups (**eFigure 1**) using computer-generated randomization codes (sequential permuted blocks of length 4) stratified for center and clinical N category (cN0 *vs* cN1–2). Randomization was performed centrally and patient assignment was conducted using a web interface hosted by the WiSP Research Institute (Langenfeld, Germany), ensuring that the next assignment in the sequence was masked (allocation concealment). This was an open trial and, hence, patients and physicians were not masked to treatment allocation.

Pathology objectives

These have been described before in detail[1]. Central pathology review was not performed.

Statistical analysis

Sample size planning (reported previously)

The purpose of the present randomized phase 2 trial was to identify and select the more promising total neoadjuvant therapy (TNT) sequence for further phase 3 evaluation against standard preoperative CRT[1]. Our hypothesis was that TNT could increase the pCR rate up to 25%, whereas the pCR rate is approximately 15% after preoperative CRT[2, 3]. Each group was designed as a single-group study to discriminate between 15% pCR (standard preoperative CRT) and 25% pCR (TNT group). At the same time, the sample size was selected to ensure accurate selection of the more promising sequence for phase 3 evaluation. The sample size was calculated as follows: For a power of 80% at an exploratory significance level of 20% (two-sided), a sample size of 144 patients per group (1:1 ratio) was required assuming pCR of 25% and 15%. As pCR, the primary endpoint, was assessed at surgery, only a small drop-out of 5% was expected. Thus, a planned sample size of 152 patients per group (304 patients in total) was required to complete the study. Considering an individual group, the comparison of the estimated pCR rate hypothesized to be 25% with a fixed rate of 15% at a significance level of 10% (one-sided) has a power of 95% given the sample size of 144 patients. The power reduces to 87% when the usual two-sided significance level of 5% (or 2.5% one-sided) is used. The sample size gives a probability of 98% to pick the better treatment regimen for phase 3 evaluation assuming the pCR rates are 15% and 25%. Of note, the present study was not powered to demonstrate a significant difference between the TNT sequences at the usual

significance level of 5% two-sided. Instead it was designed to select the better TNT sequence as a “pick the winner” strategy, as reviewed before[4].

pCR analyses (reported previously)

The primary endpoint analysis for pCR was conducted according to the intention-to-treat principle, i.e. all randomized patients were included in the analyses and in the TNT groups they were randomized to. For the primary endpoint, the proportions of patients with pCR are reported with 95% confidence intervals (CI) and one-sided p-values for the comparison against the historical control of 15%. Toxicity, compliance, and further efficacy endpoints (TRG, NAR score, R0, sphincter preserving surgery) were also previously reported[1]. The treatment groups were compared in a logistic regression with pCR as dependent variable, and treatment and stratification parameters of the randomization (i.e. center and clinical N category) as factors. Centers including no more than 4 patients were combined. The odds ratio describing the treatment effect is reported with 95% CI and two-sided p-value testing the null hypothesis of no effect. Drop-out was dealt with as independent right censoring.

Definition of oncological endpoints, statistical tests and Wexner stool incontinence score (present analysis)

Regarding the definition of secondary endpoints, DFS was defined as the time between randomization and either macroscopically visible gross tumor after surgery (R2 resection), no resection due to tumor progression, locoregional recurrence after R0/1 resection of the primary tumor, distant metastases, or death from any cause, whichever occurred first; a local regrowth in patients with cCR and NOM was censored if salvage surgery resulted in a R0/1 resection, as recently recommended[5]. The cumulative incidence of locoregional recurrence and distant metastases was defined as the time between randomization and occurrence of any locoregional recurrence (after R0/1 resection of the primary tumor) and distant metastases, respectively, irrespective of whether this was a first event or not. Overall survival (OS) was defined as time from randomization to death from any cause.

All analyses were done in the intention-to-treat population. For all time-to-event outcomes, we performed Cox proportional hazards regression analyses including treatment group, center and N status as factors, and reported hazard ratios (HRs) for the treatment effect with 95% confidence intervals (CI) and two-sided p-values for the null hypothesis of no treatment difference (i.e. HR=1). The cumulative incidences of locoregional recurrence and distant metastases were assessed with death (for any cause) as competing event using the Aalen-

Johansen estimator. Exploratory subgroup analyses of DFS on associations of baseline characteristics and the treatment effect are presented in a forest plot; P-values for interactions between the baseline characteristic and the treatment effect are derived by including the baseline characteristic as well as the interaction of the baseline characteristic with treatment in the Cox regression models described above.

The Wexner stool incontinence score (summarized in **eTable 1)** is shown as box plot (Tukey definition). The distribution of scores is shown as box plots, with the lower part of the box representing the second quartile and the upper part of the box representing the third quartile. The line within the box represents the median, and the whiskers extend up to 1.5 times of the interquartile range (IQR) from the box, whereas dot points (outliers) represent data outside this range. The results for the other secondary endpoints (chronic toxicity, QoL/PROMs and Wexner score) reported here were merely descriptive.

eTable 1. Multivariable Cox Proportional Hazard Model for DFS Including TNT Group and Baseline Parameters With Potential Imbalances

Variable	Multivariable, complete model HR (95% CI)	P-value
TNT Group	1.01 (0.65 – 1.56)	0.98
ECOG performance status	1.24 (0.77 – 2.00)	0.38
Distance of tumor mesorectal fascia	1.22 (0.72 – 2.06)	0.46
Location from anal verge		
>5-10 cm	0.83 (0.52 – 1.32)	0.43
>10 cm	0.95 (0.42 – 2.15)	0.91

eTable 1. Multivariable Cox proportional hazard model for DFS including TNT group and baseline parameters with potential imbalances. In total 293 patients with 83 events were included; 13 observations were missing and were not included.

eTable 2. Wexner Incontinence Score

Type of incontinence	Frequency				
	Never	Rarely	Sometimes	Usually	Always
Solid	0	1	2	3	3
Liquid	0	1	2	3	3
Gas	0	1	2	3	3
Wears pad	0	1	2	3	3
Lifestyle alteration	0	1	2	3	3

eTable 2. Wexner incontinence score[6]. Never, 0; rarely, <1/month; sometimes, <1/week, 1/month; usually, <1/day, 1/week; always, 1/day. 0, perfect; 20, complete incontinence

eTable 3. Severity of Stool Incontinence Based on the Wexner Score

Wexner score of stool incontinence (range)	Severity of stool incontinence
0	Normal
1-8	Minor
9-14	Average
15-20	Complete

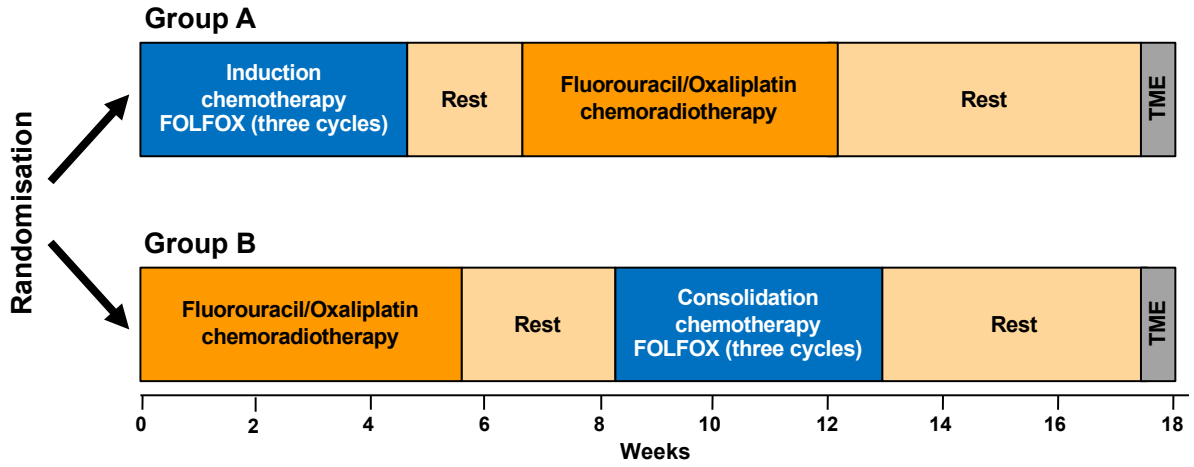
eTable 3. Severity of stool incontinence based on the Wexner score, as previously described[7]

eTable 4. Chronic Toxicity at 12 Months and 36 Months in Patients With Clinical Complete Response (cCR) That Refused Surgery

Chronic Toxicity (NCI-CTC v4.0)	Toxicity at 12 months				Toxicity at 36 months			
	TNT group A (n=6)		TNT group B (n=3)		TNT group A (n=4)		TNT group B (n=3)	
	Grade 1-2	Grade 3-4	Grade 1-2	Grade 3-4	Grade 1-2	Grade 3-4	Grade 1-2	Grade 3-4
Gastrointestinal								
Anastomotic stenosis	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Diarrhea	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (25%)	0 (0%)	2 (66.7%)	0 (0%)
Proctitis/rectal pain	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (33.3%)	0 (0%)
Nervous system								
Oxaliplatin-induced neurotoxicity: Wassermann score[8]	1 (16.7%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Peripheral sensory neuropathy	3 (50%)	0 (0%)	0 (0%)	0 (0%)	1 (25%)	0 (0%)	1 (33.3%)	0 (0%)
Genitourinary								
Cystitis noninfective	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Urinary fistula	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Urinary retention	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (25%)	0 (0%)	0 (0%)	0 (0%)
Erectile dysfunction	2 (40%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Vaginal dryness	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Vaginal fistula	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Other								
Radiation dermatitis	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)

eTable 4. Chronic toxicity at 12 months and 36 months in patients with clinical complete response (cCR) that refused surgery. Please, note that in group A toxicity data were missing in two of the six patients with cCR at 36 months (1 due to previous local regrowth that received salvage surgery; 1 not documented). In group B toxicity data were missing in one of the four patients at 12 and 36 months (not documented).

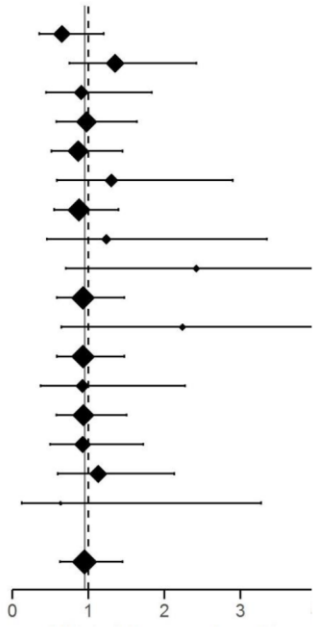
eFigure 1. Treatment Schedule



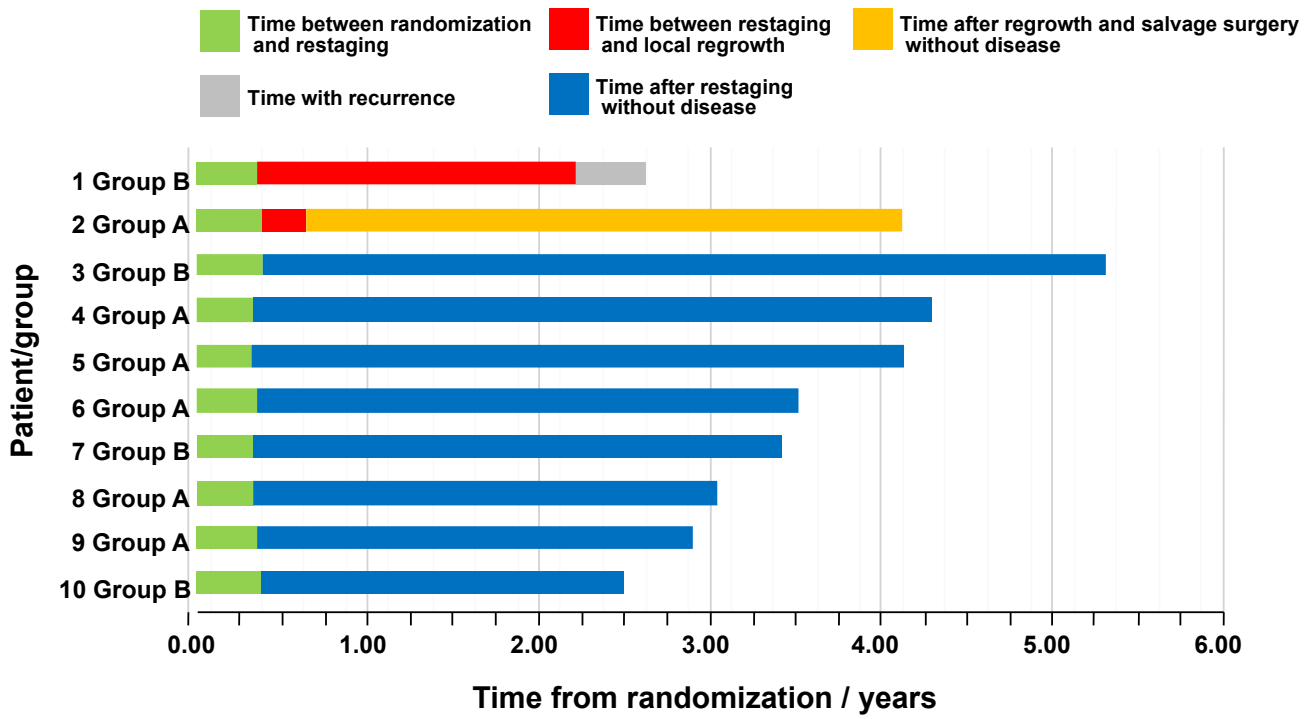
eFigure 1. Treatment schedule. Patients were randomized to receive total neoadjuvant treatment (TNT) with either induction chemotherapy prior to chemoradiotherapy (group A), or consolidation chemotherapy following chemoradiotherapy (group B). Concurrent chemotherapy was similar in both groups and was administered using continuous infusion of fluorouracil (250 mg/m^2) on days 1–14 and 22–35 of radiotherapy, and a 2-h infusion of oxaliplatin (50 mg/m^2) on days 1, 8, 22, and 29 of radiotherapy. Induction-/consolidation chemotherapy consisted of oxaliplatin (100 mg/m^2) administered as a 2-h infusion, followed by a 2-h infusion of leucovorin (400 mg/m^2), followed by a continuous 46-h infusion of fluorouracil (2400 mg/m^2) (FOLFOX), repeated on day 15 for a total of 3 cycles. Total mesorectal excision (TME) surgery was scheduled on day 123 (approximately week 18) after start of TNT in both groups.

eFigure 2. Forest Plot of the Effect of Treatment on Disease-Free Survival (DFS) According to Pretreatment Characteristics

Baseline characteristics		Group A Events/ patients	Group B Events/ patients	HR (95% CI)	$p_{interaction}$
Age	<61 years	26/73	17/72	0.65 (0.35-1.2)	.088
	≥61 years	20/83	26/78	1.35 (0.75-2.42)	
Sex	female	16/50	15/50	0.91 (0.45-1.83)	.864
	male	30/106	28/100	0.98 (0.58-1.64)	
ECOG	ECOG 0	34/118	25/100	0.87 (0.52-1.45)	.397
	ECOG 1	9/32	18/48	1.3 (0.58-2.9)	
Clinical T category	cT2-cT3	40/138	32/122	0.88 (0.55-1.4)	.534
	cT4	6/18	11/27	1.24 (0.46-3.35)	
Clinical N category	cN0	4/16	8/14	2.42 (0.7-8.33)	.145
	cN1-2	37/134	35/135	0.93 (0.59-1.48)	
Clinical disease stage	Stage II	4/16	8/14	2.24 (0.65-7.74)	.183
	Stage III	37/134	35/135	0.93 (0.59-1.48)	
Distance of tumor to mesorectal fascia	≤ 1mm	12/48	8/33	0.93 (0.38-2.28)	.983
	> 1mm	34/108	35/117	0.94 (0.58-1.5)	
Location from anal verge	0-5 cm	21/64	19/62	0.93 (0.5-1.72)	.774
	>5 - 10 cm	17/67	22/73	1.13 (0.6-2.13)	
	>10 cm	5/15	2/11	0.63 (0.12-3.27)	
Overall estimate	All	46/156	43/150	0.95 (0.63-1.45)	

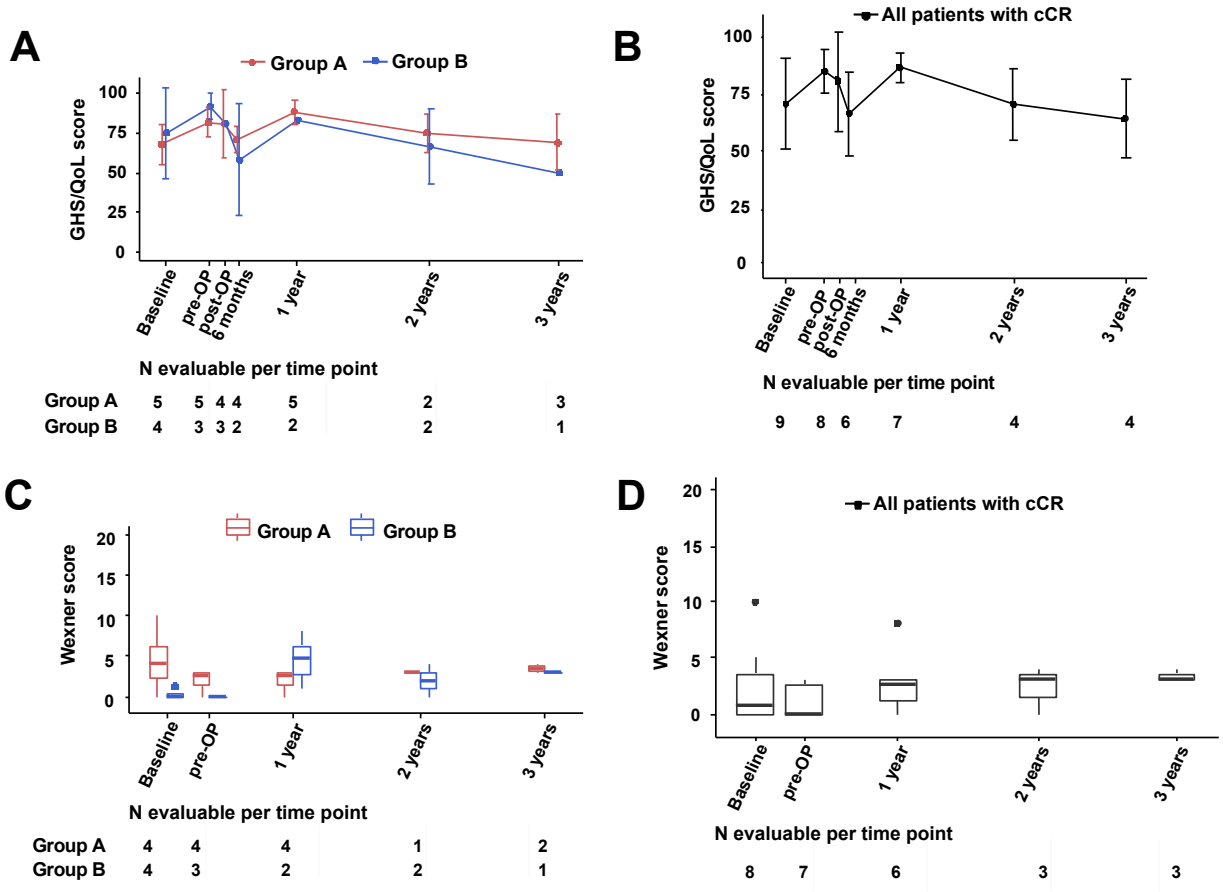


eFigure 2. Forest plot of the effect of treatment on disease-free survival (DFS) according to pretreatment characteristics. The size of the quadrats represents the proportion of patients. HR, hazard ratio.



eFigure 3. Clinical Outcome in the 10 Patients With Clinical Complete Response (cCR) That Denied Surgery and Were Managed With a "Watch and Wait" Strategy

eFigure 4. Quality of Life and Incontinence Changes Over Time in the 10 Patients With Clinical Complete Response (cCR) That Denied Surgery and Were Managed With a "Watch and Wait" Strategy



eFigure 4. Quality of life and incontinence changes over time in the 10 patients with clinical complete response (cCR) that denied surgery and were managed with a "Watch and Wait" strategy. (A) Global health status (GHS)/QoL score, assessed by the EORTC QLQ-C30 questionnaire in the two groups, as indicated. Only disease-free patients were included in the analysis, whereas patients that refused surgery due to cCR were excluded. The x-axis indicates the time from randomization. The scores are linearly converted to a 100-point scale. Data are shown as mean score ± SD. A higher mean score represents a better level of GHS/QoL score. **(B)** GSH/QoL score in all patients with cCR. Note that the numbers decrease over time due to tumor regrowth/recurrence (one patient in each group) and lack of documentation. Stool

incontinence assessed by the Wexner Score in **(C)** the two groups and **(D)** in all patients combined together. The distribution of scores is shown as box plots, with the lower part of the box representing the second quartile and the upper part of the box representing the third quartile. The line within the box represents the median, and the whiskers extend up to 1.5 times of the interquartile range from the box, whereas dot points (outliers) represent data outside this range. A higher score represents worse incontinence status. The x-axis indicates the time from randomization. Data were missing in some patients due to lack of documentation.

eReferences

1. Fokas E, Allgauer M, Polat B et al. Randomized Phase II Trial of Chemoradiotherapy Plus Induction or Consolidation Chemotherapy as Total Neoadjuvant Therapy for Locally Advanced Rectal Cancer: CAO/ARO/AIO-12. J Clin Oncol 2019; 37: 3212-3222.
2. Rodel C, Graeven U, Fietkau R et al. Oxaliplatin added to fluorouracil-based preoperative chemoradiotherapy and postoperative chemotherapy of locally advanced rectal cancer (the German CAO/ARO/AIO-04 study): final results of the multicentre, open-label, randomised, phase 3 trial. Lancet Oncol 2015; 16: 979-989.
3. Rodel C, Liersch T, Becker H et al. Preoperative chemoradiotherapy and postoperative chemotherapy with fluorouracil and oxaliplatin versus fluorouracil alone in locally advanced rectal cancer: initial results of the German CAO/ARO/AIO-04 randomised phase 3 trial. Lancet Oncol 2012; 13: 679-687.
4. Chow SC, Chang M. Adaptive design methods in clinical trials - a review. Orphanet J Rare Dis 2008; 3: 11.
5. Fokas E, Glynne-Jones R, Appelt A et al. Outcome measures in multimodal rectal cancer trials. Lancet Oncol 2020; 21: e252-e264.
6. Jorge JM, Wexner SD. Etiology and management of fecal incontinence. Dis Colon Rectum 1993; 36: 77-97.
7. Dulskas A, Kavaliauskas P, Pilipavicius L et al. Long-term bowel dysfunction following low anterior resection. Sci Rep 2020; 10: 11882.
8. Wasserman E, Cuvier C, Lokiec F et al. Combination of oxaliplatin plus irinotecan in patients with gastrointestinal tumors: results of two independent phase I studies with pharmacokinetics. J Clin Oncol 1999; 17: 1751-1759.