Detection of residual disease after neoadjuvant chemoradiotherapy for oesophageal cancer (preSANO): a prospective multicentre, diagnostic cohort study

*Bo Jan Noordman, Manon C W Spaander, Roelf Valkema, Bas P L Wijnhoven, Mark I van Berge Henegouwen, Joël Shapiro, Katharina Biermann, Ate van der Gaast, Richard van Hillegersberg, Maarten C C M Hulshof, Kausilia K Krishnadath, Sjoerd M Lagarde, Grard A P Nieuwenhuijzen, Liekele E Oostenbrug, Peter D Siersema, Erik J Schoon, Meindert N Sosef, Ewout W Steyerberg, J Jan B van Lanschot, for the SANO study group**

Summary

Background After neoadjuvant chemoradiotherapy for oesophageal cancer, roughly half of the patients with squamous cell carcinoma and a quarter of those with adenocarcinoma have a pathological complete response of the primary tumour before surgery. Thus, the necessity of standard oesophagectomy after neoadjuvant chemoradiotherapy should be reconsidered for patients who respond sufficiently to neoadjuvant treatment. In this study, we aimed to establish the accuracy of detection of residual disease after neoadjuvant chemoradiotherapy with different diagnostic approaches, and the optimal combination of diagnostic techniques for clinical response evaluations.

Methods The preSANO trial was a prospective, multicentre, diagnostic cohort study at six centres in the Netherlands. Eligible patients were aged 18 years or older, had histologically proven, resectable, squamous cell carcinoma or adenocarcinoma of the oesophagus or oesophagogastric junction, and were eligible for potential curative therapy with neoadjuvant chemoradiotherapy (five weekly cycles of carboplatin [area under the curve 2 mg/mL per min] plus paclitaxel [50 mg/m² of body-surface area] combined with 41·4 Gy radiotherapy in 23 fractions) followed by oesophagectomy. 4–6 weeks after completion of neoadjuvant chemoradiotherapy, patients had oesophagogastroduodenoscopy with biopsies and endoscopic ultrasonography with measurement of maximum tumour thickness. Patients with histologically proven locoregional residual disease or no-pass during endoscopy and without distant metastases underwent immediate surgical resection. In the remaining patients a second clinical response evaluation was done (PET–CT, oesophagogastroduodenoscopy with biopsies, endoscopic ultrasonography with measurement of maximum tumour thickness, and fine-needle aspiration of suspicious lymph nodes), followed by surgery 12–14 weeks after completion of neoadjuvant chemoradiotherapy. The primary endpoint was the correlation between clinical response during clinical response evaluations and the final pathological response in resection specimens, as shown by the proportion of tumour regression grade (TRG) 3 or 4 (>10% residual carcinoma in the resection specimen) residual tumours that was missed during clinical response evaluations. This study was registered with the Netherlands Trial Register (NTR4834), and has been completed.

Findings Between July 22, 2013, and Dec 28, 2016, 219 patients were included, 207 of whom were included in the analyses. Eight of 26 TRG3 or TRG4 tumours (31% [95% CI 17–50]) were missed by endoscopy with regular biopsies and fine-needle aspiration. Four of 41 TRG3 or TRG4 tumours (10% [95% CI 4–23]) were missed with bite-on-bite biopsies and fine-needle aspiration. Endoscopic ultrasonography with maximum tumour thickness measurement missed TRG3 or TRG4 residual tumours in 11 of 39 patients (28% [95% CI 17–44]). PET–CT missed six of 41 TRG3 or TRG4 tumours (15% [95% CI 7–28]). PET–CT detected interval distant histologically proven metastases in 18 (9%) of 190 patients (one squamous cell carcinoma, 17 adenocarcinomas).

Interpretation After neoadjuvant chemoradiotherapy for oesophageal cancer, clinical response evaluation with endoscopic ultrasonography, bite-on-bite biopsies, and fine-needle aspiration of suspicious lymph nodes was adequate for detection of locoregional residual disease, with PET–CT for detection of interval metastases. Active surveillance with this combination of diagnostic modalities is now being assessed in a phase 3 randomised controlled trial (SANO trial; Netherlands Trial Register NTR6803).

Funding Dutch Cancer Society.

Introduction

Oesophageal cancer is an aggressive malignancy: the proportion of patients who achieve 5-year survival after primary oesophagectomy rarely exceeds 35%.[1] Overall survival has improved substantially in the past two decades, however, mainly as a result of the widespread use of neoadjuvant chemo-radiotherapy.[2] Five weekly cycles of carboplatin (area under the curve 2 mg/mL per min) plus

Lancet Oncol 2018

Published **Online** May 31, 2018 http://dx.doi.org/10.1016/S1470-2045(18)30201-8

See **Online/Comment** http://dx.doi.org/10.1016/S1470-2045(18)30291-2

*Members listed in the appendix

Department of Surgery (B J Noordman MD, B P L Wijnhoven MD, J Shapiro MD, S M Lagarde MD, Prof J J B van Lanschot MD), **Department of Gastroenterology** (M C W Spaander MD, Prof P D Siersema MD), **Department of Radiology and Nuclear Medicine** (R Valkema MD), **Department of Pathology** (K Biermann MD), **Department of Medical Oncology** (A van der Gaast MD), **and Department of Public Health** (Prof E W Steyerberg MD), **Erasmus MC, University Medical Centre Rotterdam, Rotterdam, Netherlands; Department of Surgery** (Prof M I van Berge Henegouwen MD), **Department of Radiation Oncology** (M C C M Hulshof MD), **and Department of Gastroenterology** (Prof K K Krishnadath MD), **Academic Medical Centre, Amsterdam, Netherlands; Department of Surgery, University Medical Centre Utrecht, Utrecht, Netherlands** (Prof R van Hillegersberg MD); **Department of Surgery** (G A P Nieuwenhuijzen MD) **and Department of Gastroenterology** (E J Schoon MD), **Catharina Hospital, Eindhoven, Netherlands; Department of Gastroenterology** (L E Oostenbrug MD) **and Department of Surgery** (M N Sosef MD), **Zuyderland Medical Centre, Heerlen, Netherlands;**

Department of Gastroenterology, Radboud University Medical Centre, Nijmegen, Netherlands (Prof P D Siersema); and Department of Medical Statistics and Bioinformatics, Leiden University Medical Centre, Leiden, Netherlands (Prof E W Steyerberg)

Correspondence to: Dr Bo Jan Noordman, Deptarment of Surgery, Erasmus MC, University Medical Centre, 3000 CA Rotterdam, Netherlands
b.noordman@erasmusmc.nl

See Online for appendix

Research in context

Evidence before this study

We did not do a formal search of published work before this trial. The randomised CROSS trial established chemoradiotherapy (weekly administration of carboplatin and paclitaxel plus 41·4 Gy concurrent radiotherapy) followed by surgery as the standard of care for patients with oesophageal cancer, compared with surgery alone. However, after neoadjuvant chemoradiotherapy plus surgery, 29% of treated patients achieve a pathological complete response (as measured by histological examination of resection specimens), which provides a rationale for an active surveillance approach after neoadjuvant chemoradiotherapy with oesophagectomy offered only to patients with proven locoregional recurrence and without evidence of distant metastases. In a systematic review of four small retrospective studies, promising overall survival outcomes were associated with active surveillance after neoadjuvant chemoradiotherapy in patients with oesophageal cancer who had a clinically complete response. Additionally, previous small retrospective studies of single diagnostic modalities (endoscopy plus biopsy, endoscopic ultrasonography, or ^{18}F-fluorodeoxyglucose PET–CT) for detection of residual disease after neoadjuvant chemoradiotherapy have shown poor diagnostic accuracy. So far, the optimal combination of diagnostic tests for detection of residual disease in patients with oesophageal or gastroesophageal junction cancer after neoadjuvant chemoradiotherapy is unknown.

Added value of this study

By contrast with previous studies, in this multicentre, prospective cohort study, all available diagnostic modalities used for pre-treatment staging in clinical practice were applied to detect residual disease during active surveillance. These findings establish the optimal set of diagnostic modalities to accurately detect residual disease after neoadjuvant therapy in patients with oesophageal cancer, allowing the stratification of patients who would benefit from active surveillance versus radical oesophagostomy.

Implications of all the available evidence

Clinical response evaluations after neoadjuvant chemoradiotherapy for oesophageal cancer should consist of endoscopic ultrasonography with bite-on-bite biopsies and fine-needle aspiration of suspicious lymph nodes for detection of locoregional residual disease and PET–CT for detection of interval metastases. The promising diagnostic results of this study provide the rationale for a phase 3, randomised, controlled trial of active surveillance versus standard surgery in patients with oesophageal cancer treated with neoadjuvant chemoradiotherapy. Our results can be used to define the composition of the clinical response evaluations and subsequent surveillance examinations in future trials that could establish a new management protocol for patients with oesophageal cancer.

paclitaxel (50 mg/m² of body-surface area) plus 41·4 Gy radiotherapy in 23 fractions followed by oesophagectomy significantly improved overall survival at 5 years compared with oesophagectomy alone (47% [95% CI 39–54] in the neoadjuvant group *vs* 33% [26–40] in the surgery only group; hazard ratio 0·68 [95% CI 0·53–0·88]; log-rank p=0·003).[3,4] In 47 (29%) of 161 patients with oesophageal carcinoma (18 [49%] of 37 with squamous cell carcinoma and 28 [23%] of 121 with adenocarcinoma), a pathological complete response was noted after neoadjuvant chemoradiotherapy—ie, no viable tumour cells were detected in the resected specimen during conventional histological examination.[3]

This high frequency of pathological complete response provides a rationale to reconsider the necessity of standard oesophagectomy after neoadjuvant chemoradiotherapy. Theoretically, active surveillance could be feasible in patients without locoregional or disseminated disease, given that oesophagectomy probably does not affect oncological outcomes in patients with no viable tumour cells. In a pan-active surveillance approach, patients would undergo regular clinical investigations after neoadjuvant chemoradiotherapy, and oesophagectomy would be offered only to those with proven locoregional recurrence and no evidence of distant metastases.[3,5–7] However, an active surveillance approach would only be justified if the associated oncological outcomes were non-inferior to those achieved with standard surgery. To select patients for active surveillance, disease should be restaged after neoadjuvant chemoradiotherapy by means of clinical response evaluations, which need to accurately classify patients as complete or incomplete responders. We aimed to establish which combination of diagnostic tests for clinical response evaluation most accurately detects residual disease after neoadjuvant chemoradiotherapy in patients with oesophageal cancer.

Methods

Study design and participants

We did a prospective, multicentre, diagnostic cohort study at six centres in the Netherlands (appendix p 6); the study protocol has been previously published.[8] After adjuvant chemoradiotherapy, patients underwent a first clinical response evaluation. Patients found to be complete responders during the first clinical response evaluation (ie, those with no locoregional or disseminated disease proven by cytohistology) were offered postponed surgical resection, and in the week preceding surgery, a second clinical response evaluation was done before patients without distant metastases underwent oesophagectomy. If the planned operation was postponed for more than 4 weeks after the second clinical response evaluation (eg, because the patient had not sufficiently recovered from neoadjuvant chemoradiotherapy), a third clinical response evaluation was recommended a week before surgery.

Eligible patients were aged 18 years or older, had histologically proven, resectable, squamous cell carcinoma or adenocarcinoma of the oesophagus or oesophagogastric junction, and were eligible for potential curative therapy with neoadjuvant chemoradiotherapy followed by oesophagectomy.[3] Patients with a severe stricture (no-pass) on initial endoscopic ultrasonographic staging at baseline (ie, pre-treatment) were also included. The study protocol was approved by the medical ethics committee of the Erasmus MC (Rotterdam, MEC-2013-211). All patients provided written informed consent.

Procedures

All patients underwent primary clinical staging at baseline, including oesophagogastroduodenoscopy with biopsies, endoscopic ultrasonography with measurement of maximum tumour thickness,[9] CT of the neck, chest, and upper abdomen, and ^{18}F-fluorodeoxyglucose (^{18}F-FDG) PET–CT. Most patients were consciously sedated with midazolam during endoscopic ultrasonography; general anaesthesia was not routinely used.

The neoadjuvant chemoradiotherapy regimen consisted of five weekly cycles of carboplatin (area under the curve 2 mg/mL per min) plus paclitaxel (50 mg/m² of body-surface area) combined with 41·4 Gy radiotherapy in 23 fractions, as per the CROSS trial recommendations.[3] 4–6 weeks after completion of the last cycle of neoadjuvant chemotherapy, patients underwent a first clinical response evaluation before surgery to identify non-responders. During this clinical response evaluation, all patients underwent oesophagogastroduodenoscopy with biopsies and radial endoscopic ultrasonography with measurement of maximum tumour thickness and area. Patients with a severe stricture at endoscopy (no-pass) or histological evidence of locoregional residual disease underwent PET–CT to exclude distant metastases. If no distant metastases were detected, eligible patients underwent surgery within 2 weeks of the PET–CT assessment. Patients without histological evidence of residual disease during the first clinical response evaluation were offered postponed surgery scheduled approximately 6–8 weeks after the first clinical response evaluation and approximately 12–14 weeks after neoadjuvant chemoradiotherapy (compared with 6–8 weeks after neoadjuvant chemoradiotherapy for those with histological evidence of residual disease). In the week before surgery, we did a second clinical response evaluation to detect any residual disease that had developed or was previously undetected. The second clinical response evaluations comprised PET–CT, followed by oesophagogastroduodenoscopy with biopsies, radial endoscopic ultrasonography for measurement of maximum tumour thickness and area, and linear endoscopic ultrasonography plus fine-needle aspiration of any suspicious lymph nodes or ^{18}F-FDG-avid lesions. After the second clinical response evaluation, all patients without distant metastases underwent oesophagectomy. The third clinical response evaluation (in patients who needed it) was similar to the second.

When we designed the trial, endoscopy with random, conventional mucosal biopsies of the primary tumor site and of any other suspected lesions in the oesophagus were prespecified by protocol as part of the clinical response evaluations,[8] because the safety of deep bite-on-bite biopsies after neoadjuvant chemoradiotherapy was unknown. An interim safety analysis was pre-planned after inclusion of 60 patients to monitor serious complications and assess the radicality of the performed operations. On April 20, 2015, after about 95 patients were enrolled and had a regular biopsy, the interim safety analysis showed no biopsy-related adverse events, and the protocol was amended on June 22, 2015, to change the biopsy strategy. Thereafter, bite-on-bite biopsies were done instead of conventional biopsies during clinical response evaluations.[10]

During bite-on-bite biopsies, a second, deep, biopsy sample is taken at the same location as the first to increase the chance of detecting residual disease—especially submucosal—tumours (appendix p 7). Biopsies were taken from at least four different locations from the primary tumour site and from any suspicious lesions in the oesophagus. The regular biopsy procedure consisted of one biopsy of at least four different locations on the site of the primary tumour and from any suspicious lesion. All endoscopy reports and endoscopic ultrasonography images were reviewed by an experienced upper-gastrointestinal gastroenterologist (MCWS), who was blinded to pathological response results in the resected specimen after surgery.

During the second clinical response evaluation, fine-needle aspiration was done on any suspicious lymph nodes (round, hypoechogenic, and greater than 5 mm in diameter), or any lymph nodes adjacent to the primary tumour. Potential contamination from the primary tumour during fine-needle aspiration of adjacent lymph nodes was not an issue because the source of residual disease was not a variable considered in the trial outcome analyses. Maximum tumour thickness was measured as reported previously;[9] maximum tumour thickness of 6 mm or greater during the second clinical response evaluation was classified as non-complete response.[9]

Biopsies done at the first clinical response evaluation with uncertain outcome were deemed negative to reduce the risk of false-positive biopsies, whereas those with uncertain outcomes done at the second clinical response evaluation were judged positive to reduce the risk of false-negative biopsies, in light of a future active-surveillance strategy in patients with a clinical complete response after second clinical response evaluation.[11,12] Fine-needle aspirates taken from suspicious lymph nodes with uncertain outcomes or that were not representative (ie, no lymphoid tissue present) were considered positive for residual disease.

PET–CT was done according to the European Association of Nuclear Medicine guidelines.[13] All scans were reviewed by an experienced PET–CT radiologist (RV), who was blinded to pathological response results. ^{18}F-FDG PET–CT scans were visually assessed, including intensity of uptake and ^{18}F-FDG uptake in the environment (eg, adjacent oesophagus). A qualitative judgment was made, and results were scored from 1 to 5: 1 (benign), 2 (probably benign), 3 (equivocal), 4 (probably malignant), and 5 (malignant). To ensure we did not exclude any tumour residue, we defined all scores of 2 or higher as ^{18}F-FDG positive. For locoregional response assessment, PET–CT scans from the first clinical response evaluation were analysed, whereas for distant dissemination scans from both clinical response evaluations, if available, were used.

Transthoracic or transhiatal oesophagectomy was done depending on patient characteristics and local preference. A wide excision, including removal of regional lymph nodes and standard dissection of the lymph nodes around the coeliac axis, was done in all patients with the aim of removing at least 15 lymph nodes.

Resected tumours reviewed by an experienced upper-gastrointestinal pathologist (KB [an author] and MD [a collaborator]) following a standard protocol, and classified and graded according to the Union for International Cancer Control TNM Cancer Staging (7th edn).[14] We used the Chirieac modified tumour regression grade (TRG) system,[15] the most commonly used system in the Netherlands, to classify pathological response in the resected specimens as no residual carcinoma (TRG1), 1–10% residual carcinoma (TRG2), 11–50% residual carcinoma (TRG3), and greater than 50% residual carcinoma (TRG4).[15] All negative biopsies in patients with TRG3 or TRG4 tumours were re-reviewed by KB and MD.

Serious adverse events, which resulted in death, were life threatening, required hospital admission or prolongation of hospital stay, resulted in persistent or significant disability or incapacity, or were considered serious by the treating physician, were monitored continuously from the first clinical response evaluation until the day that the patient underwent surgery.

Participants could leave the study at any time for any reason if they wished to do so, and investigators could withdraw participants from the study for urgent medical reasons.

Outcomes

The primary outcome was to establish the accuracy of residual disease detection after neoadjuvant chemoradiotherapy, as reflected by the proportion of tumours classified as TRG3 or TRG4 that was missed during clinical response evaluations. The secondary outcome was the proportion of patients who had an R0 resection, defined as a resection with no gross or microscopic tumour cells present. Other prespecified outcomes were correlations between individual diagnostic modality (endoscopic examinations, PET–CT, and analysis of cytohistological biopsies) and pathological findings in the resection specimen, and optimal cutoffs with maximal distinction between patients with and without clinically relevant residual disease. Results for R0 resection will be published elsewhere.

Statistical analysis

We hypothesised that TRG3 and TRG4 tumours could be detected reliably with the described clinical response evaluations. The estimated maximum percentage of clinically false-negative TRG3 and TRG4 tumours accounted for was 10%.[8] Initially, we aimed to enrol 120 patients, approximately 45 (38%) of whom were estimated to have TRG3 or TRG4 residual tumours after surgery as per the CROSS trial results.[3] However, because of the change in biopsy strategy as per protocol amendment (June 22, 2015) after about 95 patients were enrolled and had a regular biopsy, the total sample size was increased to 215, to ensure that at least 120 patients would undergone bite-on-bite biopsies during the clinical response evaluations.

Outcomes were analysed separately for both biopsy strategies (regular biopsies *vs* bite-on-bite biopsies). For endoscopic biopsies, results from both clinical response evaluations were combined (if either was positive, the patient was classified as having residual disease). Outcomes of endoscopic ultrasonography with measurement of maximum tumour thickness and PET–CT were analysed in the overall patient population, because these modalities were not amended during the trial. Patients who did not have neoadjuvant chemoradiotherapy or who withdrew consent, and those with missing index tests because of protocol violation or death were excluded from all analyses. Patients with missing reference standard (ie, TRG) were excluded from the primary analysis. Perioperatively irresectable tumours (T4b) confirmed with frozen section analysis were classified as TRG4. 95% CIs were calculated according to the Wilson procedure, without a correction for continuity. Results of PET–CT and endoscopy with biopsies and fine-needle aspiration, and maximum tumour thickness measurement were correlated to TRG with the χ^2 test. An interim safety analysis (the results of which will be published elsewhere) was done to assess the radicality of the performed operations after a total inclusion of 60 patients. The pre-planned stopping rule established that if the proportion of patients with a radical resection was 70% or less in the first 60 patients, the trial would be stopped.

We calculated sensitivity, specificity, positive predictive value, and negative predictive value for TRG2, TRG3, and TRG4 combined versus TRG1. Patients with TRG2 tumours were not excluded from sensitivity, specificity, negative predictive value and positive predictive value analyses because this would bias results. As a secondary sensitivity analysis, we used multiple imputation of TRG

for patients who had active surveillance (instead of surgery) after clinical response evaluations to calculate the proportion of TRG3 and TRG4 residual tumours that was missed during clinical response evaluations, and sensitivity, specificity, positive predictive values, and negative predictive values for combined TRG2–4 versus TRG1.[16] We used a significance level of 0·05, based on two-sided tests. All analyses were done in SPSS (version 21.0). This study is registered with the Netherlands Trial Register (NTR4834).

Role of the funding source

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

Results

Between July 22, 2013, and Dec 28, 2016, 219 patients were enrolled (appendix p 6); 12 (6%) were excluded from further analyses: eight patients withdrew consent and four did not received the complete neoadjuvant chemoradiotherapy regimen (one had neoadjuvant chemotherapy only, two had definitive chemoradiotherapy, and one received palliative chemotherapy). Of 207 patients who underwent clinical response evaluations, 84 (41%) had clinical response evaluations with upper endoscopy and regular biopsies, of whom 61 (73%) were included in the analyses, and 123 (59%) had bite-on-bite biopsies, of whom 115 (93%) were included in the corresponding analyses (figure; appendix pp 8–9). Of the 207 patients who underwent clinical response evaluations, 113 (55%) were included in the endoscopic ultrasonographic examination of maximum tumour thickness, and 129 (62%) were included in the PET–CT analysis (figure; appendix pp 10–11). Baseline characteristics of all patients who underwent clinical response evaluations are shown in table 1.

Outcomes of regular biopsies and fine-needle aspiration during clinical response evaluations were significantly associated with the TRG of resected specimens (p=0·0036; table 2). Eight of 26 patients who had a regular biopsies and fine-needle aspiration, with a passable endoscopy, had negative biopsies despite having a TRG3 or TRG4 tumour (proportion of clinically false-negative cases 31% [95% CI 17–50]). 31 (51%) of 61 patients had positive biopsies, positive fine-needle aspiration, or no-pass. Sensitivity, specificity, negative predictive value, and positive predictive value of TRG2–4 versus TRG1 were 54% (95% CI 38–68; 20 of 37), 69% (42–87; nine of 13), 35% (19–54; nine of 26), and 83% (64–93; 20 of 24), respectively (table 2). Four patients with TRG1 residual tumours had false-positive results (one had a positive biopsy, one no-pass, one had an uncertain biopsy at the second clinical response evaluation, and one non-representative fine-needle aspiration specimen from suspicious lymph node).

Outcomes of bite-on-bite biopsies and fine-needle aspiration during clinical response evaluations were significantly associated with the TRG of resected specimens (p<0·0001; table 2). Four of 41 patients who had bite-on-bite biopsies and fine-needle aspiration had negative results despite having TRG3 or TRG4 tumours (proportion of clinically false negative cases 10% [95% CI 4–23]; table 2). 69 (60%) of 115 patients had positive bite-on-bite biopsies, positive fine-needle aspiration, or no-pass at endoscopy. After the first clinical response evaluation 45 (39%) of 115 patients who had bite-on-bite biopsies had a positive index test. All four of 41 patients with a negative bite-on-bite biopsy (false-negative

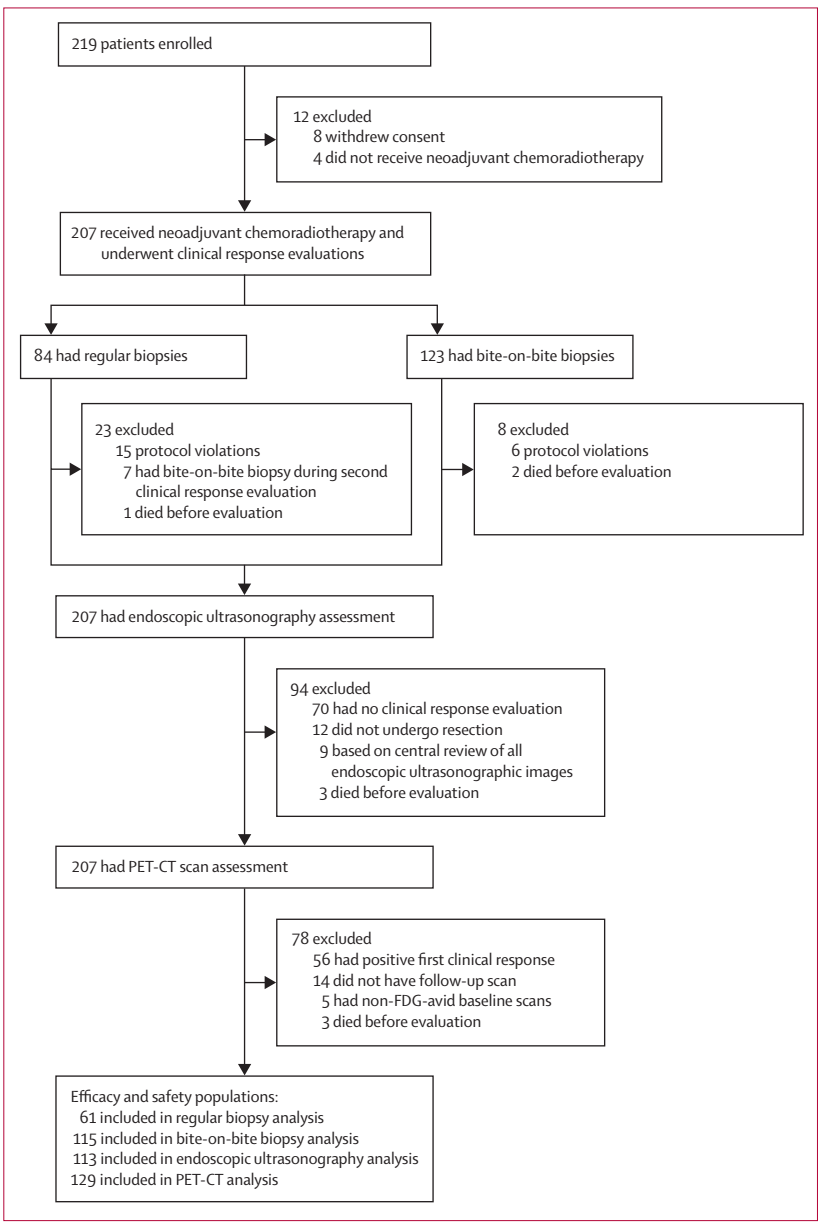

Figure: **Study profile**

	Regular biopsies (N=84)	Bite-on-bite biopsies (N=123)	Overall (N=207)
Median age, years (IQR)	65 (60–70)	66 (60–71)	66 (60–71)
Sex			
Male	72 (86%)	101 (82%)	173 (84%)
Female	12 (14%)	22 (18%)	34 (16%)
Tumour type			
Squamous cell carcinoma	22 (26%)	21 (17%)	43 (21%)
Adenocarcinoma	61 (73%)	102 (83%)	163 (78%)
Adenosquamous cell carcinoma	1 (1%)	0 (0%)	1 (<1%)
Clinical tumour stage*			
cT1	0 (0%)	1 (1%)	1 (<1%)
cT2	14 (17%)	26 (21%)	40 (19%)
cT3	66 (79%)	88 (72%)	154 (74%)
cT4	3 (4%)	8 (7%)	11 (5%)
Missing	1 (1%)	0 (0%)	1 (<1%)
Clinical lymph node stage			
N0	21 (25%)	42 (34%)	63 (30%)
N1	32 (38%)	48 (39%)	80 (39%)
N2	29 (35%)	28 (23%)	57 (28%)
N3	1 (1%)	5 (4%)	6 (3%)
Missing	1 (1%)	0 (0%)	1 (<1%)

Data are n (%), unless otherwise specified. Percentages might not total to 100% because of rounding. *Assessed by endoscopic ultrasonography or CT and classified according to the International Union against Cancer's TNM classification (7th edn).

Table 1: **Baseline characteristics of patients who underwent clinical response evaluations**

	TGR1	TRG2	TRG3	TRG4	p value*
Endoscopy with regular biopsies and fine-needle aspiration					0·0036
Positive	4/13 (31%)	2/11 (18%)	4/9 (44%)	14/17 (82%)	..
Negative	9/13 (69%)	9/11 (82%)	5/9 (56%)	3/17 (18%)	..
Endoscopy with bite-on-bite biopsies and fine-needle aspiration					<0·0001
Positive	5/18 (28%)	17/29 (59%)	20/24 (83%)	17/17 (100%)	..
Negative	13/18 (72%)	12/29 (41%)	4/24 (17%)	0/17	..
Endoscopic ultrasonography with maximum tumour thickness					0·035
Positive	11/27 (41%)	13/29 (45%)	13/20 (65%)	15/19 (79%)	..
Negative	16/27 (59%)	16/29 (55%)	7/20 (35%)	4/19 (21%)	..
PET–CT					0·191
Positive	17/27 (63%)	22/30 (73%)	17/19 (89%)	18/22 (82%)	..
Negative	10/27 (37%)	8/30 (27%)	2/19 (11%)	4/22 (18%)	..

*Calculated with the χ^2 test.

Table 2: **Clinical response evaluation outcomes per diagnostic modalities and tumour regression grade in patients who underwent oesophagectomy after neoadjuvant chemoradiotherapy**

cases 10%, 95% CI 11–21; table 2) had TRG3 residual disease—one patient had squamous cell carcinoma and three had adenocarcinomas. Sensitivity, specificity, negative predictive value, and positive predictive value of TRG2–4 versus TRG1 were 77% (95% CI 66–85; 54 of 70), 72% (49–88; 13 of 18), 45% (28–62; 13 of 29), and 92% (82–96; 54 of 59), respectively (table 2). Of the five patients with TRG1 residual tumours who had false-positive results, four (80%) were no-pass, and one (20%) had a non-representative fine-needle aspiration specimen from a suspicious lymph node.

Of the 69 patients with positive bite-on-bite biopsies, fine-needle aspirate, or no-pass, seven (10%) had positive fine-needle aspirates, but negative biopsies and a passable tumour (ie, seven of the 24 positive second clinical response evaluations were based on positive fine-needle-aspiration results only). On the basis of biopsy results only (ie, without fine-needle aspiration data), eight (31%) of 26 TRG3 or TRG4 tumours were missed with regular biopsies, and seven (17%) of 41 with bite-on-bite biopsies.

95 (84%) of 113 patients included in the endoscopic ultrasonographic examination during the second clinical response evaluation underwent oesophagectomy. Maximum tumour thickness of 6 mm or greater during the second clinical response evaluation was significantly associated with TRG of resection specimens (p=0·035; table 2). 11 (28%) of 39 patients with TRG3 or TRG4 residual disease had maximum tumour thickness of less than 6 mm at their second or third clinical response evaluation (proportion of clinically falsenegative cases 28% [95% CI 17–44]; table 2). Sensitivity, specificity, negative predictive value, and positive predictive value for TRG2–4 versus TRG1 residual disease were 60% (95% CI 48–71; 41 of 68), 59% (41–75; 16 of 27), 37% (24–52; 16 of 43), and 79% (66–88; 41 of 52), respectively (table 2).

Outcomes of PET–CT during the second clinical response evaluation were not significantly associated with tumour regression grades (p=0·191; table 2). Six of 41 patients with TRG3 or TRG4 residual tumours had negative PET–CT results (proportion of clinically false negative cases 15% [95% CI 7–28]; table 2). 102 (79%) of 129 patients had positive PET–CT results during the second or third clinical response evaluation. The six patients with false-negative PET–CTs comprised two (33%) patients with TRG3 tumours and four (67%) patients with TRG4 tumours. Sensitivity, specificity, negative predictive value, and positive predictive value for TRG2–4 versus TRG1 were 80% (95% CI 70–88; 57 of 71), 37% (22–56; ten of 27), 42% (24–61; ten of 24), and 77% (66–85; 57 of 74), respectively (table 2).

190 (92%) of 207 patients were included in the analysis of interval distant metastases (17 patients with missing follow-up scans were excluded: one participating centre did not do follow-up scans after a positive first clinical response evaluation). In 38 (20%) of 190 patients, PET–CT identified possible metastases, resulting in 18 (9%) cases of histologically proven metastases (one squamous cell carcinoma, 17 adenocarcinomas). Detection of distant metastases was more sensitive with PET–CT than with low-dose CT: ^{18}F-FDG-positive metastases would have been missed by CT in three patients; in another

	False-negative cases (95% CI)*	Sensitivity (95% CI)	Specificity (95% CI)	Negative predictive value (95% CI)	Positive predictive value (95% CI)
Endoscopy with regular biopsies and fine-needle aspiration	31% (13–49)	54% (38–70)	69% (44–94)	35% (16–53)	83% (68–98)
Endoscopy with bite-on-bite biopsies and fine-needle aspiration	11% (1–21)	74% (64–83)	77% (59–95)	45% (29–62)	92% (85–99)
Endoscopic ultrasonography with maximum tumour thickness (second clinical response evaluation)	29% (15–43)	59% (48–70)	58% (40–75)	38% (25–52)	76% (64–87)
PET–CT (second clinical response evaluation)	14% (3–24)	82% (73–90)	38% (21–55)	44% (26–63)	77% (68–87)

Accuracy estimates were calculated as TRG1 *vs* TRG2–4 after multiple imputation (for age, sex, histology, tumour grading, clinical T stage, clinical tumour stage, clinical lymph-node stage, WHO performance score, number of cycles of chemotherapy, total radiation dose, and results from endoscopic biopsies, fine-needle aspiration, maximum tumour thickness measurement, and PET–CT) per diagnostic modality for patients who had active surveillance instead of surgery after clinical response evaluations. Totals per group cannot be calculated, since this is a multiple imputation analysis. TRG=tumour regression grade. *Calculated as the proportion of TRG3 and TRG4 residual tumours missed during clinical response evaluations per diagnostic modality.

Table 3: **Sensitivity analysis for accuracy of residual tumour detection in clinical response evaluations and predictive value of the tumour regression grades**

three patients, distant lymph nodes smaller than 6 mm in diameter would probably not have been scored positive without a positive ^{18}F-FDG-PET. In at least two of the remaining 12 patients, the positive findings on PET increased the confidence.

No biopsy-related or fine-needle-aspiration-related serious adverse events were encountered during any clinical response evaluation in any patients included in the analyses. One patient had a mucosal tear during endoscopy, but this event did not have treatment implications. Two patients died during the study (one because of an aorto-oesophageal fistula and one because of pulmonary failure). Neither death was related to clinical response evaluations.

Sensitivity analysis after imputation of the TRG for patients who had active surveillance after clinical response evaluations showed a proportion of false-negative case rates for detection of residual tumour similar to those in the main analysis (table 3), for endoscopy with regular biopsies and fine-needle aspiration (31% [95% CI 13–49]), endoscopy with bite-on-bite biopsies and fine-needle aspiration (11% [1–21]), endoscopic ultrasonography with maximum tumour thickness at the second clinical response evaluation (29% [15–43]), and PET–CT at the second clinical response evaluation (14% [3–24]). The appendix shows outcomes for patients who were excluded from the analyses (p 2).

Discussion

To our knowledge, our trial is the first prospective study designed to assess the optimal composition of clinical response evaluations and the accuracy of residual disease detection in patients with oesophageal or junctional cancer. Repeated endoscopic ultrasonography with bite-on-bite biopsies and fine-needle aspiration of suspicious lymph nodes was safe, and missed 10% (95% CI 4–23) of TRG3 and TRG4 residual tumours after neoadjuvant chemoradiotherapy for oesophageal or junctional cancer. Endoscopic ultrasonography with regular biopsies and fine-needle aspiration, measurement of maximum tumour thickness, and PET–CT were less accurate to detect locoregional residual disease, as shown by the high proportion of false-negative cases. Results were similar in the sensitivity analyses, which used multiple imputation of TRG for patients who received active surveillance instead of oesophagectomy. PET–CT scans after neoadjuvant chemoradiotherapy detected new interval metastases in 9% (95% CI 6–14) of patients who had a pre-treatment or baseline PET–CT scan. These results provide insight into the optimal composition of clinical response evaluations after neoadjuvant chemoradiotherapy for patients with oesophageal or junctional cancer, and might help to stratify patients who would benefit from active surveillance and those who should undergo oesophagectomy.

In view of the substantial postoperative morbidity and mortality associated with surgery, and the effect of surgery on quality of life, an active surveillance approach could improve outcomes, not only for patients who do not show signs of disease after neoadjuvant chemoradiotherapy, but also for those with subclinical distant metastases after neoadjuvant chemoradiotherapy.[3,5–7] All available diagnostic modalities used in clinical practice for pre-treatment staging were applied to detect residual disease, and compared to establish the optimal composition of future active surveillance strategies. Previous studies[17–22] of clinical response evaluations were retrospective and examined a single diagnostic modality for residual disease detection. Furthermore, the main objective of diagnostic examinations in previous studies[17–22] was not to detect residual disease to identify patients who might benefit from active surveillance. Therefore, diagnostic accuracy might have not been accurately estimated.

Biopsies were more accurate in our study than reported previously.[19,22,23] Possible explanations for this increased accuracy are the timepoints chosen for the first and second clinical response evaluations, and the adherence to a strict, pre-specified protocol in our trial, including

random biopsies from the site of the primary tumour and targeted biopsies from any suspicious lesions. Fine-needle aspiration of suspicious lymph nodes also increased the sensitivity of the clinical response evaluation assessments in patients with negative biopsies. The percentage of TRG3 or TRG4 residual tumours that was missed by endoscopy plus regular biopsies and fine-needle aspiration decreased from 31% (95% CI 17–50) to 10% (4–23) after introduction of bite-on-bite biopsies, and the negative predictive value increased from 35% (95% CI 19–54) to 45% (28–62). Residual disease is often located in the oesophageal mucosa, or the deeper submucosa, but can be rarely also present in isolated remnants within the muscle layer or the surrounding stroma (deeper than the submucosa).[10] Bite-on-bite biopsies are thought to increase the chance of detecting residual cancer cells in deeper layers of the oesophagus, such as the submucosa, compared with regular biopsies, which rarely penetrate the submucosa (appendix p 7).

Although the diagnostic accuracy of PET–CT for the detection of locoregional residual disease is poor, PET–CT was useful for detection of interval distant metastases (in 9% [95% CI 6–14] of all patients) during clinical response evaluations. The extended period from the end of neoadjuvant chemoradiotherapy to PET–CT during the second clinical response evaluation supposedly improved the signal-to-noise ratio, because artifacts related to radiation-induced oesophagitis were expected to have diminished. Nevertheless, results were similar to those noted in previous studies.[17,21,24] During active surveillance, PET-CT is expected to detect distant metastases, thereby preventing oesophagectomy in patients with initially subclinical distant metastases. In view of the high frequency of false positivity (63% [95% CI 44–78] of TRG1 tumours) of PET–CT for detection of locoregional disease and the limited additional value as an adjunct to endoscopy with bite-on-bite biopsies and fine-needle aspiration, we propose that PET–CT should primarily be used for detection of distant metastases during response evaluations. However, during active surveillance, serial PET–CT might prove valuable for detection of local regrowths: an increase in ^{18}F-FDG-avidity theoretically suggests disease recurrence, whereas a decrease is more likely to depict recovery from oesophagitis.

Results of measurement of maximum tumour thickness were similar to those from an earlier study.[9] However, diagnostic accuracy of maximum tumour thickness was worse than that of endoscopic ultrasonography with bite-on-bite biopsies and fine-needle aspiration, with 28% (95% CI 17–44) of TRG3 and TRG4 tumours missed and a frequency of false-positive diagnosis of 41% (95% CI 25–59) for TRG1 tumours. Taken together, we recommend that clinical response evaluations after neoadjuvant chemoradiotherapy in patients with oesophageal or gastroesophageal junction cancer should consist of repeated endoscopy with bite-on-bite biopsies and fine-needle aspiration of suspicious lymph nodes for detection of locoregional residual disease and PET–CT for detection of interval metastases.

The minimum diagnostic accuracy for safe active surveillance will continue to be debated until a clinical trial is done to establish it. Even a very small amount of residual disease (eg, TRG2) should ideally not be missed during clinical response evaluations, because patients with residual viable cancer cells do not benefit from an active surveillance strategy and should have oesophagectomy as soon as possible. Conversely, if locoregional residual disease is initially missed, but can be detected during active surveillance while the tumour is still resectable, oncological outcomes should not be worse. Evidence of successful active surveillance strategies in patients with head and neck, rectal, or bladder cancer,[25–27] supports the adoption of active surveillance after neoadjuvant chemoradiotherapy in patients with oesophageal cancer. A systematic review[7] showed that postponed radical resection was associated with good survival outcomes (ie, similar to those with standardsurgery; median overall survival 58 months [95% CI 27·7 to not reached]) in most patients with oesophageal cancer who showed locoregional regrowth during active surveillance.[7,26,28,29] This median overall survival in complete responders to neoadjuvant chemoradiotherapy managed with active surveillance is similar to that of patients with a complete clinical response who undergo surgery after neoadjuvant chemoradiotherapy.[28,29] In these studies, clinical response was assessed by endoscopy with regular biopsies and PET–CT. The use of bite-on-bite biopsies and the addition of fine-needle aspiration from suspicious lymph nodes could increase diagnostic accuracy. The promising results of our study in combination with those of previous publications justify a randomised, controlled, phase 3 trial of active surveillance versus standard surgery. The results of our study could serve to define the composition of the clinical response evaluations and the subsequent surveillance examinations in such trials.

Investigators of the ongoing, randomised, phase 3 ESOSTRATE and SANO trials are comparing both treatment strategies.[12] Both trials aim to include 300 patients with squamous cell carcinoma or adenocarcinoma of the oesophagus who were clinical complete responders after neoadjuvant chemoradiotherapy. Although pathological complete responses are more likely in patients with squamous cell carcinoma (49%), they are also common in those with adenocarcinoma (23%) after carboplatin and paclitaxel combined with 41·4 Gy radiotherapy with low toxicity.[3] The activity–toxicity ratio in both histological subtypes is the rationale for the use of this regimen in the preSANO and SANO trials, rather than a definitive chemoradiotherapy regimen without surgery.[3] Furthermore, our results show that the risk of false-negative biopsies during clinical response evaluations is not higher in patients with adenocarcinoma than in those with squamous cell

carcinoma. The primary endpoint of the ESOSTRATE trial is disease-free survival and overall survival in the SANO trial. On the basis of the results of our study, clinical response evaluations in the SANO trial consist of repeated endoscopic ultrasonography with bite-on-bite biopsies and fine-needle aspiration, plus PET–CT scans for detection of distant metastases. Patients with negative results in the first and second clinical response evaluations after neoadjuvant chemoradiotherapy will be classified as clinical complete responders and allocated to either active surveillance or immediate surgery on the basis of stepped-wedge cluster randomisation.

The diagnostic accuracy of clinical response evaluations is expected to improve, reducing the number of patients who need postponed oesophagectomy or who have irresectable regrowths during active surveillance. Dynamic contrast-enhanced MRI and diffusion-weighted MRI are promising new techniques that will need to be assessed in larger diagnostic trials.[30,31] Furthermore, the incorporation of liquid biopsies to analyse circulating cell-free tumour DNA derived from blood samples might improve the prediction of response to neoadjuvant chemoradiotherapy and the detection of disease recurrence during active surveillance.

Limitations of our study include the change in the biopsy strategy during the trial. The improved diagnostic accuracy of bite-on-bite biopsies compared with regular biopsies could be explained by a learning-curve effect. The assumption that TRG2 residual disease can be safely missed during initial response evaluation assessments is based on the hypothesis that these tumours can be reliably detected as they progress to stage TRG3 or TRG4, and that surgery will still be a curative option at this point. However, we acknowledge that this assumption can be only formally tested by comparing active surveillance with standard resection in a randomised, controlled, clinical trial. In the SANO trial, strict stopping rules have been prespecified by protocol for timely detection of resectable locoregional regrowth (any T stage <T4b) and the feasibility of achieving radical resection in the active surveillance arm. Because of the small number of patients with squamous cell carcinoma included in this study, the extent to which our results can be generalised is unclear. However, active surveillance after definitive chemo-radiotherapy is a standard of care in many centres for patients with squamous cell carcinoma, and is the recommend standard of care in some guidelines.[32] Nevertheless, if a patient has a clinical complete response based on endoscopy with bite-on-bite biopsies and fine-needle aspiration, the risk that there is any residual disease left seems similar in both subgroups of patients with oesophageal cancer, those with squamous cell carcinoma and those with adenocarcinoma. Repeat CT of the thorax, abdomen, and pelvis was not done as part of the first clinical response evaluation. Furthermore, use of a limited range of diagnostic modalities during the first clinical response evaluation could have reduced the accuracy of residual disease detection in the first clinical response evaluation compared with the second. Additionally, to include any degree of possible tumour residue, we defined all PET–CT scores of 2 (probably benign) as ^{18}F-FDG-positive, and thus probably included some cases with radioisotope uptake due to oesophagitis rather than oesophageal malignancy, resulting in overdiagnosis. For the same reason, fine-needle aspiration specimens taken from suspicious lymph nodes with uncertain outcomes or that were not representative were classified as positive, but should not be considered to be correctly diagnosed. Finally, overall and progression-free survival data according to TRG will be published when follow-up is sufficient.

In conclusion, clinical response evaluation comprising endoscopic ultrasonography with bite-on-bite biopsies and fine-needle aspiration of suspicious lymph nodes for detection of locoregional residual disease in combination with PET–CT for detection of interval metastases after neoadjuvant chemoradiotherapy for oesophageal or gastroesophageal junctional cancer is an adequate strategy for clinical response evaluation. The ongoing, randomised, phase 3 SANO trial (Netherlands Trial Register: NTR6803) has incorporated this diagnostic strategy and will compare active surveillance with standard resection in patients who achieve a complete response after neoadjuvant chemoradiotherapy.

Contributors

BJN, MCWS, RV, MIvBH, JS, KB, AvdG, RvH, MCCMH, KKK, SML, GAPN, LEO, PDS, EJS, MNS, and EWS had roles in study design. BJN drafted the Article. BPLW and JJBvL initiated the trial and supervised drafting of the Article, which was critically revised and approved by all authors.

Declaration of interests

EWS has received royalties from Springer for a book on prediction models. JJBvL has received research grants from the Dutch Cancer Society, Coolsingel Stichting, and the Erasmus MC/MRace Fund. All other authors declare no competing interests.

Acknowledgments

We thank Daan Nieboer (Department of Public Health, Erasmus MC, University Medical Centre, Rotterdam, Netherlands) for assistance with statistical analysis and data interpretation, and Maartje Valkema (Departments of Surgery and Radiology and Nuclear Medicine, Erasmus MC, University Medical Centre, Rotterdam, Netherlands) for assistance with analysis of PET–CT data.

References

1 Hulscher JBF, van Sandick JW, de Boer AGEM, et al. Extended transthoracic resection compared with limited transhiatal resection for adenocarcinoma of the esophagus. *N Engl J Med* 2002; **347:** 1662–69.

2 Sjoquist KM, Burmeister BH, Smithers BM et al. Survival after neoadjuvant chemotherapy or chemoradiotherapy for resectable oesophageal carcinoma: an updated meta-analysis. *Lancet Oncol* 2011; **12:** 681–92.

3 Van Hagen P, Hulshof MC, van Lanschot JJ, et al. Preoperative chemoradiotherapy for esophageal or junctional cancer. *N Engl J Med* 2012; **366:** 2074–84.

4 Shapiro J, van Lanschot JJB, Hulshof MCCM, et al. Neoadjuvant chemoradiotherapy plus surgery versus surgery alone for oesophageal or junctional cancer (CROSS): long-term results of a randomised controlled trial. *Lancet Oncol* 2015; **16:** 1090–98.

5 Noordman BJ, Verdam MGE, Lagarde SM, et al. Effect of neoadjuvant chemoradiotherapy on health-related quality of life in esophageal or junctional cancer: results from the randomized CROSS trial. *J Clin Oncol* 2018; **36:** 268–75

6 Noordman BJ, Verdam MGE, Lagarde SM, et al. Impact of neoadjuvant chemoradiotherapy on health related quality of life in long-term survivors of esophageal or junctional cancer: results from the randomized CROSS trial. *Ann Oncol* 2018; **29:** 445–51.

7 Noordman BJ, Wijnhoven BPL, Lagarde SM, et al. Active surveillance in clinically complete responders after neoadjuvant chemoradiotherapy for esophageal or junctional cancer. *Dis Esophagus* 2017; **30:** 1–8.

8 Noordman BJ, Shapiro J, Spaander MC, et al. Accuracy of detecting residual disease after CROSS neoadjuvant chemoradiotherapy for esophageal cancer (preSANO Trial): rationale and protocol. *JMIR Res Protoc* 2015; **4:** e79.

9 Jost C, Binek J, Schuller JC, et al. Endosonographic radial tumor thickness after neoadjuvant chemoradiation therapy to predict response and survival in patients with locally advanced esophageal cancer: a prospective multicenter phase ll study by the Swiss Group for Clinical Cancer Research (SAKK 75/02). *Gastrointest Endosc* 2010; **71:** 1114–21.

10 Shapiro J, ten Kate FJ, van Hagen P, Biermann K, Wijnhoven BP, van Lanschot JJ. Residual esophageal cancer after neoadjuvant chemoradiotherapy frequently involves the mucosa and submucosa. *Ann Surg* 2013; **258:** 678–89.

11 Shapiro J, van Hagen P, Lingsma HF, et al. Prolonged time to surgery after neoadjuvant chemoradiotherapy increases histopathological response without affecting survival in patients with esophageal or junctional cancer. *Ann Surg* 2014; **260:** 807–14.

12 Noordman BJ, Wijnhoven BPL, Lagarde SM, et al. Neoadjuvant chemoradiotherapy plus surgery versus active surveillance for oesophageal cancer: a stepped-wedge cluster randomised trial. *BMC Cancer* 2018; **18:** 142.

13 Boellaard R, O'Doherty MJ, Weber WA, et al. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. *Eur J Nucl Med Mol Imaging* 2010; **37:** 181–200.

14 Sobin LH, Gospodarowicz MK, Wittekind C, International Union against Cancer. TNM classification of malignant tumors. New York: Wiley-Blackwell, 2009.

15 Chirieac LR, Swisher SG, Ajani JA, et al. Posttherapy pathologic stage predicts survival in patients with esophageal carcinoma receiving preoperative chemoradiation. *Cancer* 2005; **103:** 1347–55.

16 Naaktgeboren CA, de Groot JA, Rutjes AW, Bossuyt PM, Reitsma JB, Moons KG. Anticipating missing reference standard data when planning diagnostic accuracy studies. *BMJ* 2016; **352:** i402.

17 Heneghan HM, Donohoe C, Elliot J, et al. Can CT-PET and endoscopic assessment post-neoadjuvant chemoradiotherapy predict residual disease in esophageal cancer? *Ann Surg* 2016; **264:** 831–38.

18 Piessen G, Petyt G, Duhamel A, Mirabel X, Huglo D, Mariette C. Ineffectiveness of (1)(8)F-fluorodeoxyglucose positron emission tomography in the evaluation of tumor response after completion of neoadjuvant chemoradiation in esophageal cancer. *Ann Surg* 2013; **258:** 66–76.

19 Schneider PM, Metzger R, Schaefer H, et al. Response evaluation by endoscopy, rebiopsy, and endoscopic ultrasound does not accurately predict histopathologic regression after neoadjuvant chemoradiation for esophageal cancer. *Ann Surg* 2008; **248:** 902–08.

20 Kim MK, Ryu JS, Kim SB, et al. Value of complete metabolic response by (18)F-fluorodeoxyglucose-positron emission tomography in oesophageal cancer for prediction of pathologic response and survival after preoperative chemoradiotherapy. *Eur J Cancer* 2007; **43:** 1385–91.

21 Elliott JA, O'Farrell NJ, King S, et al. Value of CT-PET after neoadjuvant chemoradiation in the prediction of histological tumour regression, nodal status and survival in oesophageal adenocarcinoma. *Br J Surg* 2014; **101:** 1702–11.

22 Cheedella NK, Suzuki A, Xiao L, et al. Association between clinical complete response and pathological complete response after preoperative chemoradiation in patients with gastroesophageal cancer: analysis in a large cohort. *Ann Oncol* 2013; **24:** 1262–66.

23 Sarkaria IS, Rizk NP, Bains MS, et al. Post-treatment endoscopic biopsy is a poor-predictor of pathologic response in patients undergoing chemoradiation therapy for esophageal cancer. *Ann Surg* 2009; **249:** 764–67.

24 Blom RL, Schreurs WM, Belgers HJ, Oostenbrug LE, Vliegen RF, Sosef MN. The value of post-neoadjuvant therapy PET-CT in the detection of interval metastases in esophageal carcinoma. *Eur J Surg Oncol* 2011; **37:** 774–78.

25 Mehanna H, Wong WL, McConkey CC, et al. PET-CT surveillance versus neck dissection in advanced head and neck cancer. *N Engl J Med* 2016; **374:** 1444–54.

26 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:** 174–83.

27 Rödel C, Wiess C, Sauer R. Trimodality treatment and selective organ preservation for bladder cancer. *J Clin Oncol* 2006; **24:** 5536–44.

28 Castoro C, Scarpa M, Cagol M, et al. Complete clinical response after neoadjuvant chemoradiotherapy for squamous cell cancer of the thoracic oesophagus: is surgery always necessary? *J Gastrointest Surg* 2013; **17:** 1375–81.

29 Taketa T, Xiao L, Sudo K, et al. Propensity-based matching between esophagogastric cancer patients who had surgery and who declined surgery after preoperative chemoradiation. *Oncology* 2013; **85:** 95–99.

30 Heethuis SE, van Rossum PS, Lips IM, et al. Dynamic contrast-enhanced MRI for treatment response assessment in patients with oesophageal cancer receiving neoadjuvant chemoradiotherapy. *Radiother Oncol* 2016; **120:** 128–35.

31 Van Rossum PS, van Lier AL, van Vulpen M et al. Diffusion-weighted magnetic resonance imaging for the prediction of pathologic response to neoadjuvant chemoradiotherapy in esophageal cancer. *Radiother Oncol* 2015; **115:** 163–70.

32 Lordick F, Mariette C, Haustermans K, Obermannová R, Arnold D, for the ESMO Guidelines Committee. Oesophageal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2016; **27** (suppl): v50–57.