STS/ASTRO Updated Clinical Practice Guidelines on Multimodality Therapy for Locally-Advanced Cancer of the Esophagus or Gastroesophageal Junction

Stephanie G. Worrell, MD, Chair
Karyn A. Goodman, MD, Vice-chair
Nasser K. Altorki, MD
Jonathan B. Ashman, MD
Traves D. Crabtree, MD
Jennifer Dorth, MD
Scott Firestone
David H. Harpole, MD
Wayne L. Hofstetter, MD

Theodore S. Hong, MD
Kalie Kissoon
Geoffrey Y. Ku, MD
Daniela Mola, MD
Joel E. Tepper, MD
Thomas J. Watson, MD
Terence Williams, MD, PhD
Christopher Willett, MD

1Department of Surgery, Thoracic Surgery Section, University of Arizona College of Medicine, Tucson, AZ
2Department of Radiation Oncology, Icahn School of Medicine at Mount Sinai, New York, NY
3Division of Thoracic Surgery, Weill Cornell Medicine, New York-Presbyterian Hospital, New York, NY
4Department of Radiation Oncology, Mayo Clinic, Phoenix, AZ
5Division of Cardiothoracic Surgery Department of Surgery, Southern Illinois University School of Medicine, Springfield, IL
6Department of Radiation Oncology, Seidman Cancer Center, University Hospitals, Cleveland, OH
7The Society of Thoracic Surgeons, Chicago, IL
8Division of Cardiovascular and Thoracic Surgery, Department of Surgery, Duke University Medical Center, Durham, NC
9Thoracic and Cardiovascular Surgery, University of Texas MD Anderson Cancer Center, Houston, TX
10Department of Radiation Oncology, Massachusetts General Hospital, Boston, MA
11Department of Medicine, Gastrointestinal Medical Oncology Service, Memorial Sloan Kettering Cancer Center, New York, NY
12Department of Surgery, Division of Thoracic Surgery, Memorial Sloan Kettering Cancer Center New York, NY
13Department of Radiation Oncology, University of North Carolina, Chapel Hill, NC
14Thoracic Surgery Group, Beaumont Health, Royal Oak, MI
15Department of Radiation Oncology, Beckman Research Institute, City of Hope National Medical Center, Duarte, CA
16Department of Radiation Oncology, Duke University Medical Center, Durham, NC
Introduction

Outcomes for patients with esophageal cancer have improved over the last decade with the implementation of multimodality therapy (1). However, the specific way in which multimodality care is implemented varies widely based on physician bias and the wide range of available literature. There are currently no comprehensive guidelines addressing multi-disciplinary management of esophageal cancer that have incorporated the input of surgeons, radiation oncologists and medical oncologists. Most published guidelines in the literature are developed by individual medical specialty societies. To address the need for multidisciplinary input in the management of esophageal cancer and to meet current best practices for clinical practice guidelines, the current guidelines were created as a collaboration between the Society of Thoracic Surgeons (STS), American Society for Radiation Oncology (ASTRO) and the American Society of Clinical Oncology (ASCO). These practice guidelines address seven key clinical questions pertinent to the care of patients with locally advanced, resectable thoracic esophageal cancer (excluding cervical location). These questions include the use of induction chemotherapy, chemotherapy versus chemoradiation prior to surgery, the timing of esophagectomy, the value of esophagectomy, the approach and extent of lymphadenectomy, the use of minimally invasive esophagectomy, and the value of adjuvant therapy after resection.

Methodology
In 2020, the STS Workforce on Evidence-Based Surgery assembled a Task Force with representation from ASTRO and ASCO, to update the 2014 STS Practice Guidelines on the Role of Multimodality Therapy for the Treatment of Esophageal Cancer.

The members of the writing committee submitted conflict of interest disclosure forms, which were reviewed by the Chair and STS staff before confirmation for potential conflicts from relevant relationships with industry.

The writing committee reviewed the topics covered by the 2014 Guidelines and developed nine questions in the Population, Intervention, Comparator, and Outcomes format (PICO) intended to focus on the highest priority and most clinically impactful areas for a systematic review. The PICO questions were sent to a research librarian in March 2021 to develop a strategy to identify all relevant articles published in English since 2000. Strategies were developed for both MEDLINE and Embase, the details for which may be found in Appendix 1. Reference lists were manually scanned for additional relevant articles. This strategy resulted in 2,133 potentially relevant abstracts after duplicate studies were removed, and an additional 60 studies were identified as potentially relevant, for a total of 2,193 abstracts. In June 2022, two additional PICO questions on radiation dose were added, which resulted in 1,972 additional abstracts using MEDLINE and Embase. The results were refined using filters designed to identify randomized controlled trials (RCTs) (2) and comparative studies (3). Two authors (S.F., K.K.) screened the results, identifying a total of 227 articles which met the inclusion criteria (Figure 1). Reasons for exclusion were that the study design, intervention, or the primary outcomes were not relevant to the PICO questions posed.
Two authors (S.F., K.K.) developed an evidence table of the relevant papers (APPENDIX 2) and rated the studies for risk of bias. The Newcastle-Ottawa scale was used for observational studies (APPENDIX 3), and a custom-made checklist was used for RCTs and meta-analyses (APPENDIX 4). A meta-analysis was performed using Review Manager Version 5.4 (The Cochrane Collaboration) when several studies reporting the same outcome were found and there were no previous meta-analyses in the literature. Random effects models were chosen due to heterogeneity in study populations and the assumption that there is between-study uncertainty in addition to that within studies.

Voting on recommendations used a modified Delphi method of two rounds of voting to reach consensus, in which responses were required by 80% of the authors, with 75% agreement on class and level of evidence as defined by the American College of Cardiology (ACC)/American Heart Association (AHA) Classification System (APPENDIX 5).

The resulting manuscript was reviewed by the STS Workforce on Evidence-Based Surgery, the STS Council Operating Board on Quality, Research, and Patient Safety, and the Executive Committee, along with a two-week member comment period available to members of all participating societies. Three peer reviewers were selected by ASTRO to review the manuscript, after which the Board of Directors of ASTRO also reviewed the document.

These guidelines were developed by the participating societies without commercial support and will be reviewed for a potential update within five years of publication.

**Induction Chemotherapy**
• Induction chemotherapy prior to preoperative chemoradiation with a PET-based response assessment and adaptation of the regimen accordingly during chemoradiation may be reasonable in patients with resectable esophageal adenocarcinoma (Class IIB, Level of Evidence B-NR)

• Administration of induction chemotherapy prior to neoadjuvant chemoradiotherapy without early response assessment and response-adapted therapy during radiotherapy is not recommended. (Class III: No Benefit, Level of Evidence B-R)

The role of induction chemotherapy followed by chemoradiation for patients with resectable, Stage II-III esophageal cancer is a subject of ongoing debate. The rationale for the addition of chemotherapy in the neoadjuvant setting is to address the high rate of distant failure and poor tolerance of adjuvant chemotherapy. Adjuvant chemotherapy trials have demonstrated poor tolerance of additional chemotherapy following preoperative therapy and surgery (4, 5).

Nonetheless, retrospective studies evaluating the use of induction chemotherapy followed by pre-operative chemoradiation versus pre-operative chemoradiation have had conflicting results (6-9).

**Treatment Response Assessment by PET**

A benefit of induction chemotherapy is that it permits for response assessment by PET imaging prior to combining chemotherapy with radiotherapy, allowing for early evaluation of treatment response to a specific therapeutic regimen and the potential to change the chemotherapy during the radiotherapy if there is not an optimal response. The CALGB 80803 Phase II Randomized Trial of PET-Response-Adapted Combined Modality Therapy for Esophageal Cancer, evaluated this question of changing therapy based on PET response after induction.
chemotherapy (10). This study randomized 257 esophageal and GEJ cancer patients after a baseline PET scan to induction carboplatin/paclitaxel or FOLFOX followed by re-staging PET scan. PET-responders (SUVmax decreased by $\geq 35\%$ from baseline) continued on with the same chemotherapy with pre-operative radiotherapy, whereas PET-nonresponders (SUVmax decreased by $<35\%$ from baseline) crossed over to the alternate chemotherapy during chemoradiation. Half of patients receiving induction carboplatin/paclitaxel and 57\% of those receiving induction FOLFOX were PET-responders to induction chemotherapy.

**Pathological Complete Response (pCR)**

In the CALGB 80803 trial, the primary objective was to improve pCR rates in the PET-nonresponders from 5\% to 20\% by changing chemotherapy. A total of 198 patients who completed induction chemotherapy and chemoradiation and underwent surgical resection were analyzed for the primary endpoint of pCR. Among the entire group, the pCR rate was 22.7\% and the pCR rate for PET nonresponders who switched from FOLFOX to carboplatin/paclitaxel was 18\% and for those who switched from carboplatin/paclitaxel to FOLFOX was 20\% which met the pre-specified efficacy criteria. Although CALGB 80803 was not intended to evaluate the induction chemotherapy question itself since all patients received induction chemotherapy, it did demonstrate a benefit to changing chemotherapy after patients were determined to be PET-non responders after induction chemotherapy. Of note, the highest pCR rate of 40\% occurred in PET responders who received both induction and concurrent FOLFOX.

Ajani and colleagues conducted a randomized phase II trial at MD Anderson Cancer Center comparing induction chemotherapy added to preoperative chemoradiotherapy compared to
chemoradiotherapy alone, followed by surgery, with the primary end point to improve the rate of pCR (8). A total of 126 patients received 50.4 Gy of radiation therapy given with weekly oxaliplatin/5-FU for 5 weeks, with or without the inclusion of four cycles of oxaliplatin/5-FU prior to the start of radiation therapy. The authors reported a non-significant trend toward an improved pCR rate on the induction chemotherapy arm compared to the chemoradiation-only patients (26% vs. 13%, p=0.094). Unlike the CALGB 80803 study, there was no selection of concurrent chemotherapy during radiotherapy based on response; all patients received the same chemotherapy for the induction phase and concurrent chemoradiation phase. A smaller, Phase II randomized trial of pre-operative chemoradiotherapy with or without induction chemotherapy of tegafur-uracil (S1) and oxaliplatin was reported by Yoon and colleagues. The majority (98%) of these patients had esophageal SCC in contrast to the MD Anderson study which was only adenocarcinomas. These investigators found no improvement in pCR rates with induction chemotherapy (11, 12). Of note, the addition of induction S1 and oxaliplatin before preoperative chemoradiotherapy resulted in a lower pCR rate than in the upfront chemoradiation arm, thus suggesting that the induction chemotherapy approach is not advisable for patients with resectable esophageal SCC.

**Overall Survival (OS)**

In the CALGB 80803 trial, with a median follow-up of 5.2 years, median OS was 48.8 months for PET responders and 27.4 months for non-responders which was not significantly different (p=0.1). The MUNICON Trial, which evaluated early PET response after induction chemotherapy demonstrated that PET non-responders had a significantly worse OS than PET responders when
they continue on with the same chemotherapy (13). Thus, these results suggest that changing chemotherapy based on PET response moved the survival curve in the PET non-responder group close to that of the PET-responder group. Moreover, the patients who were PET-responders to induction FOLFOX had the highest 5-year OS of 53% and the median survival was not reached (10). Recent retrospective data, comparing outcomes among 451 esophageal adenocarcinoma patients treated with neoadjuvant chemoradiation with concurrent CP, induction CP or induction FOLFOX followed by PET-adaptive therapy during chemoradiation, demonstrated that use of induction FOLFOX led to higher rates of pathologic response and 2-year DFS than either induction CP followed by PET-directed chemoradiation or neoadjuvant chemoradiation with CP alone. In addition, this approach did not increase the risk of postoperative complications (14). Thus, induction FOLFOX appears to be the better regimen if using PET-adaptive therapy for patients with resectable esophageal adenocarcinoma.

Ajani et al. reported no improvement in median OS for patients who received induction chemotherapy (43.7 vs. 45.6 months, p=0.69) (12). However, an analysis of the long-term data demonstrated that induction chemotherapy significantly prolonged OS of the well- and moderately-differentiated adenocarcinoma patients (15), suggesting that the chemotherapy regimen may not have been optimal for the high-grade/poorly differentiated patients. However, based on the improvement in OS for the well- and moderately-differentiated adenocarcinoma subgroup, the induction chemotherapy approach for esophageal adenocarcinoma patients should be evaluated in larger prospective trials, possibly including a response-adapted design. In Yoon et al, there was also no survival benefit with induction chemotherapy in the esophageal SCC population. Thus, based on these studies, the benefit for
induction chemotherapy prior to chemoradiation is limited to patients with esophageal adenocarcinomas, particularly when response assessment after induction chemotherapy and tailoring of subsequent therapy is performed.

**Limitations**

There are numerous limitations to all of the studies, including varying histologies, pathological stages, chemotherapeutic and radiotherapy regimens, surgical techniques, all of which make interpretation and comparisons difficult. All of the retrospective studies share a limitation that patients who received induction CT may not be representative of the larger population with locally-advanced resectable esophageal cancer, and no study attempted to match groups apart from one that limited matching only to those who fit the definition of “high risk” according to a prognostic model they developed.

The small sample sizes in the prospective studies raise the possibility that the negative results might be due to a lack of statistical power. Moreover, both RCTs were assessed as having a risk of bias, particularly the trial by Yoon et al that failed to report a number of methodological details and had a rather high rate of patients who did not continue on to surgery. The subsequent analysis of the Phase II trial from MD Anderson Cancer Center opens the door for a larger trial to assess different chemotherapeutic regimens targeted at well- and moderately differentiated adenocarcinomas versus poorly-differentiated adenocarcinomas. Difficulties in interpreting a heterogeneous dataset aside, it is unlikely that induction CT in and of itself, without early response-assessment and changes in chemotherapy for non-responders, is beneficial to all patients presenting with resectable esophageal cancer, particularly those with esophageal SCC. Future studies should focus on identifying the optimal biomarkers of
treatment response to better tailor neoadjuvant therapy for patients with esophageal cancers of both histologies.

Neoadjuvant Chemotherapy vs. Neoadjuvant Chemoradiotherapy

- In patients with locally-advanced SCC of the esophagus, neoadjuvant chemoradiation is reasonable to choose over neoadjuvant chemotherapy. (Class IIA, Level B-R)

- In patients with locally-advanced adenocarcinoma of the esophagus or gastroesophageal junction, either neoadjuvant chemoradiation or neoadjuvant chemotherapy are reasonable to choose. (Class IIA, Level B-R)

Neoadjuvant chemoradiotherapy (nCRT) has consistently been demonstrated to result in superior pathological complete response rates (pCR) and complete (R0) resection rates than neoadjuvant chemotherapy (nCT) in patients with resectable adenocarcinoma (AC). These results are even more pronounced in squamous cell carcinoma (SCC), regardless of regimen and radiation dose (16). Whether this translates to better survival is debated. Pooled data from meta-analyses point to better overall survival with nCRT, with the benefit more pronounced in SCC over AC.

pCR

Four meta-analyses reported significantly increased rates of pathological complete response for nCRT over nCT, with RRs/ORs ranging from 2.90 to 6.48. The most recent of these analyses by Han et al had the most precise estimate from over 2,000 patients with low heterogeneity [26.1% vs. 6.0%, RR: 3.61 (95% confidence interval (CI), 2.66–4.90) p < 0.001] (17), although this
precision is likely to be misleading due to use of a fixed-effect model despite varying patient populations. The other three, although all highly significant, had considerably wider CI's with little to no heterogeneity between studies (18-20).

The meta-analyses by Deng et al and Han et al both investigated the effect by histology and found that the rate of pCR was higher regardless of whether the patients had AC or SCC. Deng reported a higher RR in SCC pts (6.73 vs. 4.69), although the sample sizes were relatively small (18). Additionally, the SCC sub-analysis had considerable heterogeneity ($I^2 = 71.8\%$). Han reported more similar effects between histological subtypes with greater precision (AC RR: 3.48 vs. SCC RR: 3.68), again with higher heterogeneity in the SCC analysis (17).

### OS

Several meta-analyses report 3- or 5-year survival, while others report overall survival of undetermined duration. The Han meta-analysis reported a minor but statistically significant benefit at 3 years for nCRT over nCT [RR: 1.15 (95% CI, 1.05–1.25) $p = 0.003$ $I^2 = 30.1\%$] that did not hold up at 5 years (17). The survival advantage was entirely due to patients with SCC, a finding that was supported by a meta-analysis by Deng et al (18). Another meta-analysis by Fan et al found a benefit for nCRT [HR 0.73 (95% CI, 0.61–0.89) $p = 0.02$ $I^2 = 0.0\%$] with both AC and SCC patients included (19). A network meta-analysis by Pasquali found no significant difference in OS, although nCRT was given a higher probability of being the better treatment option than nCT (21).

Interestingly, meta-analyses by Huang et al, Li et al, and Montagnani et al in only SCC patients did not report such a clear benefit for nCRT. Li et al reported a marginally-significant benefit with low heterogeneity [HR: 0.72 (95% CI, 0.52–0.99) $p = 0.046$], although the other two studies
were negative (20, 22, 23). Montagnani et al barely missed significance, and as a result, in a
network meta-analysis it was found to have a higher probability of being the second treatment
option over nCT, with definitive CRT being the best (23).

Individual prospective randomized studies have consistently found negative results, although
this may generally be a result of a Type II error rather than lack of an effect. A recently
published abstract for a randomized phase III study also points to no difference in survival
outcomes (24).

**DFS**
The meta-analysis by Fan et al was the only study that reported pooled data on DFS, finding a
benefit for nCRT, albeit with a small sample size and high heterogeneity. $[\text{HR}: 0.73, \ (95\% \ CI, 0.54–0.98) \ p = 0.037, \ I^2 = 64\%] \ (19)$. Three RCTs reported DFS, with two negative results by Von Döbeln et al and Burmeister et al (25, 26). The POET trial by Stahl et al reported a lower HR (0.37) for nCRT (95% CI, 0.16-0.85; $p = 0.01$) with a more specific inclusion criteria of only AC patients with Siewert I-II tumors and a longer CT regimen (27).

**Limitations**
The prospective randomized evidence comes from relatively small unblinded trials, albeit with a
pretty low risk for methodological bias otherwise. Meta-analyses pool highly heterogeneous
patient populations, neoadjuvant regimens, mixes of histology, and study designs, but quite
often do not find evidence of statistical heterogeneity in many outcomes of interest.

**Optimal Dose for Radiation Therapy**
Patients Undergoing nCRT

- When radiation therapy is planned as part of pre-operative chemoradiotherapy, a
dose of 41.4 Gy-50.4 Gy is reasonable (Class IIA; level B-NR).

No randomized controlled trials have directly compared radiation doses in the pre-operative
setting. Dose selection has largely been based on multiple prospective trials in which pre-
operative chemoradiation was included in the study design. In the United States, 50.4Gy in 28
fractions was the dose fractionation for pre-operative radiation in CALGB 9871 (28),CALGB
80803 (10) and more recently NRG Oncology/RTOG 1010 (29). Lower radiation doses such as 40
Gy in 20 fractions or 41.4 Gy in 23 fractions have been preferred in China or Europe as best
exemplified by the NEOCRTEC5010 Trial (30) or CROSS trial (16). No significant differences in
disease-free or overall survival have been detected for lower or higher pre-operative radiation
doses in multiple meta-analyses or database analyses (31-34). Although one NCDB analysis
detected a statistically significant increase in pathologic complete response (pCR) using higher
doses compared to lower doses (35), several other such studies failed to detect significant
differences (34, 36-38). Only one single center retrospective study reported on toxicities and
found no differences in pulmonary complications between patients receiving less than 50 Gy or
at least 50 Gy (36). No other studies have compared radiation doses with respect to peri-
operative complications, long term toxicities such as cardiopulmonary effects, or patient quality
of life. Therefore, for a patient with a high likelihood of proceeding on to surgery after
chemoradiation, the dose of 41.4 Gy is reasonable. However, patients who are older and have
multiple co-morbidities may not be medically operable to proceed to esophagectomy after
chemoradiation. In these patients, 50-50.4 Gy remains an appropriate alternative to 40 Gy or 41.4 Gy as it represents a reasonable prescription dose for cases treated with either pre-operative or definitive intent. Intensity-modulated radiation therapy (IMRT) is increasingly being used for esophageal cancer compared to 3D conformal radiation techniques. There have been retrospective studies suggesting reductions in radiation dose to critical organs such as the lungs and heart, improved dose homogeneity and conformality, and improved clinical outcomes (39-41). In cases where 3D techniques can not sufficiently reduce dose to organs at risk to meet required dose objectives, IMRT is recommended.

Patients Undergoing dCRT

- A dose of 50-50.4 Gy in 25-28 fractions is recommended for patients treated with definitive intent chemoradiation (Class I; level A).

A dose of 50.4 Gy given concurrently with chemotherapy was established as a standard-of-care for definitively-treated patients in RTOG 85-01 (42, 43). Four phase III randomized controlled trials have tested the potential benefit of dose escalation for patients treated with chemoradiation alone without surgery. The INT 0123 study (Minsky et al.) found no benefit of 64.8 Gy compared to 50.4 Gy. Despite several on treatment deaths in the experimental high dose arm before 50 Gy and only one death apparently attributable to high dose radiation, statistical analysis determined futility for the higher dose arm to improve survival. Given that INT 0123 was performed with older radiation techniques and chemotherapy regimens, there
has been significant interest in dose escalation in the modern era. Recently, two additional
phase III randomized trials have been published comparing 50-50.4 Gy to 60-61.8 Gy using IMRT
(44, 45), while the abstract for a third comparing 50 Gy to 66 Gy has also been published (46).
No differences in local or local-regional progression free survival or overall survival could be
identified between the high or standard dose arms in any study. Only patients with squamous
cell carcinoma were included in Xu et al. In Hulshof et al. similar outcomes were achieved with
patients with either squamous cell carcinoma or adenocarcinoma. Toxicities were not different
between the two arms in Hulshof et al., while the rate of grade 3 pneumonitis was doubled in
Xu et al in the higher dose group. (7.5% vs. 3.1%; P = 0.03). Treatment prescriptions using
standard fractionation of 1.8-2.0 Gy per fraction over 25-28 fractions to 50-50.4 Gy is
recommended based on the above trials with lack of evidence supporting alternative
fractionation patterns and total doses exceeding 50-50.4 Gy. As is true in the above section,
IMRT is recommended when maximum target doses to organs at risk can not be achieved by 3D
conformal radiation.

### Value of Surgery

- **Surgery after CRT is recommended as the standard of care in patients with**
adeno-carcinoma. (Class I, Level C-LD)
- **Surgery is recommended in medically operable patients with SCC when a cCR is not**
achieved after CRT. (Class I, Level B-NR)
- **Either surgery or observation are reasonable in low operative risk patients with SCC**
who achieve a cCR after CRT. (Class IIA, Level B-NR)
Surgery has consistently been included in the multi-modality approach to esophageal cancer (1,2,20). However, whether all patients who undergo nCRT or nCT require surgery to achieve optimal oncologic and quality of life outcomes remains an important question.

### OS

Three prospective, phase III trials have been published comparing definitive chemoradiotherapy (dCRT) and TMT (47-49). The primary endpoint for both the Stahl and Bedenne trials was non-inferiority of 2-year overall survival (OS) for patients treated with dCRT compared to TMT.

Patients with SCC comprised 100% and approximately 90% of enrollment in the Stahl and Bedenne trials, respectively. Neither trial used positron emission tomography (PET) for staging, and endoscopic ultrasound (EUS) was not required in the Bedenne study. Importantly, patients were randomized prior to treatment initiation for the Stahl trial but only after demonstrating a response to the initial course of chemoradiation in the Bedenne trial.

Both the Stahl and Bedenne trials demonstrated statistical equivalence in OS for patients treated with dCRT compared to nCRT plus esophagectomy (47, 48). Vellapayyan et al. (2017) performed a meta-analysis of these two trials. The authors concluded that esophagectomy after chemoradiation improved local control but not OS based on moderate-quality and high-quality evidence, respectively (50).

A comprehensive meta-analysis of 35 prospective and retrospective studies comparing dCRT to TMT found no apparent benefit with surgery for OS after balancing baseline patient factors but confirmed a strong benefit in reducing local recurrence (51). There was significantly less short-term (90-day) treatment-related mortality in patients treated with dCRT compared to TMT with no heterogeneity [RR: 0.2 (95% CI 0.10-0.43) p < 0.0001]. A smaller meta-analysis of four
retrospective studies of only complete clinical responders identified a potential survival benefit
of surgery after nCRT compared to dCRT alone at 2 years, but this advantage lost statistical
significance at 5 years (52).

DFS
A third randomized trial used modern techniques such as PET, EUS, and endoscopy after CRT to
determine patients with complete clinical response (cCR), who were then randomized to
observation or esophagectomy (49). The primary endpoint was 2-year DFS, however, the trial
closed early after enrolling and randomizing only 82 and 37 patients, for observation and
surgery, respectively. No statistically significant differences could be identified between the
dCRT and TMT groups, but strong conclusions could not be drawn due to small patient
numbers.

The meta-analyses by Vellayappan (93% SCC) and Wang (96% SCC) both report superior 2-year
results for nCRT, while Voeten reported a lower overall local recurrence rate for nCRT overall
and in SCC patients only. This benefit was not maintained at 5 years per Wang et al, although
this outcome had significant heterogeneity (50-52).

The more recent matched and registry studies also find less recurrence with nCRT. However,
patients with more advanced disease (stage III) continued to have an increased risk of
recurrence regardless of nCRT in a recent retrospective study published by Jung et al (53).

Complications
Very little data exists on complications related to surgery. In the trials by Stahl and Bedenne,
post-operative complications contributed to higher short-term mortality in patients treated
with surgery (47, 48), although Vellappayan rated this as a low-quality of evidence (50). The
meta-analysis by Voeten et al found an 80% reduction in 90-day mortality in the no surgery arm (51).

Quality of Life

Only the Bedenne study reported quality of life (QoL) data; QoL scores in the dCRT patients were superior to patients in the TMT arm at 6 months but were equivalent between the treatment arms at later time-points (48).

For patients considered to be at low operative risk, the benefits of TMT likely outweigh the risks, even in the setting of cCR. Finally, for patients with SCC at higher operative risk where cCR is achieved after CRT, deferral of surgery is reasonable. Close surveillance is recommended to potentially provide the opportunity for salvage surgery if local recurrence is detected in the absence of metastatic disease. Close surveillance in the SANO trial consisted of endoscopic evaluation every three months for the first year, every four months in the second year, every six months in the third year and annual until the fifth year (54).

Limitations

The recommendations on the value of surgery after chemotherapy and radiation are based on the ability to accurately identify patients with a complete pathologic response. The current recommendations by the NCCN are FDG-PET/CT, chest/abdominal CT scan with contrast, and endoscopy with biopsy. These modalities all have limitations making close follow-up and a critical discussion with the patient paramount. The consequences of inaccurately defining a complete pathologic response is that if disease is found at a time further out from completion of chemotherapy and radiation then the patient will require a delayed or salvage esophagectomy if feasible. Salvage esophagectomy historically has been associated with an
increased risk of morbidity and mortality. Data from recent literature is mixed with some
suggesting similar outcomes and others still reporting higher rates of morbidity and mortality
(55, 56). These factors make a thoughtful discussion with the patient important prior to
following non-operative recommendations for patients with a cCR. Further data on the ability
to safely avoid surgery will be available once the SANO trial concludes (ref from above).

**Timing of Esophagectomy After nCRT**

- In patients who have recovered sufficiently and are ready for surgery, timing of
  surgery prior to 7-8 weeks after nCRT may result in a slight overall survival advantage
  with a lower risk of perioperative morbidity and mortality and is reasonable when
  possible. (Class IIA, Level B-NR)

- For patients undergoing surgery after nCRT, surgery should not be scheduled prior to 4
  weeks after completion of nCRT (Class III: Harm, Level C-LD).

The ideal timing of surgery after completion of nCRT is not well settled, and studies have used
varying definitions of short versus long duration varied by study, as well as analyzing multiple
time ranges in the same dataset. As such, several meta-analysis addressed this by defining any
surgery at less than 7-8 weeks as a short interval (57-59). The lowest cutoff for any single
individual study was 30 days, and few patients overall were operated on before then.

Using the 7-8 week cutoff, an association with increased perioperative mortality and
pneumonia was assessed. Dichotomizing into generic “shorter” and “longer” groups revealed a
potential increased risk of anastomotic leak with a longer duration to surgery.
Two of the available meta-analyses have shown an increased rate of R0 resection with a shorter duration to operation (58, 60) and one reported that pCR is superior with a longer duration (59), but overall the signal for these outcomes was neither strong nor consistent, and it is not clear that an improved pCR would translate to improved outcomes for the patient.

These results provide no strong evidence of worse overall survival with a longer delay to surgery. Inclusion of the NCDB data suggests modest but statistically significant benefits might be possible at 2 and 5 years with a longer duration to surgery, and this advantage may be more robust in patients with AC. However, not all analyses confirm this. Given the small size of the potential effects, it is not possible to make a definitive conclusion.

Operative Oncologic Outcomes

Individual studies have not shown a difference in the rate of R0 resection between a shorter and longer duration from nCRT to surgery. However, the meta-analyses by Lin et al and Tie et al with a selective inclusion criteria both report significantly worse odds of R0 with a longer duration to surgery (58, 60). Conversely, a much larger analysis by Qin et al including NCDB data and a national database study from The Netherlands report no effect on R0 resection rate in over 8,000 patients, albeit with moderate heterogeneity (59).

The meta-analysis by Qin at al report significantly better pCR in longer duration (>7-8 weeks) patients (RR: 1.13, p = 0.001) (59), while the other two meta-analyses report no difference (58, 60). However, this difference in pCR was determined with a fixed-effect model, which leads to false precision (61). The effect of timing on pCR did not show statistical significance when a random-effects model was used more appropriately (RR 1.09, p = 0.18) (Figure 2).
OS
The available meta-analyses had conflicting results in terms of OS, differing based on inclusion criteria and statistical methodology. A fourth meta-analysis published in 2020 (57) which only reported OS and DFS outcomes found a significant advantage for a shorter duration to surgery. In a subgroup analysis, this was largely due to the effect in AC patients, as the results in SCC were not significant.

Qin et al found modestly lower long-term mortality in patients who had surgery in 7-8 weeks or sooner at both two years (RR 0.94, p = 0.002) and five years (RR 0.88, p = 0.0009) with little to no heterogeneity (vs. reference >7-8 weeks). These results were dominated by the large series from the NCDB.

The meta-analysis by Lin et al reported the same benefit at two years, but this was not maintained at five years (58). Tie et al found no significant effect, regardless of whether they used the cutoff of 7-8 weeks, or 8 weeks-60 days (60). All four analyses report low to moderate heterogeneity.

DFS
The meta-analysis by Shang et al was the only one to report DFS, finding no difference between groups (57). This analysis incorporated data from only three studies.

Complications
Overall, there were two complications (anastomotic leak and pneumonia) post-esophagectomy that showed differences in outcomes based on the time from nCRT to surgery. When dichotomizing studies into “short” and “long” delays without specifying a particular cutoff, Tie et al found a significant increase in overall anastomotic complications associated with the
longer delay (60). However, with a cutoff of 7-8 weeks, three meta-analyses found that
duration to surgery did not impact the rate of anastomotic leak (58-60).

Analysis of the four studies that reported pneumonia found an increased risk of pneumonia
with a longer duration (>7-9 weeks), despite different histologies and cutoff ranges (Figure 3)
(62-65). The effect remains when excluding Tsang et al in order to set a tighter 7-8 week cutoff
(Figure 4). A funnel plot is also included (Figure 5).

30-day Mortality

Individual studies have not shown increased 30-day mortality after longer duration between
the end of nCRT and surgery. Pooled data in the Qin et al meta-analysis, however found an
increased 30-day mortality associated with a duration of longer than 7-8 weeks to surgery after
nCRT (RR 1.51, p = 0.0006) (59). Neither the meta-analysis by Tie et al nor Lin et al found such
an association (58, 60).

Limitations

The relevant studies were all retrospective in nature, as such this introduces bias. Due to this
bias, it is possible that patients who had a long delay from time of neoadjuvant therapy to
surgery were different than those who had a shorter delay. None of the studies statistically
matched the groups, although multivariable analysis was consistently used to address some of
the potential confounding variables. Overall, the studies rate moderately well to high quality on
the Newcastle-Ottawa scale, generally getting scores of 6-8.

The available meta-analyses used different inclusion criteria and frequently reported
contrasting results. While large registry data such as that from the NCDB has its limitations,
there does not seem to be a strong rationale to specifically exclude it from pooled analysis of
retrospective studies. The meta-analyses by Qin et al (when there is little to no heterogeneity) and Shang et al (for survival) appear to have the most reliable data, although the questionable use of fixed-effect models by Qin makes that paper’s conclusion on pCR difficult to parse.

Patients who had a longer delay to surgery may have had significant complications during nCRT, failure to thrive or other co-morbidities that delayed their ability to undergo surgery. These factors may have impacted perioperative morbidity and mortality and long-term survival. Those who had perioperative morbidity may have had lower long-term survival due to the impacts of the in-hospital complications and not from the impact of delay to surgery.

There were very few complications found to be associated with timing from nCRT to surgery. This may have been in part because the majority of studies do not report an effect in rates of pulmonary complications, reintubation, reoperation, or recurrent laryngeal nerve paralysis. This may be a result of lack of statistical power, as trends are generally suggestive that a longer duration could lead to more complications in larger studies.

Transhiatal Esophagectomy vs. Transthoracic Esophagectomy

- In the setting of nCRT, both TTE and THE are reasonable approaches (Class IIA, Level B-NR)

Numerous techniques have been described for resection of the esophagus and regional lymph nodes. Transthoracic esophagectomy (TTE) offers the advantage of a more extensive mediastinal lymphadenectomy compared to approaches where a chest incision is avoided, such as transhiatal esophagectomy (THE). TTE does however have an increased risk of pain, longer operative times,
the need for single lung ventilation, and the potential for more perioperative respiratory complications compared to THE. A randomized trial comparing THE versus TTE without the use of neoadjuvant therapy did not support one approach over the other (66). While THE was associated with lower morbidity than TTE with extended en bloc lymphadenectomy, a trend toward improved long-term survival at five years was noted with TTE. In the setting of neoadjuvant chemoradiotherapy (nCRT), the critical question is whether the more extensive lymphadenectomy afforded by TTE is beneficial, because the tumor spread to regional lymph nodes may be eradicated by preoperative treatment.

Lymph Node Yield versus Perioperative Morbidity/Mortality
The most extensive data regarding surgical approach in the setting of nCRT come from a retrospective cohort study of a Dutch national registry that included over 4,000 patients, most with esophageal adenocarcinoma (>85%) (67). After propensity score matching, 1,532 patients were included in the analysis. While TTE was associated with a more thorough oncologic resection with a higher number of lymph nodes (LNs) harvested (transthoracic median 19 vs. transhiatal median 14; \( p < 0.001 \)), no differences were noted between approaches in the number of positive LNs. The use of a thoracotomy, however, came at the cost of an increased rate of respiratory complications (35.5% vs. 26.1%; \( p < 0.001 \)), longer hospital (median 14 vs. 11 days; \( p < 0.001 \)) and ICU stays (median 3 vs. 1 day; \( p < 0.001 \)), more reoperations (14.8% vs. 9.3%; \( p = 0.002 \)), and higher mortality in-hospital or within 30 days of surgery (4.0% vs. 1.7%; \( p = 0.009 \)) compared to THE. Two smaller studies did not find differences in complication rates between the surgical approaches, though were underpowered (68, 69).
Six studies have compared TTE and THE following nCRT with regards to overall survival (OS) (69-74). Of these, only one small study of 58 patients found a survival difference between approaches, favoring TTE (70). This survival advantage, however, was likely explained by a selection bias with older patients having significant pulmonary or cardiac comorbidities undergoing THE. In a Dutch study of 2,698 patients, improved OS was associated with ≥15 LNs resected (HR 0.77, 95% CI: 0.68-0.86), but not with use of THE versus TTE (HR 0.89, 95% CI: 0.79-1.1) on multivariate analysis (74).

Limitations

A significant limitation in the reported studies is that they did not assess the impact of performing a TTE with a cervical anastomosis, since patients with either a cervical or thoracic anastomosis were grouped together within the TTE category. It is possible that patients with TTE with a cervical anastomosis could have less severe anastomotic leaks, which might reduce perioperative morbidity and mortality. Additionally, studies did not investigate the impact of surgeon and hospital volume on perioperative mortality after TTE relative to THE, factors found to be important in prior reports (75, 76). It is not clear how to apply our recommendation in the context of a planned TTE with a cervical anastomosis at a high-volume, tertiary hospital.

Minimally-Invasive Esophagectomy

- Minimally-invasive esophagectomy has the potential to reduce perioperative pulmonary complications and improve short-term quality of life and is reasonable to consider. (Class IIA, Level B-R)
The potential for minimally-invasive esophagectomy (MIE) to reduce mortality, major morbidity, and improve quality of life (QoL) seems intuitive. However, after nearly two decades of primarily retrospective publications comparing MIE to open esophagectomy, the data suggests the potential for modest improvement for MIE over open surgery. The major caveat in comparing MIE versus open esophagectomy is that the data come primarily from high-volume centers experienced with MIE technique and included a variety of definitions for both MIE and open surgery. Our analysis focused on the few existing RCTs, as well as meta-analyses of the extensive retrospective data. Many of these meta-analyses had a wide range of definitions for MIE including completely minimally invasive (robotic or scope based), hybrid laparoscopic/thoracotomy, and hybrid thoracoscopic/laparotomy. Open surgeries were commonly limited to TTE, although again with some variety.

Six meta-analyses reported long-term survival outcomes with up to five years of follow-up and little to no heterogeneity. The meta-analysis by Gottlieb-Vedi, et al. of over 14,000 patients reported a 15% and 18% reduced hazard of mortality at three and five years, respectively, for MIE (77). The authors noted evidence for publication bias in the three-year outcome. Siaw-Acheampong, et al found a benefit at one year across the board for totally MIE, laparoscopic hybrid, and thoracoscopic hybrid, but it was not maintained at three or five years (78). Guo et al reported better survival at two years for totally MIE, but not at one or five years (79).

Two multicenter, open label, randomized controlled trials have provided long term survival results of patients undergoing open esophagectomy (OE) versus minimally invasive esophagectomy (MIE). In the first trial, 115 patients from five European hospitals with
resectable intrathoracic esophageal or gastroesophageal junction carcinoma, were randomized between OE (n=56) and MIE (n=59) with curative intent. No differences were observed for overall survival in patients who underwent MIE compared with OE with a 3-year survival rate of 41.2% (95% CI 27.5-54.9%) in the OE group and 42.9% (95% CI 30.4-55.4%) in the MIE group, log-rank, \( P=0.633 \) (80). In the second trial, 207 patients with resectable cancer of the middle or lower third of the esophagus were randomly assigned to undergo transthoracic open esophagectomy (open procedure) or hybrid MI esophagectomy (hybrid procedure). Hybrid surgery comprised a two-field abdominal-thoracic operation with laparoscopic gastric mobilization and open right thoracotomy. At 3 years, overall survival was 67% (95% CI, 57 to 75) in the hybrid-procedure group (103 patients) as compared with 55% (95% CI 45 to 64) in the open procedure group (104 patients) (81). These differences were not statistically significant.

A single-center randomized controlled trial assigned 112 patients with resectable intrathoracic esophageal cancer to either robot-assisted minimally invasive thoracolaparoscopic esophagectomy (RAMIE) or open TTE. All patients were included in the overall survival analysis. At a median follow up of 40 months, there were no statistically significant differences in overall survival (log rank test, \( P=0.427 \)) between the two treatment arms (82).

**Operative Mortality**

Three RCTs (81-83) have compared 30-day mortality, and two RCTs (82, 83) compare in-hospital mortality between MIE and OE. Overall, mortality is very low in both groups, with a non-significant trend toward higher mortality in OE.

Several meta-analyses evaluate in-hospital and 30-day mortality comparing MIE to OE. When large pools of studies are used (84-86), MIE had a reduced 30 day and in-hospital mortality.
Although most of the individual studies were negative with non-significant trends favoring MIE, they were not powered to detect a significant difference in such a rare outcome. Although these meta-analyses did not find evidence of statistical heterogeneity, the previously mentioned issues with varying techniques and treatment complicate the analysis. Moreover, more recent meta-analyses by Akhtar et al. (87) and Lv et al (88) did not confirm these findings using stricter inclusion criteria (88).

The above data are all from studies that included hybrid techniques under MIE. The network meta-analysis by Siaw-Acheampong et al. (78) investigated open vs. totally MIE and vs. hybrid techniques separately and did not find a difference between the techniques.

Complications
Two RCTs (81, 82) have investigated differences in major complication rates (Clavien-Dindo ≥ 2) within 30 days of esophagectomy between MIE and open esophagectomy. In both studies, MIE resulted in a significantly lower incidence of intraoperative and postoperative major complications.

While the rate of major complications was lower in MIE patients, the reoperation rate was similar between the techniques as shown by two RCTs (83, 89). Regarding recurrent laryngeal nerve injuries, three RCTs have shown higher vocal cord paralysis rates in patients randomized to open esophagectomy, but the difference was only statistically significant in the study by Biere et al. (83). Interestingly, there was no difference between the groups after one year of follow-up, as reported by Maas et al. (80). A large matched cohort study by Takeuchi et al. (90) using a national database in Japan found an increased rate of nerve injury in the MIE group (MIE 361 (10.3%) vs. Open 285 (8.1%), p=0.002), although this may be related to center-specific
lymph node dissection approaches. The meta-analysis by Xiong et al., including 3 RCTs (80, 83, 91) and 2 prospective studies (92, 93) with 488 patients in total supports this finding [OR: 0.300 (95% CI, 0.101-0.864) \( p = 0.026, I^2 = 2.2\% \)]. However, several other meta-analyses, often with larger pools of patients report no difference (79, 84, 88, 94-96).

While the mortality data is ambiguous and non-conclusive, data suggesting that MIE results in fewer pulmonary complications are comparatively more robust and generally lacking heterogeneity.

The best data in favor of this comes from the meta-analysis by Lv et al. (88), which performed separate subgroup analyses for RCTs and nonrandomized studies. In the RCT data, MIE was associated with significantly reduced pulmonary complications [9.8% vs. 28.4%, RR: 0.34 (95% CI, 0.21-0.53) \( p < 0.00001, I^2 = 0\% \)]. The subsequent randomized trial by Mariette et al. (81) also added to the evidence in support of MIE. The Siaw-Acheampong meta-analysis (78) by MIE technique suggests that totally MIE and thoracoscopic hybrid lead to this reported benefit, although perhaps not in laparoscopic hybrid procedures. Further, the reduced rate of pulmonary complications is consistent across nearly all studies.

There is no substantial evidence that MIE affects the rate of anastomotic leakage. The RCT subgroup analysis by Lv et al. (88) found no difference for MIE, although the number of included patients (n=363) was low. Regardless, no individual study or meta-analysis suggests a clinically relevant difference between MIE and open esophagectomy.

No data suggests that MIE reduces the rate of reoperation or renal failure.
Quality of Life

The 2017 meta-analysis by Kauppila et al pooled together data from 1,806 patients from studies that compared QOL outcomes in either totally MIE or hybrid MIE vs. open esophagectomy (transhiatal approach excluded). The investigators found that minimally invasive surgery is generally followed by better postoperative outcomes regarding global quality of life, physical function, fatigue and pain for up to 3 months after surgery but these differences fail to persist at 6 or 12 months (97).

Additionally, two more recent retrospective studies have investigated the impact MIE has on QoL. In one, Wang et al compared a minimally-invasive Ivor-Lewis approach with the open Sweet approach, which is common in China. They reported statistically and clinically significant benefits at three, six and 12 months postoperatively for MIE, but this did not persist at 24 months (98).

In addition, a one-year follow-up analysis of the quality of life was conducted for patients participating in the randomized trial in which minimally invasive esophagectomy (59 patients) was compared with open esophagectomy (56 patients). A response compliance of 82% by patients was obtained. There were significantly better quality of life scores in certain domains after 1 year follow up for the MIE group as compared to the OE group. These differences were present in three domains: physical activity [SF36: 50 (6;48-53) vs 0.45 (9;42-48) p=0.003]; global health [C30:79 (10;76-83) vs 67 (21; 60-75) p=0.004] and pain [OES18:6(9;2-8) versus 16 (16;10-22)p=0.001]. These are clinically meaningful differences. However, mental component scores and the degree of improvement were not superior for MIE. The investigators concluded that MIE is associated with a better mid-term one-year quality of life compared to OE (80).
Limitations

Only three small randomized trials (112 to 207 patients) compare the long term survival of esophageal cancer patients by OE or MIE. The statistical power of these three trials is limited to 3-year follow-up. Only one of these studies provided QoL analyses and this was limited to one year follow-up.

Dozens of studies, mostly single-center retrospective studies have been performed in the last two decades that compare MIE vs. OE, although conclusions are complicated by varying patient groups, pre-and postoperative treatment protocols, and surgical techniques. Likewise, the multiple meta-analyses available vary greatly in inclusion criteria, adding to the complexity of reaching any clear conclusions.

Additionally, several of the meta-analyses used fixed-effects models with heterogeneity as high as 50%, and made other questionable methodological choices that undermined the strength of their conclusions. Further, combining RCTs and cohort studies into a single pooled estimate, particularly without separate subgroup analysis, is questionable.

Adjuvant Systemic Therapy

- Adjuvant nivolumab is recommended in patients with residual disease after neoadjuvant chemoradiotherapy and no contraindications (Class I, Level of Evidence B-R).

Conclusions regarding the efficacy of adjuvant chemotherapy in patients with residual disease are complicated by varying patient populations, histology, chemotherapy regimens, etc.

Although patients are now largely receiving neoadjuvant chemoradiotherapy, there are only a
few studies investigating adjuvant chemotherapy in patients who have received preoperative therapy (99-102). For squamous cell carcinoma patients who have undergone neoadjuvant therapy, no compelling data yet exists that adjuvant chemotherapy improves either OS or DFS.

For SCC patients who did not receive neoadjuvant therapy, there is limited data suggesting a potential clinical benefit for adjuvant chemotherapy in both OS and DFS, although these results are somewhat mixed and mostly from small, single-center studies.

While several retrospective analyses that combine AC and SCC patients suggest some degree of survival benefit, the questionable comparability of the two histologies, lack of standardization in preoperative therapy and adjuvant treatment, and lack of high-quality prospective evidence prevent making any conclusions about this approach.

Adjuvant nivolumab represents a hopeful treatment option for patients with residual disease after nCRT, and interim analysis after a median of 24.4 months of follow-up suggests the benefits of treatment may appear in as early as 6 months in terms of improved DFS in patients. This survival difference was not impacted by PD-L1 expression.

**Squamous Cell Carcinoma**

**OS**

It is now a matter of historical interest only but the earliest randomized studies to evaluate adjuvant chemotherapy following surgery for esophageal SCC were published by Ando and colleagues from the Japan Clinical Oncology Group (JCOG) (5, 103), and were designed to compare surgery alone vs. adjuvant cisplatin/vindesine or 5-FU. Results were mixed and did not clearly define a role for adjuvant use of cytotoxic agents, with the JCOG9204 study indicating a potential benefit in 5-year DFS for patients with lymph node involvement (52% vs. 38%, p =
0.041), and the subsequent JCOG9907 study showing an inferior 5-year OS for postoperative 5-
FU/cisplatin vs. preoperative treatment (43% vs. 55%, p = 0.04).

Aside from these randomized studies, other reports of adjuvant chemotherapy in SCC involve
retrospective series. Retrospective analyses by Kim et al. (104) and Sohda et al. (105) suggest a
potential survival benefit for adjuvant chemotherapy vs. surgery alone, although the most
contemporary study by Zheng et al. failed to show a benefit in 1, 3, or 5-year OS (106). Patients
who received adjuvant chemotherapy tended to be younger and had more locally-advanced
disease.

A recent retrospective analysis by Matsuura et al. evaluated 113 patients who received
neoadjuvant chemotherapy followed by radical surgery and had three or more pathologic
positive lymph nodes and were either treated with adjuvant docetaxel + paclitaxel + S-1, or
docetaxel + S-1 (n=40) or no further therapy (n=73). Both 2-year and 5-year OS were equivalent
between groups (99).

Two meta-analyses including both RCT and observational data investigated use of adjuvant
therapy regardless of preoperative treatment. Zhang et al. performed separate analyses for RCT
and nonrandomized data and found no advantage for adjuvant therapy for 3-year OS from the
RCTs only [RR: 0.95 (95% CI, 0.78-1.15), p = 0.59 I² = 0%], although only 309 patients were
included. Of note, the three randomized studies in this meta-analysis were the JCOG study with
cisplatin/vindesine, the JCOG9204 study and a French study from the 1990s by Pouliquen et al.
that included patients who had undergone incomplete resections and who had metastatic
disease. Likewise, no advantage was found among 1,392 nonrandomized patients (107).
However, a more recent meta-analysis by Zhao et al found an overall advantage for adjuvant therapy in a combined analysis of RCT and nonrandomized studies [Adjuvant HR: 0.78 (95% CI, 0.66–0.91) p = 0.002, I² = 0%) (108). This discrepancy is likely a result of the inclusion of four trials published after the previous meta-analysis. In addition, the authors in the earlier study included two publications in Chinese that were not part of the more recent analysis.

DFS
The two meta-analyses present a mixed picture on DFS, as well. In RCTs only, DFS was superior at 1 year for the adjuvant therapy group, but not at 3 years, reflecting the negative results of all three included RCTs. The combined analysis by Zhao et al, however, found a clear benefit for adjuvant therapy [adjuvant HR: 0.72 (95% CI, 0.60–0.86) p < 0.001, I² = 0%) (107, 108).

In the study by Matsuura, adjuvant chemotherapy did not improve DFS after 2 years (30.0% vs. 28.8%, p = 0.47). However, after evaluating 11 variables for inclusion in multivariable analysis and selecting number of positive lymph nodes, therapeutic effect grade, and adjuvant chemotherapy for the model, adjuvant chemotherapy was found to improve DFS [HR: 0.53 (95% CI, 0.32–0.88; p = 0.01)] (99).

Zhang et al also reported superior 3-year and 5-year DFS [HR: 0.448 (95% CI, 0.260–0.773; p = 0.004)] (107). Sohda et al (5-year 59.0% vs. 43.1%, p = 0.002) (105) and Kim et al [MVA only - HR: 0.56 (0.35-0.90) p = 0.018] (104) also reported similar findings.

As with OS, Zheng et al found a trend toward increased DFS throughout the duration of follow-up, but it was not statistically significant (106).
Locoregional Control

The study by Kim et al of 130 SCC patients who did not undergo neoadjuvant therapy was the only study to report on locoregional control-related outcomes. In this study, adjuvant chemotherapy patients showed a non-significant trend toward reduced 5-year locoregional recurrence [HR: 0.58 (95% CI, 0.31-1.08; p = 0.068)] (104).

Studies with AC and SCC Patients

OS

For patients who had received neoadjuvant therapy, large registry studies by Samson et al, Burt et al, and Nevala-Plagemann et al are all suggestive of a survival benefit for adjuvant chemotherapy, whether in AC patients or overall, and particularly in patients with pathologically-confirmed nodal involvement (100-102). In patients who received induction CT, Samson et al matched 214 patients with nodal involvement who received adjuvant CT with 214 patients who did not receive adjuvant chemotherapy and found a survival benefit (median 33.1 mo vs. 26.2 mo; p = 0.03). When performing multivariable analysis on a full cohort of 3,100 patients who predominantly did not receive adjuvant treatment, adjuvant chemotherapy was associated with a reduced mortality hazard [0.71 (95% CI, 0.59-0.87; p = 0.001)] (102).

Burt et al reported a trend of improved survival in patients with residual disease after nCRT [HR: 0.87 (95% CI, 0.75-1.00)] which increased to significance in residual nodal disease both overall [HR 0.70 (0.57-0.85)] and in AC patients alone [HR: 0.69 (0.57-0.85)] (101).

The study by Nevala-Plagemann et al was able to show a benefit after applying a multivariable analysis to address confounding variables among 6,785 patients who received nCRT [p = 0.38 when comparing median survival overall, but HR: 0.77 (0.66 — 0.89) p < 0.001 in MVA] (100).
DFS

Adjuvant immunotherapy has been evaluated in the recent CheckMate 577 Phase III trial of 794 patients randomized 2:1 to receive either adjuvant nivolumab or placebo after neoadjuvant chemoradiation and surgery reported a significantly increased DFS for nivolumab after a median follow-up of 24.4 months [22.4 mo vs. 11.0 mo, HR: 0.69 (96.4% CI, 0.56–0.86) p <0.001] (109). Post hoc analysis showed that this benefit was maintained regardless of histological type. Per the Kaplan-Meier estimate, the DFS curves start to separate at 6 months between treatment groups and remain separated.

The retrospective studies testing adjuvant cytotoxic agents in combined AC and SCC patients did not report DFS.

Limitations

While the sample size issues of SCC-only studies are largely rectified in the combined histology trials, conclusions are softened due to the uncertain impact that histological type has on treatment efficacy. As with the previous set of studies, selection of the patients who received adjuvant therapy may have been a result of bias. This is particularly the case in the large registry studies, where patients who did not receive adjuvant therapy may have had complications after surgery and may not have been well enough to have it.

The post hoc analyses by subtype in the randomized trial of nivolumab should be interpreted with caution, however, the main findings are the result of a methodologically-sound trial.
References


41


Figure 1

PRISMA Flow Diagram:
STS/ASTRO Updated Clinical Practice Guidelines on Multimodality Therapy for Esophageal Cancer

Records identified through database searching
- Medline (n = 1187)
- Embase (n = 1670)

Additional records identified through other sources (n = 62)

Records after duplicates removed (n = 2133 + 62)

Records screened (n = 2195)

Records excluded (n = 2017)

Full-text articles assessed for eligibility (n = 178)

Full-text articles excluded, with reasons (n = 62)

Studies included in qualitative synthesis (n = 112)
Figure 2: Reanalysis of Qin et al as a random effects model showing no significant difference in pCR related to a 7-8 week cutoff for surgery after CRT

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>&gt; 7-8 Weeks</th>
<th>≤ 7-8 Weeks</th>
<th>Weight</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chiu et al 2013</td>
<td>27</td>
<td>136</td>
<td>55</td>
<td>55.0% 0.75 [0.48, 1.16]</td>
</tr>
<tr>
<td>Harsley et al 2014</td>
<td>39</td>
<td>154</td>
<td>60</td>
<td>61.0% 1.11 [0.73, 1.69]</td>
</tr>
<tr>
<td>Kathiravetpillai et al 2016</td>
<td>31</td>
<td>121</td>
<td>18</td>
<td>18.6% 0.62 [0.35, 1.09]</td>
</tr>
<tr>
<td>Kim et al 2012</td>
<td>150</td>
<td>356</td>
<td>116</td>
<td>116.0% 1.05 [0.47, 1.16]</td>
</tr>
<tr>
<td>Lee et al 2016</td>
<td>2012</td>
<td>1672</td>
<td>800</td>
<td>800.0% 1.62 [1.11, 2.37]</td>
</tr>
<tr>
<td>Muller et al 2015</td>
<td>11</td>
<td>44</td>
<td>14</td>
<td>14.0% 0.60 [0.34, 1.05]</td>
</tr>
<tr>
<td>Rued et al 2010</td>
<td>16</td>
<td>60</td>
<td>20</td>
<td>20.0% 0.60 [0.34, 1.05]</td>
</tr>
<tr>
<td>Shaikh et al 2015</td>
<td>14</td>
<td>44</td>
<td>7</td>
<td>7.0% 0.69 [0.48, 0.99]</td>
</tr>
<tr>
<td>Tessler et al 2014</td>
<td>43</td>
<td>135</td>
<td>30</td>
<td>30.0% 1.36 [0.87, 2.13]</td>
</tr>
<tr>
<td>Tsang et al 2017</td>
<td>13</td>
<td>53</td>
<td>19</td>
<td>19.0% 0.70 [0.36, 1.36]</td>
</tr>
<tr>
<td>van der Weij et al 2018</td>
<td>813</td>
<td>2067</td>
<td>288</td>
<td>288.0% 1.68 [0.62, 1.18]</td>
</tr>
<tr>
<td>Wang et al 2015</td>
<td>90</td>
<td>385</td>
<td>159</td>
<td>159.0% 1.25 [0.64, 2.42]</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>5731</td>
<td>5059</td>
<td>100.0%</td>
<td>1.09 [0.96, 1.23]</td>
</tr>
<tr>
<td>Total events</td>
<td>1389</td>
<td>13014</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: Tau² = 0.02; Chi² = 15.62, df = 11 (P = 0.06); P = 42%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 1.34 (P = 0.19)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3: Forest plot for pneumonia dichotomized into “shorter” and “longer duration”

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Shorter Duration</th>
<th>Longer Duration</th>
<th>Weight</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Furukawa et al 2018</td>
<td>18</td>
<td>108</td>
<td>10</td>
<td>10.0% 0.92 [0.75, 1.13]</td>
</tr>
<tr>
<td>Kathiravetpillai et al 2016</td>
<td>18</td>
<td>85</td>
<td>42</td>
<td>42.0% 0.73 [0.45, 1.18]</td>
</tr>
<tr>
<td>Kheir et al 2020</td>
<td>33</td>
<td>344</td>
<td>39</td>
<td>39.0% 0.74 [0.48, 1.18]</td>
</tr>
<tr>
<td>Tsang et al 2017</td>
<td>5</td>
<td>54</td>
<td>3</td>
<td>3.0% 0.84 [0.51, 1.38]</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>559</td>
<td>505</td>
<td>100.0%</td>
<td>0.68 [0.49, 0.96]</td>
</tr>
<tr>
<td>Total events</td>
<td>73</td>
<td>94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: Tau² = 0.92; Chi² = 3.74, df = 3 (P = 0.26); P = 20%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 2.21 (P = 0.04)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4: Forest plot for pneumonia with a 7-8 week cutoff (Tsang et al excluded)

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>≤ 7-8 Weeks</th>
<th>&gt; 7-8 Weeks</th>
<th>Weight</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Furukawa et al 2018</td>
<td>16</td>
<td>106</td>
<td>10</td>
<td>10.0% 0.92 [0.75, 1.13]</td>
</tr>
<tr>
<td>Kathiravetpillai et al 2016</td>
<td>16</td>
<td>85</td>
<td>42</td>
<td>42.0% 0.73 [0.45, 1.18]</td>
</tr>
<tr>
<td>Kheir et al 2020</td>
<td>33</td>
<td>344</td>
<td>39</td>
<td>39.0% 0.74 [0.48, 1.18]</td>
</tr>
<tr>
<td>Tsang et al 2017</td>
<td>5</td>
<td>54</td>
<td>3</td>
<td>3.0% 0.84 [0.51, 1.38]</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>515</td>
<td>452</td>
<td>100.0%</td>
<td>0.66 [0.49, 0.98]</td>
</tr>
<tr>
<td>Total events</td>
<td>65</td>
<td>81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: Tau² = 0.00; Chi² = 2.13, df = 2 (P = 0.34); P = 8%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 2.69 (P = 0.007)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 5: Funnel plot for pneumonia