Response rate and diagnostic accuracy of early PET-CT during neo-adjuvant therapies in oesophageal adenocarcinoma: a systematic review and meta-analysis

Abstract

Purpose

Only 25% of oesophageal adenocarcinoma (OAC) patients have a pathological response to neo-adjuvant therapy (NAT) before oesophagectomy. Early response assessment using PET imaging may help guide management of these patients. We performed a systematic review and meta-analysis to synthesise the evidence detailing response rate and diagnostic accuracy of early PET-CT assessment.

Methods

We systematically searched several databases including MEDLINE and Embase. Studies with mixed cohorts of histology, tumour location, and a repeat PET-CT assessment after more than one cycle of NAT were excluded. Reference standard was pathological response, defined by Becker or Mandard classifications. Primary outcome was metabolic response rate after one cycle of NAT defined by a reduction in maximum standardised uptake value (SUVmax) of 35%. Secondary outcome was diagnostic accuracy of treatment response prediction, defined as the sensitivity and specificity of early PET-CT using this threshold. Quality of evidence was also assessed. Random-effects meta-analysis pooled response rates and diagnostic accuracy. This study was registered with PROSPERO (CRD42019147034).

Results

Overall, 1341 articles were screened, and six studies were eligible for analysis. These studies reported data for 518 patients (aged 27-78 years; 452 [87.3%] were male) between 2005-2020. Pooled sensitivity of early metabolic response to predict pathological response was 77.2% (95%CI 53.2%-100%). Significant heterogeneity existed between studies (I^2=80.6% (95%CI 38.9%-93.8%), p=0.006). Pooled specificity was 75.0% (95%CI 68.2%-82.5%), however no significant heterogeneity between studies existed (I^2=0.0% (95%CI 0.0%-67.4%), p=0.73).

Conclusion

High-quality evidence is lacking, and few studies met the inclusion criteria of this systematic review. The sensitivity of PET using a SUVmax reduction threshold of 35% was suboptimal and varied widely. However, specificity was consistent across studies with a pooled value of 75.0%, suggesting early PET assessment is a better predictor of treatment resistance than of pathological response. Further research is required to define optimal PET-guided treatment decisions in OAC.

Keywords: oesophageal neoplasm; adenocarcinoma; positron emission tomography; treatment response; neo-adjuvant therapy; diagnostic accuracy

Review criteria: how did you gather, select and analyze the information you considered in your review?

- A comprehensive search strategy using text words and Medical Subject Heading (MeSH) was designed.
- This strategy was conducted in MEDLINE, Embase, Cochrane Library, Cumulative Index of Nursing and Allied Health Literature (CINAHL), Scopus, Web of Science, International Clinical Trials Registry Platform (ICTRP) Search Portal and ClinicalTrials.gov.
- The search was limited to articles published in English from 2005 onwards. Study filters for randomised control trials and observational study types were applied.

Message for the clinic: what is the ‘take-home’ message for the clinician?

- Specificity of early PET response was consistent across studies with a pooled value of 75.0%, suggesting early PET assessment is a better predictor of treatment resistance than of pathological response.
- The sensitivity of PET using a SUVmax reduction threshold of 35% was found to be suboptimal (77.2%) and varied widely.
- The pooled early response rate defined by PET imaging was estimated to be 44.7%, but high-quality evidence is lacking.

Introduction

The incidence of oesophageal cancer is increasing worldwide, with more than 450,000 patients diagnosed each year [1]. The prognosis of oesophageal cancer is poor, especially in locally advanced and metastatic disease [2]. Only 20-30% of patients are suitable for surgical management [3], the majority of these receive neo-adjuvant therapy which aims to reduce the volume of disease prior to resection to improve survival.

A number of neo-adjuvant and perioperative chemotherapy trials have shown an overall survival benefit over surgery alone [4]; these include the neo-adjuvant chemotherapy Medical Research Council OE02 and peri-operative Adjuvant Gastric Infusional Chemotherapy (MAGIC) and ACCORD-07 trials [2, 5–7]. Similarly, the Chemoradiotherapy for Esophageal Cancer followed by Surgery Study (CROSS) trial showed improved survival benefit for neo-adjuvant chemoradiotherapy over surgery alone [8, 9]. More recently, perioperative chemotherapy with FLOT chemotherapy (5-fluorouracil (5-FU), Leucovorin, Oxaliplatin and Docetaxel) is established as the new standard-of-care for patients with operable oesophago-gastric cancer [10].

Positron-emission tomography combined with computed tomography (PET-CT) is now a standard investigation in the routine staging pathway of oesophageal cancer [11]. The main advantage of PET-CT is its greater sensitivity for undetected metastases on CT, which changes management in a significant number of patients [12], thus preventing them from undergoing major surgical intervention for little potential benefit. Focus has been placed on the role of PET-CT to define treatment response, particularly at an early timepoint during neo-adjuvant treatment [13].

The decision to alter neo-adjuvant therapy based on an early assessment may differentiate metabolic responders from non-responders; the latter group could potentially be offered an alternative neo-adjuvant therapy or simply omit the remaining cycles and proceed straight to surgery, thereby reducing the exposure to potential side-effects of chemotherapy with or without radiotherapy and reducing the chance of progression during the interval before

surgery. Several recent large scale interventional trials have used a PET-directed therapy approach [14, 15].

The primary objective of the study was to systematically review the available literature reporting early metabolic response rate, defined by fluorodeoxyglucose (FDG)-PET after one cycle of neoadjuvant therapy, in patients with oesophageal adenocarcinoma. The secondary objectives were to review the literature reporting diagnostic accuracy of early PET-CT, to perform sub-group analyses of co-variates to investigate sources of heterogeneity between studies, to assess for publication bias, and to evaluate the quality of included studies.

Materials and Methods

Search strategy and selection criteria

A comprehensive search strategy using text words and Medical Subject Heading (MeSH) was designed (supplementary material). This strategy was translated and conducted in MEDLINE, Embase, Cochrane Library, Cumulative Index of Nursing and Allied Health Literature (CINAHL), Scopus, Web of Science, International Clinical Trials Registry Platform (ICTRP) Search Portal and ClinicalTrials.gov. The search was limited to articles published in English from 2005 onwards, because 3D PET became integrated into most PET-CT scanners providing more standardisation in SUVmax from this timepoint onwards [13]. Study filters for randomised control trials and observational study types were applied. Reference lists of all eligible studies were checked and underwent citation tracking for additional eligible studies. After the abstract screening was performed, the search was repeated to capture any recent articles that may have been missed. The review included randomised control trials, observational cohort, cross-sectional and case-control studies reporting original response rate data in adult human participants.

Participants were patients with biopsy-proven oesophageal, or gastro-oesophageal junction adenocarcinoma (confirmed by histopathology), who were treated with neo-adjuvant chemotherapy or chemo-radiotherapy prior to surgical resection and had an interim PET-CT

examination (after one-cycle of neo-adjuvant therapy, around day 14). Patients with recurrent oesophageal adenocarcinoma were excluded. Studies of patients with histology other than adenocarcinoma, those without an early PET-CT examination, and studies with mixed patient cohorts were also excluded.

The radioisotope 18-Fluorine (^{18}F) FDG must have been used for the PET-CT examination. The maximum standardised uptake value (SUVmax) must have been measured by an appropriately trained and experienced professional. SUVmax was defined as the voxel with the highest SUV value in a defined region of interest [16]. The threshold for metabolic response classification, defined in terms of the percentage reduction in SUVmax, must have been stated. The reference standard was the pathological tumour regression grade (TRG), defined by either validated pathological classification systems Becker [17] or Mandard [18] (supplementary material). The number of patients that progressed during neo-adjuvant therapy and were no longer suitable for surgery will also be recorded.

The a priori primary outcome of the systematic review was the early metabolic response rate, a PET response defined as a 35% SUVmax reduction from the staging PET-CT. Diagnostic accuracy was defined by calculating sensitivity and specificity, against the reference standard of pathological TRG. For diagnostic accuracy, pathological responders were classified as Becker TRG 1 or Mandard TRG 1-2, and pathological non-responders as Becker TRG 2-3 or Mandard TRG 3-5.

This study adheres to the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) guidelines [19] and the protocol is registered with PROSPERO, registration number CRD42019147034 [20].

Data Analysis

Following the systematic search, all titles and abstracts were screened by two independent authors (KF and JJ) against the defined eligibility criteria. Full text articles were obtained for all studies that meet the inclusion criteria. In cases of disagreement following screening of titles and abstracts, a third author was asked to review and decide upon the suitability of

the study (ES). Reasons for exclusion at this stage were recorded. The results of the systematic search were shared between the reviewers using an output file imported into Mendeley Desktop (version 1.19.4). The full-text articles were included in the output file. Duplicate items were identified, and one duplicate copy was deleted.

Relevant data were extracted from the final set of eligible articles and inputted into a spreadsheet designed specifically for this review (Microsoft Excel 365). Data items extracted included patient characteristics: number of patients in study, age, gender, tumour location, neo-adjuvant regimen, pathological response rate at surgery, length of survival; study characteristics: primary author, publication year, study dates, country of study, study design, number of centres, length of time between interim PET and surgery, conclusions of study; and PET-CT characteristics: timing of interim PET-CT (days after treatment inception), type of scanner and acquisition (including PET reconstruction method), length of fasting before injection, time between FDG injection and PET, PET quantification method, interpreter(s), threshold criteria for defining response, proportion of patients with early response to neo-adjuvant therapy. Where diagnostic accuracy statistics were not explicitly quoted, a 2x2 table was constructed using the published data to derive them for meta-analysis.

The methodological quality of eligible studies was assessed using the Quality Assessment of Studies of Diagnostic Accuracy Included in Systematic Reviews (QUADAS-2) criteria [21], which comprises four domains, each assessing the risk of bias and clinical application. Perceived quality will be graded low, high or unclear risk.

Meta-analysis was performed with a random-effects model [22], using the meta package for R version 3.6.1 [23]. Early metabolic response rates, sensitivity and specificity were pooled across studies with corresponding 95% CIs. A weighted symmetric summary ROC (sROC) curve with a 95% CI was computed using diagnostic accuracy results of included studies [24]. Heterogeneity was assessed between specific-study estimates using the inconsistency index (I^2 statistic [25]). Sources of heterogeneity between studies were investigated using subgroup analyses by stratifying original co-variates according to methodological quality (QUADAS-2 score), sample size, PET injection time, neo-adjuvant therapy regimen and

histopathological response. Publication bias and small-study effects were assessed using funnel plots. The overall strength of the evidence was rated using Grading of Recommendations Assessment, Development, and Evaluation (GRADE) methodology [26].

Results

The initial search identified 1338 studies, of which 13 were duplicates. Three additional studies were identified through other sources. The titles and abstracts of 1328 studies were screened by the two independent reviewers.

After screening, 1296 records were excluded for being irrelevant to this study, wrong study type or wrong patient group, leaving 32 full-text articles for review.

Both reviewers identified 25 of the 32 full-text articles (78.1%) and the remaining seven were included after agreement by the third reviewer. Of the 32 full-text articles, 26 were excluded (agreed by both reviewers) leaving six articles for inclusion. (Fig. 1) Important characteristics of the included studies [14, 27–31] are detailed in Table 1.

The 26 excluded articles were either review articles [13, 32–38], had inadequate or mixed patient cohorts [39–43], a patient cohort previously reported [44–48], inadequate imaging for example not an early PET-CT [49–52], inadequate reference standard [53, 54] and inadequate treatment for this study [55, 56]. No articles were excluded for using an alternative SUVmax reduction threshold to 35%.

In total, 518 patients were included in evidence synthesis. Of those, 450 proceeded to surgical resection allowing the determination of diagnostic accuracy. The median age range was 58-63 (minimum 27, maximum 78) and the majority were males (73.8%-93.3%). All patients had gastro-oesophageal adenocarcinoma. Three studies [27, 28, 30] used neo-adjuvant chemotherapy alone prior to surgery, one [29] used neo-adjuvant chemoradiotherapy, and two [14, 31] used initial neo-adjuvant chemotherapy followed by

radiotherapy in metabolic non-responders. (Table 2) In total, 20/519 (3.9%) of patients progressed during neo-adjuvant therapy and were no longer eligible for surgery.

All studies used fluorodeoxyglucose (FDG). Preparation was inconsistently described, mainly featuring the length of fasting prior to examination (at least 360 minutes in all studies that reported this). The length of time between injection and scan ranged from 40-75 minutes. The FDG dose was variably reported from dose per kilogram of bodyweight, overall dose range and mean dose between patients. PET acquisition and reconstruction methods were scantily reported, often briefly summarised. The most common method of SUVmax measurement was a semi-quantitative region of interest drawn by the reader of the study. The number of reporting radiologists was not reported, nor the experience of the readers.

The time between the baseline PET-CT and the early PET-CT ranged between 9-21 days, but all studies pre-specified the aim to perform the early PET-CT around day 14, one of the key inclusion criteria of this systematic review. The threshold defining a metabolic response in all studies was pre-specified as a reduction in SUVmax of 35%. In total, 205/450 (45.6%) patients who proceeded to surgery were defined as metabolic responders. A forest plot shows the pooled early response rates to neo-adjuvant therapies using a reduction threshold in SUVmax of 35%. (Fig. 2)

Early metabolic response rate

The pooled early metabolic response rate across the included studies was 44.7% (95% CI 0.37-0.52). There was significant heterogeneity between studies (I^2 = 64.4%, 95% CI 14.2%-85.2%, p=0.02).

Diagnostic Accuracy

Three studies were used to meta-analyse the diagnostic accuracy of early PET response [28–30]. These studies used consistent treatment regimens for all patients and classified pathological TRG using either Becker 1 [28, 30], or Mandard 1 or 2 [29] classifications. In total, 80/197 (40.6%) patients had a pathological response using the respective pathological

classification criteria. Table 3 shows the sensitivity and specificity of included early PET response studies to predict pathological response.

The pooled sensitivity of early metabolic response to predict pathological response was 77.2% (95% CI 53.2%-100%). A wide variation in sensitivity results was observed, from 43.8% [29] to 100% [28]. As such, there was significant heterogeneity between studies (I^2 = 80.6% (95% CI 38.9%-93.8%), p=0.006). (Fig. 3A) The pooled specificity was 75.0% (95% CI 68.2%-82.5%). Unlike sensitivity, no significant heterogeneity between studies existed (I^2 = 0.0% (95% CI 0.0%-67.4%), p=0.73). (Fig. 3B)

A summary ROC curve demonstrates the diagnostic accuracy of included studies using a threshold reduction of 35% in SUVmax during neo-adjuvant therapy to predict pathological response (supplementary material).

Meta-regression

Meta-regression analysis was performed for five variables (study sample size, neo-adjuvant treatment regimen, TRG classification system, image acquisition parameters and risk of study bias); which are potential confounders. None were significantly associated with the primary outcome (proportion of early responders to neo-adjuvant therapy).

Study sample size was not associated with the primary outcome (p=0.4583). The largest study was Barbour et al [14] with 107 patients resected, the smallest was Malik et al [29] with 37 patients resected. There was significant heterogeneity between studies (I^2 69.1%, p=0.012). No association between the primary outcome and length of interval from FDG injection to image acquisition was demonstrated (p=0.3616). There was significant variation between studies (I^2 66.6%, p=0.018). The neo-adjuvant regimen differed between studies (Table 2) but no association with the primary outcome was demonstrated (p=0.4638). There was significant variation between studies (I^2 72.2%, p=0.013). The classification system used to define pathological TRG was not associated with the primary outcome (p=0.87). Significant variation existed between studies (I^2 71.5%, p=0.007). There was no association

with perceived risk of study bias using the QUADAS-2 grading system (p=0.72). Again, there was significant variation between studies (I^2 70.6%, p=0.009).

Publication bias

A funnel plot showed no clear indication of small-study effects, including publication bias (supplementary material). Egger's test did not demonstrate any significant evidence of publication bias (p=0.60).

Quality of Evidence

Despite the strict inclusion criteria for this systematic review, there were some concerns regarding the quality of included studies according to the QUADAS-2 quality assessment tool. (Table 4) These mainly concerned patient selection and the consistency of treatment that patients received in the interval between the early PET-CT and pathological reference standard evaluation. Overall, the confidence in the results of the quantitative synthesis was summarised as moderate, according to the GRADE methodology.

Discussion

Optimising patient selection for radical curative treatment is imperative in oesophageal adenocarcinoma. The prognosis is poor with 5-year overall survival just 15%, and the majority of patients (70%) present with advanced disease [57]. Patients fit for surgery, who are deemed to have potentially resectable disease, usually have neoadjuvant treatment prior to oesophagectomy [5, 8]. However, only a minority (23%) experience a complete pathological response after neo-adjuvant chemoradiotherapy [8], and even fewer (15%) demonstrate significant tumour regression from neoadjuvant chemotherapy [58].

The systematic review found that 44.7% of patients with oesophageal adenocarcinoma had an early metabolic response on PET-CT performed around 14 days after neo-adjuvant treatment initiation, however there was substantial heterogeneity in response rates

between included studies. This review also meta-analysed diagnostic accuracy of early PET response using the 35% SUVmax reduction threshold. We found the pooled sensitivity and specificity of early metabolic response to predict pathological response at this threshold was 77.2% and 75.0%, respectively, but only three studies were eligible for diagnostic accuracy meta-analysis. Again, there was substantial heterogeneity between included studies for sensitivity, but specificity results were relatively consistent. These data indicate that early PET response is a far more consistent negative predictor of treatment resistance than a positive predictor of response, but that diagnostic accuracy currently is suboptimal at the 35% threshold level. The consistent specificity is re-assuring for future research aiming to identify non-responders to neo-adjuvant treatment in whom, alternative management strategies could be sought. Those patients who are found to demonstrate an early response should continue with that regimen until their oesophagectomy.

Overall, studies investigating early PET response assessment to neo-adjuvant treatments have reported promising data [13], but often suffer from inclusion of heterogeneous patient cohorts such as mixed populations of adenocarcinoma and squamous cell carcinoma, the two most common histological subtypes of oesophageal cancer [1]. This systematic review and meta-analysis used strict inclusion criteria in attempt to provide definitive evidence of early treatment response and diagnostic accuracy in patients with oesophageal adenocarcinoma. This systematic review confirmed that many studies were ineligible for inclusion because they contained mixed patient cohorts of histological cell type, tumour location, and treatment strategies, different definitions of pathological response, and differing timings of early PET assessment. In particular, studies with mixed histological cell types were excluded because there are reported differences in FDG-avidity between oesophageal adenocarcinoma and SCC [12, 59], therefore histological response is likely to vary. The results demonstrate that PET-based treatment decisions in oesophageal adenocarcinoma are feasible, but as this review has demonstrated, the definition of PET response must be optimised first.

In one large single-centre study, Findlay et al [60] investigated prediction of pathological response using PET response between staging and post neoadjuvant chemotherapy. The authors found that a larger reduction of SUVmax after treatment completion may be more

predictive of pathological response. An SUVmax reduction of 77.8% performed better than the PERCIST threshold of 30% [16] and the 35% threshold. A more pragmatic threshold of 75% was suggested for clinical use, which would result in a sensitivity of 73.6% and a specificity of 84.0%. This threshold must be validated in external centres, but this study adds to the evidence that the optimum response threshold is yet to be defined.

Similarly, this review found variation in the definition of pathological response. The two main histopathological classifications used were the Becker [17] and Mandard [18] TRG classifications. The definition of pathological response using histopathological grading is contentious [61]. Studies were included that used a reference standard of either Becker TRG 1 (either 1a, no residual tumour, or 1b, <10% residual tumour) or Mandard TRG 1 or 2 (no residual cancer, or rare residual cancer cells, respectively). Mandard TRG 1 and 2 was found to represent a clinically meaningful response to neo-adjuvant chemotherapy in a multi-centre study of oesophageal adenocarcinoma patients [58]. There was a significant difference in survival between pathological responders (TRG 1-2 median overall survival not reached) and non-responders (TRG 3-5; median overall survival 2.22 years (95% CI 1.94-2.51), $p<0.001$). Further research is needed to define optimum definitions of both PET and pathological response in oesophageal adenocarcinoma.

Ideally, patients treated with a neo-adjuvant regimen would have an early assessment of response to guide subsequent treatment decisions. If objective evidence of positive response was demonstrated, then a patient could continue with that regimen. However, if no response was detected, indicating an ineffective treatment regimen, then an alternative strategy could be sought. This would have the benefit of discontinuing an inefficacious treatment with potential risks of harm and would allow an alternative treatment with potentially beneficial effects to be pursued.

Modification of treatment for early oesophageal cancer based on PET-CT response has been explored using PET-CT in oesophageal adenocarcinoma. A prominent group in the assessment of early PET response in oesophageal adenocarcinoma was the German group who led the series of MUNICON studies. Lordick et al [28] conducted a phase II trial which evaluated a PET-guided treatment decision after 14 days of neo-adjuvant 5-FU and cisplatin.

The study used a pre-defined threshold of 35% reduction in SUVmax to define a metabolic response. The definition was validated from original work from Weber et al [62], though this study was not included in this systematic review because the publication date was prior to 2005 (pre-specified in the review protocol). Metabolic responders continued with the chemotherapy regimen, whereas metabolic non-responders proceeded directly to surgery. There were significant improvements in the R0 resection rate (96% vs 74%, p=0.002), major pathologic response rate (58% vs 0%, p=0.001), median event-free survival (29.7 vs 14.1 months, p=0.002), and median OS (median not reached vs 25.8 months, p=0.015) for PET responders versus PET non-responders.

Furthermore, outcomes were similar between PET non-responders and an independent cohort of patients from [30] who completed 3 months of ineffective chemotherapy before oesophagectomy, suggesting that discontinuation of ineffective chemotherapy was not harmful. The MUNICON-2 trial [31] attempted to improve outcomes of the group of PET non-responders further by evaluating the addition of radiotherapy to the neo-adjuvant chemotherapy regimen to overcome the resilient tumour biology. The PET non-responders had worse 2-year progression-free survival compared to PET responders (64% vs 33%, p=0.035) but 2-year overall survival was not significantly different (71% vs 42%, p=0.10). These results demonstrate that unfavourable tumour biology remains difficult to overcome, although the reported radiotherapy regimen that was delivered varied from standard practice (Table 2).

Further research is required to define PET-based treatment decision making in oesophageal adenocarcinoma. In other tumour sites, image-guided treatment decisions have been adopted into clinical practice [63, 64]. Two large, randomised, phase II trials in oesophageal cancer have recently published data investigating PET response to guide subsequent treatment prior to surgery. Barbour et al [14] randomised oesophageal adenocarcinoma patients using a response threshold of 35% when scanned 15 days after one cycle of cisplatin fluoropyrimidine chemotherapy and baseline PET-CT. Early PET responders continued with the same neoadjuvant chemotherapy regimen prior to oesophagectomy, but PET non-responders were switched to either an alternative chemotherapy regimen (Arm A), or a combination of chemo- and radiotherapy before oesophagectomy (Arm B). In total,

58.5% had a defined metabolic response. Median OS in PET responders was 61 months and in non-responders was 30 and 35 months in Arm A and B, respectively, exemplifying the challenges of resistant cancer biology, independent of treatment approach. The CALGB 80803 trial [15] randomised patients to alternative neo-adjuvant chemoradiotherapy regimens using a PET response criterion of 35%, when scanned 36-42 days after completion of induction chemotherapy. This strategy was deemed effective after the trial demonstrated a significantly higher complete pathological response rate of 26% in PET responders versus 18% in non-responders. This also translated into an overall survival advantage compared to historical controls.

The study protocol for this systematic review initially specified an inclusion criterion that studies should also have a PET-CT after completion of the neo-adjuvant therapy. However, only two studies were identified after screening and neither were eligible for inclusion. Wieder et al was excluded because the patient cohort had already been reported [48] and Gerbaudo et al [49] because only five patients were included in a pilot study comparing FDG and 18F-fluorothymidine (FLT). Wieder et al found an optimal SUVmax reduction threshold of 63% from baseline staging to completion of neo-adjuvant treatment resulted in a sensitivity and specificity of 75% and 87%, respectively. This threshold level is similar to that of 75% proposed by Findlay et al, in this setting.

Strengths of our study include adherence to methodological and reporting recommendations, robust data extraction and quality assessment. Despite the strengths, this study also has limitations. A relatively small number of studies met the strict inclusion criteria set in the study protocol. However, the authors felt this was justified and necessary to provide definitive evidence and this decision was validated by large heterogeneity between study designs and populations. Methodology between future studies must be consistent to define optimum PET response thresholds in external patient cohorts. There were concerns over the quality of evidence in the included studies, but as no co-variates were significant in met-regression analysis, then any clinical impact on early response rate is unlikely to be substantial. Also, there were no studies investigating PET response and neo-adjuvant FLOT regimen, the new standard of care in oesophageal adenocarcinoma.

Conclusions

The pooled early response rate defined by PET imaging was estimated to be 44.7%, but high-quality evidence is lacking, and few studies met the strict inclusion criteria of this systematic review and meta-analysis. The sensitivity of PET using a SUVmax reduction threshold of 35% was found to be suboptimal (77.2%) and varied widely. However, specificity was consistent across studies with a pooled value of 75.0%, suggesting early PET assessment is a more consistent predictor of treatment resistance than of pathological response. Further research is required to define optimal PET-guided treatment decisions in oesophageal adenocarcinoma.

References

1. Arnold M, Soerjomataram I, Ferlay J, Forman D (2015) Global incidence of oesophageal cancer by histological subtype in 2012. Gut 64:381–387.
2. Cunningham D, Allum WH, Stenning SP, et al (2006) Perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. N Engl J Med 355:11–20.
3. Crosby TD, Evans M (2009) The management of locally advanced carcinoma of the oesophagus - SCOPE for improvement? Clin Onc 21:281–284.
4. Sjoquist KM, Burmeister BH, Smithers BM, et al (2011) Survival after neoadjuvant chemotherapy or chemoradiotherapy for resectable oesophageal carcinoma: an updated meta-analysis. Lancet Oncol 12:681–692.
5. Medical Research Council Oesophageal Cancer Working Group (2002) Surgical resection with or without preoperative chemotherapy in oesophageal cancer: a randomised controlled trial. Lancet 359:1727–1733.
6. Allum WH, Stenning SP, Bancewicz J, et al (2009) Long-term results of a randomized trial of surgery with or without preoperative chemotherapy in esophageal cancer. J Clin Oncol 27:5062–5067.
7. Ychou M, Boige V, Pignon JP, et al (2011) Perioperative chemotherapy compared with surgery alone for resectable gastroesophageal adenocarcinoma: an FNCLCC and FFCD multicenter phase III trial. J Clin Oncol 29:1715–1721.
8. van Hagen P, Hulshof MCCM, van Lanschot JJB, et al (2012) Preoperative chemoradiotherapy for esophageal or junctional cancer. N Engl J Med 366:2074–2084.
9. Shapiro J, van Lanschot JJ, Hulshof MC, et al (2015) Neoadjuvant chemoradiotherapy plus surgery versus surgery alone for oesophageal or junctional cancer (CROSS): long-term results of a randomised controlled trial. Lancet Oncol 16:1090–1098.
10. Al-Batran S-E, Homann N, Pauligk C, et al (2019) Perioperative chemotherapy with fluorouracil plus leucovorin, oxaliplatin, and docetaxel versus fluorouracil or capecitabine plus cisplatin and epirubicin for locally advanced, resectable gastric or gastro-oesophageal junction adenocarcinoma (FLOT4): a randomised, phase 2/3 trial. Lancet 393:1948–1957.

11. Allum WH, Blazeby JM, Griffin SM, et al (2011) Guidelines for the management of oesophageal and gastric cancer. Gut 60:1449–1472.
12. Gillies RS, Middleton MR, Maynard ND, et al (2011) Additional benefit of (1)(8)F-fluorodeoxyglucose integrated positron emission tomography/computed tomography in the staging of oesophageal cancer. Eur Radiol 21:274–280.
13. Cremonesi M, Garibaldi C, Timmerman R, et al (2017) Interim (18)F-FDG-PET/CT during chemo-radiotherapy in the management of oesophageal cancer patients: a systematic review. Radiother Oncol 125:200–212.
14. Barbour AP, Walpole ET, Mai GT, et al (2020) Preoperative cisplatin, fluorouracil, and docetaxel with or without radiotherapy after poor early response to cisplatin and fluorouracil for resectable oesophageal adenocarcinoma (AGITG DOCTOR): results from a multicentre, randomised controlled phase II . Ann Oncol Off J Eur Soc Med Oncol 31:236–245.
15. Goodman KA, Hall N, Bekaii-Saab TS, et al (2018) Survival outcomes from CALGB 80803 (Alliance): a randomized phase II trial of PET scan-directed combined modality therapy for esophageal cancer. J Clin Oncol 36(15_suppl):4012.
16. Wahl RL, Jacene H, Kasamon Y, Lodge MA (2009) From RECIST to PERCIST: Evolving Considerations for PET Response Criteria in Solid Tumours. J Nucl Med 50:122S–50S.
17. Becker K, Mueller JD, Schulmacher C, et al (2003) Histomorphology and grading of regression in gastric carcinoma treated with neoadjuvant chemotherapy. Cancer 98:1521–1530.
18. Mandard AM, Dalibard F, Mandard JC, et al (1994) Pathologic assessment of tumor regression after preoperative chemoradiotherapy of esophageal carcinoma. Clinicopathologic correlations. Cancer 73:2680–2686.
19. Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 339:b2535.
20. Foley KG, Jeffries J, Coles B, Smyth EC (2019) Response rate and diagnostic accuracy of PET/CT during and after neo-adjuvant therapies in oesophageal adenocarcinoma: a systematic review. In: PROSPERO. https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=147034
21. Whiting PF, Rutjes AWS, Westwood ME, et al (2011) QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 155:529–536.

22. DerSimonian R, Laird N (1986) Meta-analysis in clinical trials. Control Clin Trials 7:177–188.
23. Balduzzi S, Rucker G, Schwarzer G (2019) How to perform a meta-analysis with R: a practical tutorial. Evid Based Ment Health 22(4). https://ebmh.bmj.com/content/22/4/153
24. Moses LE, Shapiro D, Littenberg B (1993) Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. Stat Med 12:1293–1316.
25. Higgins JPT, Thompson SG, Deeks JJ, Altman DG (2003) Measuring inconsistency in meta-analyses. BMJ 327:557–560.
26. Guyatt GH, Oxman AD, Vist GE, et al (2008) GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 336:924–926.
27. Harustiak T, Zemanova M, Fencl P, et al (2018) [18 F]Fluorodeoxyglucose PET/CT and prediction of histopathological response to neoadjuvant chemotherapy for adenocarcinoma of the oesophagus and oesophagogastric junction. Br J Surg 105:419–428.
28. Lordick F, Ott K, Krause BJ, et al (2007) PET to assess early metabolic response and to guide treatment of adenocarcinoma of the oesophagogastric junction: the MUNICON phase II trial. Lancet Oncol 8:797–805.
29. Malik V, Lucey JA, Duffy GJ, et al (2010) Early repeated 18F-FDG PET scans during neoadjuvant chemoradiation fail to predict histopathologic response or survival benefit in adenocarcinoma of the esophagus. J Nucl Med 51:1863–1869.
30. Ott K, Weber WA, Lordick F, et al (2006) Metabolic imaging predicts response, survival, and recurrence in adenocarcinomas of the esophagogastric junction. J Clin Oncol 24:4692–4698.
31. zum Buschenfelde CM, Herrmann K, Schuster T, et al (2011) (18)F-FDG PET-guided salvage neoadjuvant radiochemotherapy of adenocarcinoma of the esophagogastric junction: the MUNICON II trial. J Nucl Med 52:1189–1196.
32. Burtness B, Ilson D, Iqbal S (2014) New directions in perioperative management of locally advanced esophagogastric cancer. Am Soc Clin Oncol Educ book Am Soc Clin Oncol Annu Meet e172-8. https://ascopubs.org/doi/full/10.14694/EdBook_AM.2014.34.e172

33. Chen YM, Pan XF, Tong LJ, et al (2011) Can F-18-fluorodeoxyglucose positron emission tomography predict responses to neoadjuvant therapy in oesophageal cancer patients? A meta-analysis. Nucl Med Commun 32:1005–1010.
34. De Geus-Oei LF, Vriens D, Arens AIJ, et al (2012) FDG-PET/CT based response-adapted treatment. Cancer Imaging 12:324–335.
35. Kwee RM (2010) Prediction of tumor response to neoadjuvant therapy in patients with esophageal cancer with use of 18F FDG PET: a systematic review. Radiology 254:707–717.
36. Ngamruengphong S, Sharma VK, Nguyen B, Das A (2010) Assessment of response to neoadjuvant therapy in esophageal cancer: an updated systematic review of diagnostic accuracy of endoscopic ultrasonography and fluorodeoxyglucose positron emission tomography. Dis Esophagus 23:216–231.
37. Omloo JM, van Heijl M, Hoekstra OS, et al (2011) FDG-PET parameters as prognostic factor in esophageal cancer patients: a review. Ann Surg Oncol 18:3338–3352.
38. Sendler A (2010) Metabolic response evaluation by PET during neoadjuvant treatment for adenocarcinoma of the esophagus and esophagogastric junction. Recent Results Cancer Res 182:167–177.
39. Fang P, Musall BC, Son JB, et al (2018) Multimodal Imaging of Pathologic Response to Chemoradiation in Esophageal Cancer. Int J Radiat Oncol Biol Phys 102:996-1001.
40. Klaeser B, Nitzsche E, Schuller JC, et al (2009) Limited predictive value of FDG-PET for response assessment in the preoperative treatment of esophageal cancer: results of a prospective multi-center trial (SAKK 75/02). Onkologie 32:724–730.
41. Hagen P V, Heijl M V, Henegouwen MIVB, et al (2017) Prediction of disease-free survival using relative change in FDG-uptake early during neoadjuvant chemoradiotherapy for potentially curable esophageal cancer: A prospective cohort study. Dis esophagus Off J Int Soc Dis Esophagus 30:1–7.
42. van Heijl M, Omloo JM, van Berge Henegouwen MI, et al (2011) Fluorodeoxyglucose positron emission tomography for evaluating early response during neoadjuvant chemoradiotherapy in patients with potentially curable esophageal cancer. Ann Surg 253:56–63.
43. Vihervaara H, Algars A, Kemppainen J, et al (2019) Perioperative EOX treatment in operable locally advanced gastroesophageal adenocarcinoma: Prediction of tumor

response by FDG -PET and histopathology. Surg Oncol 28:42–49.

44. Langer R, Ott K, Specht K, et al (2008) Protein Expression Profiling in Esophageal Adenocarcinoma Patients Indicates Association of Heat-Shock Protein 27 Expression and Chemotherapy Response. Clin Cancer Res 14:8279–8287.
45. Siewert JR, Lordick F, Ott K, et al (2007) Induction chemotherapy in Barrett cancer: influence on surgical risk and outcome. Ann Surg 246:624–631.
46. Weber MA, Bender K, von Gall CC, et al (2013) Assessment of diffusion-weighted MRI and 18F-fluoro-deoxyglucose PET/CT in monitoring early response to neoadjuvant chemotherapy in adenocarcinoma of the esophagogastric junction. J Gastrointest Liver Dis 22:45–52.
47. Wieder HA, Beer AJ, Lordick F, et al (2005) Comparison of changes in tumor metabolic activity and tumor size during chemotherapy of adenocarcinomas of the esophagogastric junction. J Nucl Med 46:2029–2034.
48. Wieder HA, Ott K, Lordick F, et al (2007) Prediction of tumor response by FDG-PET: comparison of the accuracy of single and sequential studies in patients with adenocarcinomas of the esophagogastric junction. Eur J Nucl Med Mol Imaging 34:1925–1932.
49. Gerbaudo VH, Killoran JH, Kim CK, et al (2018) Pilot study of serial FLT and FDG-PET/CT imaging to monitor response to neoadjuvant chemoradiotherapy of esophageal adenocarcinoma: correlation with histopathologic response. Ann Nucl Med 32:165–174.
50. Harada K, Wang X, Shimodaira Y, et al (2018) Early Metabolic Change after Induction Chemotherapy Predicts Histologic Response and Prognosis in Patients with Esophageal Cancer: Secondary Analysis of a Randomized Trial. Target Oncol 13:99–106.
51. Ku GY, Kriplani A, Janjigian YY, et al (2016) Change in chemotherapy during concurrent radiation followed by surgery after a suboptimal positron emission tomography response to induction chemotherapy improves outcomes for locally advanced esophageal adenocarcinoma. Cancer 122:2083–2090.
52. Xiong JF, Yu W, Ma JC, et al (2018) The Role of PET-Based Radiomic Features in Predicting Local Control of Esophageal Cancer Treated with Concurrent Chemoradiotherapy. Sci Rep 8(1).

53. Kim SJ, Koo PJ, Chang S (2016) Predictive value of repeated F-18 FDG PET/CT parameters changes during preoperative chemoradiotherapy to predict pathologic response and overall survival in locally advanced esophageal adenocarcinoma patients. Cancer Chemother Pharmacol 77:723–731.
54. Xi M, Liao Z, Hofstetter WL, et al (2017) (18)F-FDG PET Response After Induction Chemotherapy Can Predict Who Will Benefit from Subsequent Esophagectomy After Chemoradiotherapy for Esophageal Adenocarcinoma. J Nucl Med 58:1756–1763.
55. De Vita F, Orditura M, Martinelli E, et al (2011) A multicenter phase II study of induction chemotherapy with FOLFOX-4 and cetuximab followed by radiation and cetuximab in locally advanced oesophageal cancer. Br J Cancer 104:427–432.
56. Westerterp M, Omloo JMT, Sloof GW, et al (2006) Monitoring of response to pre-operative chemoradiation in combination with hyperthermia in oesophageal cancer by FDG-PET. Int J Hyperthermia 22:149–160.
57. National Oesophago-Gastric Cancer Audit (NOGCA) (2019) An Audit of the Care Received by People with Oesophago-Gastric Cancer in England and Wales Annual Report. https://www.nogca.org.uk/reports/2019-annual-report/
58. Noble F, Lloyd MA, Turkington R, et al (2017) Multicentre cohort study to define and validate pathological assessment of response to neoadjuvant therapy in oesophagogastric adenocarcinoma. Br J Surg 104:1816–1828.
59. Yang GY, Wagner TD, Jobe BA, Thomas CR (2008) The role of positron emission tomography in esophageal cancer. Gastrointest Cancer Res 2:3–9.
60. Findlay JM, Bradley KM, Wang LM, et al (2017) Predicting Pathologic Response of Esophageal Cancer to Neoadjuvant Chemotherapy: The Implications of Metabolic Nodal Response for Personalized Therapy. J Nucl Med 58:266–275.
61. Zhu Y, Sun Y, Hu S, et al (2017) Comparison of five tumor regression grading systems for gastric adenocarcinoma after neoadjuvant chemotherapy: a retrospective study of 192 cases from National Cancer Center in China. BMC Gastroenterol 17:41.
62. Weber WA, Ott K, Becker K, et al (2001) Prediction of response to preoperative chemotherapy in adenocarcinomas of the esophagogastric junction by metabolic imaging. J Clin Oncol 19:3058–3065.
63. Mehanna H, Wong WL, McConkey CC, et al (2016) PET-CT Surveillance versus Neck Dissection in Advanced Head and Neck Cancer. N Engl J Med 374:1444–1454.

64. Kasivisvanathan V, Rannikko AS, Borghi M, et al (2018) MRI-Targeted or Standard Biopsy for Prostate-Cancer Diagnosis. N Engl J Med 378:1767–1777.

Tables

Table 1. Studies reporting early metabolic response and diagnostic accuracy that met the inclusion criteria

Study (Year)	Study period	Region	Sites	Design	Total Patients	Number resected	Patient age range (years)	Male patients (%)	GOJ tumours (%)	Neo-adjuvant treatment	Dose	Interval baseline to early PET (range days)	Pre-injection fasting (mins)	Injection to PET interval (mins)
Barbour et al [14] (2020)	2009-2015	Australia	8	Randomised controlled phase II trial	124	107	38-78	110/124 (88.7)	124/124 (100)	NACT + RT in non-responders	4.5MBq /kg	14-21	NA	60
Harustiak et al [27] (2018)	2009-2015	Czech Republic	2	Prospective observational cohort	126	90	27-75	93/126 (73.8)	90/90 (100)	NACT	4·00 MBq / kg	12-16	360	75
Lordick et al [28] (2007)	2002-2005	Germany	1	Prospective phase 2 trial	119	104	NA	111/119 (93.3)	104/104 (100)	NACT	300–400 MBq	14	360	40
Malik et al [29] (2010)	2003-2007	Ireland	1	Prospective observational cohort	37	37	37-73	31/37 (83.8)	NA	NACRT	NA	9-14	360	60
Ott et al [30] (2006)	1999-2002	Germany	1	Prospective observational cohort	56	56	NA	50/56 (89.3)	56/56 (100)	NACT	300–400 MBq	14	360	40
zum Buschenfelde et al [31] (2011)	2005-2008	Germany	1	Prospective non-randomised phase II cohort	56	56	35-77	51/56 (91.1)	56/56 (100)	NACT + RT in non-responders	Mean 447 MBq PET1 and 406 MBq PET2	14	360	60

NA data not reported; PET positron emission tomography; NACT neo-adjuvant chemotherapy; RT radiotherapy; NACRT neo-adjuvant chemoradiotherapy

Table 2. Details of neo-adjuvant therapy regimen and survival in patient cohorts.

Study	Chemotherapy regimen	Radiotherapy (dose/schedule)	Percentage of metabolic responders	mOS responders	mOS non-responders	Median follow-up time
Barbour et al [14] (2020)	<u>Induction</u> Cisplatin 80 mg/m^2, 5-fluorouracil 1000 mg/m^2 per day over 96 hours as a continuous infusion. Early metabolic responders received a second round of CF on Day 22, followed by oesophagectomy. <u>Randomisation</u> Non-responder group Arm A = DCF (Arm A) intravenous docetaxel 35mg/m^2 on days 22, 29, and 36 and days 50, 57 and 64, and bolus cisplatin 60 mg/m^2 on days 22 and 50 after the completion of docetaxel. Following the cisplatin dose on day 22, 5-fluorouracil 150 mg/m^2 per day was given as a continuous infusion for 56 days. DCFRT (Arm B).	Metabolic responders continued with previous chemotherapy regimen then proceeded to surgery. Metabolic non-responders were randomised to either Arm A (DCF chemo) or Arm B (DCF + RT = 45 Gy in 25 fractions)	58.5%	61 months	Arm A = 30 months Arm B = 35 months	62 months
Harustiak et al [27] (2018)	3 pre-op and 3 post-op cycles of epirubicin (50mg/m^2) and cisplatin (60mg/m^2) – day 1 + fluoracil (200mg/m^2) 21 days infusion or 21 days of capecitabine(1g/m^2) for 14 days	None	37%	NA	NA	NA
Lordick et al [28] (2007)	2 cycles of cispatin (50mg/m^2) days 1,15,29 + folinic acid 500mg/m2 + fluoracil (2g/m^2) – days 1,8,15,22,29 and 36, then repeated at day 49. If eGFR less than 60mL/kg/min, then oxaliplatin (85mg/m^2) replaced cisplatin. If aged 60 or younger with good health status, then additional paclitaxel (80mg/m^2) given on days 0, 14 and 28.	None	49%	Not reached	25.8 months (95% CI 19.4-32.2)	NA
Malik et al	5-fluorouracil (15 mg/kg) on days 1 to 5 and cisplatin (75 mg/	Total dose of 40 Gy in 15	26.4%	Not reached	21 months	47 months

[29] (2010)	m^2) on day 7. Chemotherapy was repeated in week 6.	daily fractions (4 weeks after completion of neo-adjuvant chemotherapy)				
Ott et al [30] (2006)	2 cycles (49 days each) cisplatin (50mg/m^2) on day 1 + leucovorin (500mg/m^2 BSA) + fluoracil (2g/m^2 BSA), then cisplatin on days 15 and 29, then leucovorin and fluoracil on days 8, 15, 22, 29 and 36. For type 1 junctional tumours, paclitaxel (80mg/m^2 BSA) was added.	None	32%	Not reached	18 months	NA
zum Buschenfelde et al [31] (2011)	2 weeks of same regimen as in Lordick et al [28]. After 2 weeks, metabolic responders continued on this regimen. Metabolic non-responders started cisplatin 6mg/m^2 (days 1-5 and 8-12), or 5- fluorouracil (250mg/m^2) if poor renal function.	For metabolic non responders, a total of 32Gy at 1.6Gy/fraction was added (twice daily/ 10 fractions a week)	59%	Not reached	18.3 months	38 months (95% CI 14-54)

NA data not reported; mOS median overall survival; CF cisplatin and 5-fluorouracil; DCF cisplatin, 5-fluorouracil and docetaxel; DCFRT cisplatin, 5-fluorouracil, docetaxel and radiotherapy; RT radiotherapy; eGFR estimated glomerular filtration rate; Gy gray; BSA body surface area; CI confidence intervals

Table 3. Sensitivity and specificity of early PET response to predict pathological response.

Study	Total resected	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
Lordick et al [28] (2007)	104	29	21	0	54	1.00 (0.88-1.00)	0.72 (0.60-0.82)	0.58 (0.43-0.72)	1.00 (0.93-1.00)
Malik et al [29] (2010)	37	7	5	9	16	0.44 (0.20-0.70)	0.76 (0.53-0.92)	0.58 (0.28-0.85)	0.64 (0.43-0.82)
Ott et al [30] (2006)	56	8	10	2	36	0.80 (0.44-0.97)	0.78 (0.64-0.89)	0.44 (0.22-0.69)	0.95 (0.82-0.99)

TP true positives; FP false positives; FN false negatives; TN true negatives

Table 4. QUADAS-2 quality assessment of included studies

Study	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Barbour et al [14] (2020)	Low	Low	Unclear	Low	Low	Low	Unclear
Harustiak et al [27] (2018)	Unclear	Low	High	Low	Low	Low	Unclear
Lordick et al [28] (2007)	Unclear	Low	Unclear	Low	Low	Low	Low
Malik et al [29] (2010)	Unclear	Low	Low	Low	Low	Low	Low
Ott et al [30] (2006)	Unclear	Low	Low	Low	Low	Low	Low
zum Buschenfelde et al [31] (2011)	Unclear	Low	Unclear	Low	Low	Low	Low

Figure Legends

Figure 1. Study selection process.

Figure 2. Pooled estimate of early metabolic response rate after one cycle of neoadjuvant therapy.

Figure 3. Pooled estimates of sensitivity (A) and specificity (B). TP = true positives, FP = false positives, FN = false negatives, TN = true negatives.