Prognostic significance of SUV on PET/CT in patients with localised oesophagogastric junction cancer receiving neoadjuvant chemotherapy/chemoradiation: a systematic review and meta-analysis

¹W ZHU, MD, ¹L XING, MD, ¹J YUE, MD, ²X SUN, MD, ¹X SUN, MD, ¹H ZHAO, MD and ¹J YU, MD, PHD

¹Department of Radiation Oncology, Shandong Cancer Hospital, Shandong Academy of Medical Sciences, Provincial Key Laboratory of Radiation Oncology, Jinan, China, and ²Department of Nuclear Medicine, Shandong Cancer Hospital, Shandong Academy of Medical Sciences, Provincial Key Laboratory of Radiation Oncology, Jinan, China

Objective: The objective of this study was to comprehensively review the evidence for use of pre-treatment, post-treatment and changes in tumour glucose uptake that were assessed by 18-fluorodeoxyglucose (¹⁸F-FDG) positron emission tomography (PET) early, during or immediately after neoadjuvant chemotherapy/chemoradiation to predict prognosis of localised oesophagogastric junction (AEG) cancer.

Methods: We searched for articles published in English; limited to AEG; ¹⁸F-FDG uptake on PET performed on a dedicated device; dealt with the impact of standard uptake value (SUV) on survival. We extracted an estimate of the log hazard ratios (HRs) and their variances and performed meta-analysis.

Results: 798 patients with AEG were included. And the scan time for ¹⁸F-FDG-PET was as follows: prior to therapy (PET1, *n*=646), exactly 2 weeks after initiation of neoadjuvant therapy (PET2, *n*=245), and pre-operatively (PET3, *n*=278). In the two meta-analyses for overall survival, including the studies that dealt with reduction of tumour maximum SUV (SUV_{max}) (from PET1 to PET2/PET3 and from PET1 to PET2), the results were similar, with the overall HR for non-responders being 1.83 [95% confidence interval (CI), 1.41–2.36] and 2.62 (95% CI, 1.61–4.26), respectively; as for disease-free survival, the combined HR was 2.92 (95% CI, 2.08–4.10) and 2.39 (95% CI, 1.57–3.64), respectively. The meta-analyses did not attribute significant prognostic values to SUV_{max} before and during therapy in localised AEG.

Conclusion: Relative changes in FDG-uptake of AEG are better prognosticators. Early metabolic changes from PET1 to PET2 may provide the same accuracy for prediction of treatment outcome as late changes from PET1 to PET3.

Received 30 March 2011
Revised 16 September 2011
Accepted 17 October 2011

DOI: 10.1259/bjr/29946900

© 2012 The British Institute of Radiology

It is known from several studies that the outcome of patients with adenocarcinomas of the oesophagogastric junction (AEG) treated with pre-operative chemotherapy/chemoradiation is heterogeneous. To the best of our knowledge, the reasons for this unpredictability in clinical outcome are not entirely clear but could be attributed to the differences in molecular compositions of cancers [1–5] and/or patient genetics [6, 7]. Two Phase III studies indicated that pre-operative chemotherapy improved survival in patients with oesophageal adenocarcinoma and AEG [8, 9]. However, a systematic review did show only marginal effects of pre-operative chemotherapy for resectable intrathoracic oesophageal cancer [10]. Of note, in non-responders, survival seems to be similar or even worse than after surgical resection alone [11]. Therefore, early identification of patients likely to have an unfavourable outcome after pre-operative therapy is highly important for the future use of neoadjuvant therapy in AEG.

The use of 18-fluorodeoxyglucose (¹⁸F-FDG) positron emission tomography (PET) as a metabolism-based imaging technique has been increased steadily during the last decade in most malignant tumours. The prognostic values of FDG uptake before and after chemotherapy/chemoradiotherapy were reported, and other studies have been conducted to evaluate the prognostic value of a decrease in tumour metabolic activity on the survival of patients with AEG. However, the sample size of most studies was rather small and the results of the prognostic value of standardised uptake value (SUV) remained undetermined. Therefore, we performed a meta-analysis to assess the prognostic value of SUV for survival of patients with AEG receiving neoadjuvant chemotherapy/chemoradiation.

Methods and materials

Literature search

We searched the MEDLINE and Embase databases for articles published between January 1998 and January

Address correspondence to: Dr Jinming Yu, Department of Radiation Oncology, Shandong Cancer Hospital, Shandong Academy of Medical Sciences, Provincial Key Laboratory of Radiation Oncology, Jiyuan Road 440, Jinan 250117, Shandong Province, China. E-mail: yujmwin@126.com

2011, using the following terms: “oesophagogastric/oesophagogastric junction” or “oesophagogastric/oesophagogastric junction cancer” or “oesophagogastric/oesophagogastric junction carcinoma” or “oesophagogastric/oesophagogastric junction neoplasm”; “PET imaging tomography” or “PET” or “positron emission tomography” or “FDG” or “¹⁸F-FDG” or “FDG-F18” or “fluorodeoxyglucose” or “¹⁸F-fluorodeoxyglucose”; “standardized uptake value” or “SUV” or “uptake value” or “semiquantitation”; and “outcome” or “survival” or “predict” or “prognosis” or “prognostic factor”. A manual search of the cross-references for eligible articles was used to identify additional relevant articles.

Selection of studies

Four investigators reviewed each article independently and scored them according to a quality scale (described in Appendix A). The methodology quality assessment consisted of four main dimensions modified on the basis of similar studies: the scientific design, the generalisability of the results, the analysis of the study data and the PET reports [12, 13]. A points value (0–2) was attributed to each item. Each dimension was worth 0–10 points up to a maximum of 40 points. As the scores were objective, a consensus was obtained in meetings with at least three-quarters of the investigators present. Their participation guaranteed the correct interpretation of the publications. The final scores were expressed as percentages, with higher values reflecting a greater application of methodological standards. Because the scoring of quality is intrinsically subjective, the quality scores were not applied to exclude lower quality studies from the meta-analysis or to weigh the studies. The studies included in the systematic review were called “eligible” and those providing sufficient data for meta-analysis “evaluable”.

Statistical methods

SUV cut-off values to differentiate responders from non-responders were based on the definition used in each individual study. The correlation between the quality scores and the number of patients included in the studies was measured by the Spearman’s rank correlation coefficient, and testing a null hypothesis of equality to zero of the coefficient assessed its significance. Non-parametric tests (Mann–Whitney *U*-test or Kruskal–Wallis test) were applied to compare the distribution of the quality scores according to the value of a discrete variable. The effect of SUV on survival was measured by the hazard ratio (HR) between the survival distributions of two groups using the following methodology. For each study, the log HR and standard error can be calculated by extracting the unadjusted HR and confidence intervals (CIs) directly from each publication or from extracting cumulative survival data from published Kaplan–Meier plots as described by Parmar et al [14]. If authors reported survival of more than two groups, we pooled the results, making a comparison between two groups feasible.

The *Q* statistics were applied to test for heterogeneity among the evaluable studies. A fixed-effect model was applied to calculate the summary HRs, if there was no heterogeneity observed (*i.e.* *Q*-test *p* > 0.05). When heterogeneity was observed, a random-effects model was used. The *I*² statistic was applied to estimate the percentage of variation across studies due to heterogeneity rather than chance. *I*² can be calculated as following:

$$I^2 = 100\% \times (Q \text{ d.f.}) / Q \quad (1)$$

Q = Cochran’s heterogeneity statistics, d.f. = degree of freedom. We defined substantial heterogeneity within every meta-analysis as an *I*² > 50%. Publication bias was detected by performing the Egger test.

Studies about local control or recurrence were also included in the meta-analysis for disease-free survival (DFS). Survival rates on the graphical representation of the survival curves were read by Engauge Digitizer v. 2.5 (Trolltech, Oslo, Norway). HRs and their variations were calculated by Review Manager v. 5.0 (The Nordic Cochrane Centre, Copenhagen, Denmark). All reported *p*-values are two-sided and are performed at the 5% level of significance using SPSS® v. 13.0 (SPSS Inc., Chicago, IL).

Results

Study selection and characteristics analysis

The electronic and manual searches yielded 97 potentially eligible articles from all databases. Of these articles, 63 were eliminated because they were not about localised oesophagogastric junction carcinoma, without an outcome of interest, not full-text articles or non-English language. The remaining 34 full-text articles were further analysed. 24 of these studies were excluded because the log HR and its variance of the oesophagogastric junction carcinoma from the total patients with oesophageal or gastric carcinoma could not be calculated (*n* = 12) [15–26], an author published two reports on the same population (*n* = 1) [27, 28], and irrelevant references including the studies without surgery such as Murthy et al (*n* = 11) [29]. Finally, a total of 10 studies were determined to be evaluable for the actual analysis [27, 30–37]. The principal characteristics of the 10 studies evaluable for the meta-analysis are described in Table 1. The scan time for ¹⁸F-FDG-PET was as follows: prior to therapy (PET1), exactly 2 weeks after initiation of neoadjuvant therapy (PET2) and pre-operatively (PET3). Six studies dealt with the prognostic value of changes in glucose utilisation, measured by ¹⁸F-PDG-PET from PET1 to PET2 or PET3 for overall survival (OS) [30, 31, 33, 35, 37, 38]. Four studies dealt with the prognostic value of changes in maximum SUV (SUV_{max}) measured on ¹⁸F-PDG-PET for DFS at the same time [30, 34, 35, 37]. Moreover, the meta-analyses were performed for the studies that dealt only with changes in SUV_{max} from PET1 to PET2 for predictions of OS and DFS. Three studies dealt with the prognostic value of SUV_{max} measured on ¹⁸F-PDG-PET at PET1 [27, 32, 35] for OS, and three studies did it at PET2/PET3 for OS [31, 35, 36]. The meta-analyses that dealt only with SUV for DFS prediction at PET1 [35] and PET2/PET3 [35, 36] were excluded owing to too few articles.

Table 1. Principal characteristics of the 10 studies included in the meta-analysis

Study	Publication year	Number of patients	Stage	Location	Methodology score (%)
Weber et al [37]	2001	40	Tumour stage T3/T4, NX, and M0	Types I and II	81.6
Ott et al [35]	2006	56	IIa–III	Types I and II	78.8
Wieder et al [33]	2007	24	cT3–4, N0/+, M1a	Types I and II	50
Lordick et al [30]	2007	104	cT3 or cT4	Types I and II	76.3
Roedl et al [38]	2008	51	Not detailed	Type I	55.3
Smith et al [34]	2009	21	Not detailed	Not detailed	36.8
Rizk et al [27]	2009	189	II–IVa	Types I and II	57.9
Javeri et al [32]	2009	161	T2N0–M1a	Types I, II and III	55.3
Javeri et al [31]	2009	151	T2N0–M1a	Types I, II and III	57.9
Patnana et al [36]	2010	152	II–IV	Types I and II	65.8

A total of 798 patients, with a predominance of adenocarcinoma of the distal oesophagus, were included in survival analyses. SUV_{max} normalised by body weight was used in all studies and the mean SUV was not used. Only five of the six studies for OS achieved definite statistical significance in addition to three of the four for DFS. Two of the studies with significant DFS results also had significant OS results [35, 39], while one study showed an undetermined effect on both OS and DFS [37]. Furthermore, in one study, the statistical result was different according to the threshold definition [33]. Table 2 shows the main SUV_{max} characteristics reported in each article. The response threshold in SUV_{max} chosen arbitrarily between patients with high and low survival was based on the SUV_{max} values decreasing by $\geq 35\%$ from PET1 to PET2 in all studies. In the four studies that had the PET scan at PET3, a so-called “best cut-off” was used, which meant that the threshold maximised the log rank test statistic among several survival comparisons. It is known that this method may lead to a high risk of false-positive results, especially without adjustment of p -values for multiplicity. Five studies indicated that patients with a high SUV_{max} had a worse OS than

patients with a low SUV_{max} at PET1, PET2 or PET3. However, none of the studies achieved definite statistical significance. Three studies used the median as the SUV_{max} threshold, and the best cut-off value was used in two studies in addition to analysing SUV_{max} as a continuous variable in one study [31].

Quality assessment

Overall, the global quality score ranged from 36.8% to 81.6%, with a median of 57.9% (Table 1). An attempt was made to contact the authors, if necessary, to obtain missing details of methodological quality. There was a non-significant correlation between the global score and the number of patients included in the study (Spearman's correlation coefficient, $r=0.250$; $p=0.486$).

Meta-analysis

For OS, four meta-analyses were performed. In the first meta-analysis for OS, six studies that dealt with the

Table 2. Main SUV characteristics extracted from the 10 articles used for meta-analysis

Study	Type of SUV	Correction of SUV	Threshold definition	SUV threshold
Weber et al [37]	Decreased SUV_{max} (from PET1 to PET2)	Weight	Best cut-off	$>35\%$
Ott et al [35]	Initial SUV_{max} (at PET1)	Weight	Median	8.1
	FDG uptake at day 14 (at PET2)	Weight	Median	5.4
	Decreased SUV_{max} (from PET1 to PET2)	Weight	Best cut-off	$>35\%$
Wieder et al [33]	Decreased SUV_{max} (from PET1 to PET2)	Weight	Best cut-off	$>35\%$
	Decreased SUV_{max} (3–4 weeks after the completion of chemoradiotherapy, from PET1 to PET3)	Weight	Best cut-off	$>63\%$
Lordick et al [30]	Decreased SUV_{max} (from PET1 to PET2)	Weight	Previous report	$>35\%$
Roedl et al [38]	Decreased SUV_{max} (16.9 days $\pm$ 6.8 after chemoradiotherapy, from PET1 to PET3)	Weight	Best cut-off	$>43\%$
Smith et al [34]	Decreased SUV_{max} (from PET1 to PET2)	Weight	Best cut-off	$>50\%$
Rizk et al [27]	Initial SUV_{max} (at PET1)	Weight	Previous report	4.5
Javeri et al [32]	Initial SUV_{max} (at PET1)	Weight	Median	10.1
Javeri et al [31]	FDG uptake at day (12 $\pm$ 2 weeks, at PET3)	Weight	Continuous variable	—
	Decreased SUV_{max} (12 $\pm$ 2 weeks after initiation of therapy, from PET1 to PET3)	Weight	Previous report	$>52\%$
Patnana et al [36]	SUV_{max} (Approximately 5–6 weeks after the completion of chemoradiation, at PET3)	Weight	Median	4.6

FDG, fludeoxyglucose; PET, positron emission tomography; PET1, PET scan prior to therapy; PET2, PET scan exactly 2 weeks after the initiation of neoadjuvant therapy; PET3, PET scan pre-operatively; SUV, standardised uptake value; SUV_{max} , maximum standardised uptake value.

— indicates that SUV was not used as a categorical variable in the article.

(a)

(b)

response in SUV_{max} measured from PET1 to PET2 or PET3 were included (Figure 1a). The combined HR was 1.83 (95% CI, 1.41–2.36) with a fixed-effect model, meaning that metabolic responders had a better OS than metabolic non-responders. The test for heterogeneity gave no significant results (Q -test, $p = 0.31$, $I^2 = 16\%$). The evaluation of publication bias showed that the Egger test was not significant ($p = 0.237$). The funnel plots for publication bias (Figure 2) also showed some symmetry. The second meta-analysis for OS included the studies that dealt with the changes in tumour SUV_{max} measured from PET1 to PET2. The combined HR was with 2.62 (95% CI, 1.61–4.26); heterogeneity (Q -test, $p = 0.55$, $I^2 = 0\%$; Figure 1b). As for the prognostic value of SUV_{max} measured on ^{18}F -PDG-PET at PET1 and PET2/PET3 for OS, the combined HRs were 1.52 (95% CI, 0.99–2.34) and 1.05 (95% CI, 0.96–1.16), respectively (Figure 3). This means that a high primary tumour SUV_{max} was not associated with a worse survival prognosis. Similarly, there were four studies that dealt with the prognostic value of response in primary tumour SUV_{max} for DFS [30, 34, 35, 37]. With the studies that dealt with the change in tumour SUV_{max} measured from PET1 to PET2 or PET3 (Figure 4a), the combined HR was 2.92 (95% CI, 2.08–4.10). The results of the test for heterogeneity were insignificant (Q -test, $p = 0.30$, $I^2 = 18\%$). The evaluation of publication bias showed that the Egger test was insignificant ($p = 0.371$) in addition to the funnel plots. After excluding the studies that dealt with the response in SUV_{max} measured from PET1 to PET3 [30, 35, 37], the combined HR was 2.39 (95% CI, 1.57–3.64; Q -test, $p = 0.57$, $I^2 = 0\%$; Figure 4b).

Discussion

Oesophagogastric carcinoma is frequently diagnosed at an advanced stage, and the number of deaths, as a result, has continued to rise [39, 40]. For localised adenocarcinomas, surgery following neoadjuvant

Figure 1. Review: meta-analyses of the studies dealt with the prognostic value of response in maximum standardised uptake value measured by fludeoxyglucose positron emission tomography (PET) (a) at least 2 weeks [from prior to therapy (PET1) to 2 weeks after initiation of neoadjuvant therapy (PET2) or pre-operatively] and (b) exactly 2 weeks (from PET1 to PET2) after initiation of therapy for overall survival. Results in (a) indicate that metabolic responders from PET1 to PET2/PET3 had a better overall survival than metabolic non-responders; results in (b) indicate that metabolic responders from PET1 to PET2 had a better overall survival than metabolic non-responders. CI, confidence interval; HR, hazard ratio; SE, standard error.

therapy remains the mainstay of treatment [41]. However, only patients who responded to induction therapy gained survival benefits. Individual, response-guided treatment concepts in oncology are needed [42]. Despite intensive efforts to identify molecular markers, the value of each of these markers is currently not sufficiently validated to use them for the selection of patients for therapy [43–45]. Most previously reported studies showed that changes in tumour glucose uptake assessed by ^{18}F -FDG PET during or immediately after neoadjuvant therapy were associated with response and prognosis in oesophageal and oesophagogastric cancer [15, 21, 25, 35, 37, 42, 46–50]. Therefore, a meta-analysis is useful because it reduces the effect that chance plays on the individual result, and quality assessments are important for reducing bias. The methodological quality of this study was moderate considering the median score of 57.9%. Sensitivity analysis was not performed owing to insufficient data.

The current meta-analysis provides two possible findings regarding the use of ^{18}F -PDG-PET for monitoring chemotherapy of localised oesophagogastric cancer. Firstly, early metabolic changes (14 days after the start of therapy, PET2) might provide the same accuracy for the prediction of treatment outcome as late changes (after the end of pre-operative therapy, PET3). Secondly, the predictive value of relative changes in tumour FDG-uptake might be superior to measurements of tumour FDG-uptake at PET1, PET2 or PET3. In the meta-analyses of the studies that dealt with the response of SUV_{max} measured from PET1 to PET2, the results showed that the responses were associated with a significantly better OS and DFS, indicating that the pejorative impact of non-response in SUV_{max} on both OS and DFS was statistically significant. In all the meta-analyses, including the studies that dealt with the response of SUV_{max} regardless of the second scan time, the overall HR results were similar and positive. Therefore, PET might be used as a reference method for early response assessment in oesophagogastric

Figure 2. Funnel graph for the assessment of potential publication bias in studies of response in standard uptake value in patients with oesophagogastric junction cancer. (a) The studies for overall survival and (b) the studies for disease-free survival. HR, hazard ratio; SE, standard error.

Figure 3. Review: meta-analyses of the studies dealt with the prognostic values of the maximum standardised uptake value measured by fludeoxyglucose positron emission tomography (a) prior to therapy and (b) at least 2 weeks after the initiation of therapy (at 2 weeks after the initiation of neoadjuvant therapy or pre-operatively) for overall survival. Results in (a) indicate that a high primary tumour maximum standard uptake value (SUV_{max}) at positron emission tomography 1 (PET1) was not associated with a worse overall survival prognosis. Results in (b) indicate that a high primary tumour SUV_{max} at PET2/PET3 was not associated with a worse overall survival prognosis. CI, confidence interval; HR, hazard ratio; SE, standard error.

Figure 4. Review: meta-analyses of the studies dealt with the prognostic value of response in maximum standardised uptake value measured by fludeoxyglucose positron emission tomography (PET) (a) at least 2 weeks [from prior to therapy (PET1) to exactly 2 weeks after initiation of neoadjuvant therapy (PET2) or pre-operatively] and (b) exactly 2 weeks after initiation of therapy (from PET1 to PET2) for disease-free survival. Results in (a) indicate that metabolic responders from PET1 to PET2/PET3 had a better disease-free survival (DFS) than metabolic non-responders. Results in (b) indicate that metabolic responders from PET1 to PET2 had a better DFS than metabolic non-responders. CI, confidence interval; HR, hazard ratio; SE, standard error.

cancer. Furthermore, the meta-analyses did not attribute a significant prognostic value to SUV_{max} before and during therapy in localised AEG. It was encouraging for the clinical application of ^{18}F -PDG-PET for prediction of treatment response that relative changes provided the same or higher accuracy for predicting or assessing tumour response as SUV_{max} before and during therapy [51].

However, the results of this systematic review might be criticised on several points. Firstly, the number and quality of studies included in each analysis might be different, which might affect the indirect comparison of results at the different PET scan time. Secondly, SUV estimates might suffer from poor reproducibility because of the lack of standardisation of the acquisition and processing protocols leading to its assessment. For example, Boellaard et al [51] have shown in a simulation study that differences in defining regions of interest can result in a difference of $>50\%$ of measured absolute SUV_{max} . In a clinical study, Stahl et al [52] found that the SUV_{max} of gastric carcinomas increased by 60% between 40 and 90 min post injection, indicating that absolute tumour SUV_{max} are highly dependent on the timing of the data acquisition. Fortunately, SUV_{max} ratios measured by different methods varied only minimally [51]. Therefore, relative changes represented more robust parameters and were preferable for establishing the response criteria that could be used at multiple institutions, e.g. in multicentre trials. Thirdly, the prognostic value was evidenced by the broad range of SUV threshold values that have been used in the literature to distinguish between patients with low and high survival (thresholds varying from 35% to 63%). Remarkably, all studies with the prognostic effect of decreased SUV_{max} from PET1 to PET2 for OS selected $>35\%$ as the SUV threshold definition [30, 33, 35, 37]. Indeed, a meta-analysis of the individual patient data (IPD) would be needed to compensate for the large heterogeneity of the reported SUV. Summary statistics meta-analyses have the advantage of including published studies which are immediately available for analysis and whose results can be checked by others. Although IPD meta-analysis results are usually similar to previous literature-based

publications, they add some interest, such as including unpublished trials, updating results, and particularly allowing for multivariate analyses, adjusting for other variables, and subgroup analyses.

Conclusion

Although this study had some drawbacks, including being restricted to articles published in English, most retrospective studies, HRs extrapolated from the survival curves and so on, our experience has been that ^{18}F -PDG-PET scanning is useful as a tool to help guide therapy for patients with this difficult and morbid AEG carcinoma. We also found that the change in SUV_{max} on ^{18}F -PDG-PET scans taken from PET1 to PET2 may identify patients with a better prognosis after surgery and may someday be able to select patients who would benefit more from additional chemotherapeutic approaches than from resection.

References

- Chin L, Gray JW. Translating insights from the cancer genome into clinical practice. *Nature* 2008;452:553–63.
- Sawyers CL. The cancer biomarker problem. *Nature* 2008;452:548–52.
- van't Veer LJ, Bernards R. Enabling personalized cancer medicine through analysis of gene-expression patterns. *Nature* 2008;452:564–70.
- Luthra MG, Ajani JA, Izzo J, Ensor J, Wu TT, Rashid A, et al. Decreased expression of gene cluster at chromosome 1q21 defines molecular subgroups of chemoradiotherapy response in esophageal cancers. *Clin Cancer Res* 2007;13:912–19.
- Luthra R, Wu TT, Luthra MG, Izzo J, Lopez-Alvarez E, Zhang L, et al. Gene expression profiling of localized esophageal carcinomas: association with pathologic response to preoperative chemoradiation. *J Clin Oncol* 2006;24:259–67.
- Hildebrandt MA, Yang H, Hung MC, Izzo JG, Huang M, Lin J, et al. Genetic variations in the PI3K/PTEN/AKT/mTOR pathway are associated with clinical outcomes in esophageal cancer patients treated with chemoradiotherapy. *J Clin Oncol* 2009;27:857–71.
- Wu X, Gu J, Wu TT, Swisher SG, Liao Z, Correa AM, et al. Genetic variations in radiation and chemotherapy drug

- action pathways predict clinical outcomes in esophageal cancer. *J Clin Oncol* 2006;24:3789–98.
8. Cunningham D, Allum WH, Stenning SP, Thompson JN, Van de Velde CJ, Nicolson M, et al. Perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. *N Engl J Med* 2006;355:11–20.
 9. Medical Research Council Oesophageal Cancer Working Group. Surgical resection with or without preoperative chemotherapy in oesophageal cancer: a randomised controlled trial. *Lancet* 2002;359:1727–33.
 10. Malthaner RA, Collin S, Fenlon D. Preoperative chemotherapy for resectable thoracic esophageal cancer. *Cochrane Database Syst Rev* 2006;3:CD001556.
 11. Kelsen DP, Ginsberg R, Pajak TF, Sheahan DG, Gunderson L, Mortimer J, et al. Chemotherapy followed by surgery compared with surgery alone for localized esophageal cancer. *N Engl J Med* 1998;339:1979–84.
 12. Berghmans T, Dusart R, Paesmans M, Hossein-Foucher C, Buvat I, Castaigne C, et al. Primary tumor standardized uptake value (SUVmax) measured on fluorodeoxyglucose positron emission tomography (PDG-PET) is of prognostic value for survival in non-small cell lung cancer (NSCLC): a systematic review and meta-analysis (MA) by the European Lung Cancer Working Party for the IASLC Lung Cancer Staging Project. *J Thorac Oncol* 2008;3:6–12.
 13. Pan L, Gu P, Huang G, Xue H, Wu S. Prognostic significance of SUV on PET/CT in patients with esophageal cancer: a systematic review and meta-analysis. *Eur J Gastroenterol Hepatol* 2009;21:1008–15.
 14. Parmar MK, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. *Stat Med* 1998;17:2815–34.
 15. Levine EA, Farmer MR, Clark P, Mishra G, Ho C, Geisinger KR, et al. Predictive value of 18-fluoro-deoxy-glucose-positron emission tomography (18F-PDG-PET) in the identification of responders to chemoradiation therapy for the treatment of locally advanced esophageal cancer. *Ann Surg* 2006;243:472–8.
 16. Swisher SG, Erasmus J, Maish M, Correa AM, Macapinlac H, Ajani JA, et al. 2-Fluoro-2-deoxy-D-glucose positron emission tomography imaging is predictive of pathologic response and survival after preoperative chemoradiation in patients with esophageal carcinoma. *Cancer* 2004;101:1776–85.
 17. Hong D, Lunagomez S, Kim EE, Lee JH, Bresalier RS, Swisher SG, et al. Value of baseline positron emission tomography for predicting overall survival in patient with nonmetastatic esophageal or gastroesophageal junction carcinoma. *Cancer* 2005;104:1620–6.
 18. Smithers BM, Couper GC, Thomas JM, Wong D, Gotley DC, Martin I, et al. Positron emission tomography and pathological evidence of response to neoadjuvant therapy in adenocarcinoma of the esophagus. *Dis Esophagus* 2008;21:151–8.
 19. Swisher SG, Hofstetter W, Komaki R, Correa AM, Erasmus J, Lee JH, et al. Improved long-term outcome with chemoradiotherapy strategies in esophageal cancer. *Ann Thorac Surg* 2010;90:892–8.
 20. Tian J, Chen L, Wei B, Shao M, Ding Y, Yin D, et al. The value of vesicant 18F-fluorodeoxyglucose positron emission tomography (18F-FDG PET) in gastric malignancies. *Nucl Med Commun* 2004;25:825–31.
 21. Flamen P, Van Cutsem E, Lerut A, Cambier JP, Haustermans K, Bormans G, et al. Positron emission tomography for assessment of the response to induction radiochemotherapy in locally advanced oesophageal cancer. *Ann Oncol* 2002;13:361–8.
 22. Kroep JR, Van Groenigen CJ, Cuesta MA, Craanen ME, Hoekstra OS, Comans EF, et al. Positron emission tomography using 2-deoxy-2-[18F]-fluoro-D-glucose for response monitoring in locally advanced gastroesophageal cancer; a comparison of different analytical methods. *Mol Imaging Biol* 2003;5:337–46.
 23. Haley M, Konski A, Li T, Cheng JD, Maurer A, Haluszka O, et al. Influence of diabetes on the interpretation of PET scans in patients with esophageal cancer. *Gastrointest Cancer Res* 2009;3:149–52.
 24. Stahl A, Ott K, Weber WA, Wiedner H, Fink U, Schwaiger M, et al. FDG PET imaging of locally advanced gastric carcinomas: correlation with endoscopic and histopathological findings. *Eur J Nucl Med Mol Imaging* 2003;30:288–95.
 25. Swisher SG, Maish M, Erasmus JJ, Correa AM, Ajani JA, Bresalier R, et al. Utility of PET, CT, and EUS to identify pathologic responders in esophageal cancer. *Ann Thorac Surg* 2004;78:1152–60.
 26. Mochiki E, Kuwano H, Katoh H, Asao T, Oriuchi N, Endo K. Evaluation of 18F-2-deoxy-2-fluoro-D-glucose positron emission tomography for gastric cancer. *World J Surg* 2004;28:247–53.
 27. Rizk NP, Tang L, Adusumilli PS, Bains MS, Akhurst TJ, Ilson D, et al. Predictive value of initial PET-SUVmax in patients with locally advanced esophageal and gastroesophageal junction adenocarcinoma. *J Thorac Oncol* 2009;4:875–9.
 28. Rizk N, Downey RJ, Akhurst T, Gonen M, Bains MS, Larson S, et al. Preoperative 18[F]-fluorodeoxyglucose positron emission tomography standardized uptake values predict survival after esophageal adenocarcinoma resection. *Ann Thorac Surg* 2006;81:1076–81.
 29. Murthy SB, Patnana SV, Xiao L, Rohren E, Hofstetter WL, Swisher SG, et al. The standardized uptake value of 18-fluorodeoxyglucose positron emission tomography after chemoradiation and clinical outcome in patients with localized gastroesophageal carcinoma. *Oncology* 2010;78:316–22.
 30. Lordick F, Ott K, Krause BJ, Weber WA, Becker K, Stein HJ, et al. PET to assess early metabolic response and to guide treatment of adenocarcinoma of the oesophagogastric junction: the MUNICON phase II trial. *Lancet Oncol* 2007;8:797–805.
 31. Javeri H, Xiao L, Rohren E, Lee JH, Liao Z, Hofstetter W, et al. The higher the decrease in the standardized uptake value of positron emission tomography after chemoradiation, the better the survival of patients with gastroesophageal adenocarcinoma. *Cancer* 2009;115:5184–92.
 32. Javeri H, Xiao L, Rohren E, Komaki R, Hofstetter W, Lee JH, et al. Influence of the baseline 18F-fluoro-2-deoxy-D-glucose positron emission tomography results on survival and pathologic response in patients with gastroesophageal cancer undergoing chemoradiation. *Cancer* 2009;115:624–30.
 33. Wiedner HA, Ott K, Lordick F, Becker K, Stahl A, Herrmann K, et al. Prediction of tumor response by PDG-PET: comparison of the accuracy of single and sequential studies in patients with adenocarcinomas of the esophagogastric junction. *Eur J Nucl Med Mol Imaging* 2007;34:1925–32.
 34. Smith JW, Moreira J, Abood G, Aranha GV, Nagda S, Wagner RH, et al. The influence of (18)fluorodeoxyglucose positron emission tomography on the management of gastroesophageal junction carcinoma. *Am J Surg* 2009;197:308–12.
 35. Ott K, Weber WA, Lordick F, Hulshof MC, Hoekstra OS, Herrmann K, et al. Metabolic imaging predicts response, survival, and recurrence in adenocarcinomas of the esophagogastric junction. *J Clin Oncol* 2006;24:4692–8.
 36. Patnana SV, Murthy SB, Xiao L, Rohren E, Hofstetter WL, Swisher SG, et al. Critical role of surgery in patients with gastroesophageal carcinoma with a poor prognosis after chemoradiation as defined by positron emission tomography. *Cancer* 2010;116:4487–94.

37. Weber WA, Ott K, Becker K, Dittler HJ, Helmberger H, Avril NE, et al. Prediction of response to preoperative chemotherapy in adenocarcinomas of the esophagogastric junction by metabolic imaging. *J Clin Oncol* 2001;19:3058–65.
38. Roedl JB, Colen RR, Holalkere NS, Fischman AJ, Choi NC, Blake MA. Adenocarcinomas of the esophagus: response to chemoradiotherapy is associated with decrease of metabolic tumor volume as measured on PET-CT. Comparison to histopathologic and clinical response evaluation. *Radiother Oncol* 2008;89:278–86.
39. Pohl H, Welch HG. The role of overdiagnosis and reclassification in the marked increase of esophageal adenocarcinoma incidence. *J Natl Cancer Inst* 2005;97:142–6.
40. Jemal A, Siegel R, Xu J, Ward E. Cancer statistics, 2010. *CA Cancer J Clin* 2010;60:277–300.
41. Mariette C, Piessen G, Triboulet JP. Therapeutic strategies in oesophageal carcinoma: role of surgery and other modalities. *Lancet Oncol* 2007;8:545–53.
42. Juweid ME, Cheson BD. Positron-emission tomography and assessment of cancer therapy. *N Engl J Med* 2006;354:496–507.
43. Langer R, Specht K, Becker K, Ewald P, Bekesch M, Sarbia M, et al. Association of pretherapeutic expression of chemotherapy-related genes with response to neoadjuvant chemotherapy in Barrett carcinoma. *Clin Cancer Res* 2005;11:7462–9.
44. Zhu WQ, Yu JM, Sun XD, Xie P, Kong L. Serum CYFRA21-1 as a prognostic marker for patients with undifferentiated nasopharyngeal carcinoma. *Biomarkers* 2010;15:602–7.
45. Makdissi FB, Machado LV, Oliveira AG, Benvenuti TT, Katayama ML, Brentani MM, et al. Expression of E-cadherin, Snail and Hakai in epithelial cells isolated from the primary tumor and from peritumoral tissue of invasive ductal breast carcinomas. *Braz J Med Biol Res* 2009;42:1128–37.
46. Brucher BL, Weber W, Bauer M, Fink U, Avril N, Stein HJ, et al. Neoadjuvant therapy of esophageal squamous cell carcinoma: response evaluation by positron emission tomography. *Ann Surg* 2001;233:300–9.
47. Downey RJ, Akhurst T, Ilson D, Ginsberg R, Bains MS, Gonen M, et al. Whole body 18FDG-PET and the response of esophageal cancer to induction therapy: results of a prospective trial. *J Clin Oncol* 2003;21:428–32.
48. Wieder HA, Brucher BL, Zimmermann F, Becker K, Lordick F, Beer A, et al. Time course of tumor metabolic activity during chemoradiotherapy of esophageal squamous cell carcinoma and response to treatment. *J Clin Oncol* 2004;22:900–8.
49. Gillham CM, Lucey JA, Keogan M, Duffy GJ, Malik V, Raouf AA, et al. (18)FDG uptake during induction chemoradiation for oesophageal cancer fails to predict histomorphological tumour response. *Br J Cancer* 2006;95:1174–9.
50. Westerterp M, Omloo JM, Sloof GW, Hulshof MC, Hoekstra OS, Crezee H, et al. Monitoring of response to pre-operative chemoradiation in combination with hyperthermia in oesophageal cancer by PDG-PET. *Int J Hyperthermia* 2006;22:149–60.
51. Boellaard R, Krak NC, Hoekstra OS, Lammertsma AA. Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values: a simulation study. *J Nucl Med* 2004;45:1519–27.
52. Stahl A, Ott K, Schwaiger M, Weber WA. Comparison of different SUV-based methods for monitoring cytotoxic therapy with FDG PET. *Eur J Nucl Med Mol Imaging* 2004;31:1471–8.

Appendix A

The quality scale used in this study

Except when specified, the attributed value per item is 2 points if it is clearly defined in the article, 1 point if its description is incomplete or unclear, and 0 points if it is not defined or is inadequate.

Scientific design

(1) Study objective definition. (2) Study design: prospective, 2 points; retrospective or retrolective, 1 point; not defined, 0 points. (3) Outcome definition. (4) Statistical considerations: fully reported with a preliminary assessment of the patient/sample number to be included and/or analysed, 2 points; patient/sample number to be included and/or analysed justified by the number of studied variables (minimum 10 patients per variable), 1 point; not defined, 0 points. (5) Statistical methods and tests description.

Generalisability

(1) Patient selection criteria, including histological type, disease stage and treatment. (2) Patients' characteristics, including histology type, disease stage and treatment. (3) Initial workup. (4) Treatment description. (5) Number of ineligible patients with exclusion causes.

Results analysis

(1) Follow-up description, including the number of events. (2) Survival analysis according to the SUV. (3) Univariate analysis of the prognostic factors for survival: report of the relative risk with the CI, 2 points; results without evaluation of the relative risk and its CI, 1 point; not reported or inadequate, 0 points. (4) Multivariate analysis of the prognostic factors for survival: report the relative risk with the CI, 2 points; results without evaluation of the relative risk and its CI, 1 point; not reported or inadequate, 0 points.

The PET reports

(1) Patients characteristics: weight/height; blood sugar level; histological subtype. (2) ¹⁸F-FDG-PET acquisition protocol characteristics: fasting duration; injected dose of ¹⁸F-FDG; delay between injection and data acquisition. (3) Technical parameters: investigation area; delay between CT thorax and PET acquisition; SUV formula; type of SUV; type of PET engine; duration of emission time; duration of transmission time; attenuation; and reconstruction parameters. (4) The analysis of the relationship between SUV was performed without knowledge of survival results and conversely (double blind). (5) SUV cut-off definition.