The prognostic value of TP53 mutations in oesophageal adenocarcinoma: a systematic review and meta-analysis

Oliver M Fisher,1 Sarah J Lord,1,2,3 Dan Falkenback,1,4 Nicholas J Clemons,5,6 Guy D Eslick,7 Reginald V Lord1,8

ABSTRACT
Objective To clarify the prognostic role of tumour protein 53 (TP53) mutations in patients with oesophageal adenocarcinoma (OAC) as there is a need for biomarkers that assist in guiding management for patients with OAC.

Design A systematic review was conducted using MEDLINE, Embase, PubMed and Current Contents Connect to identify studies published between January 1990 and February 2015 of oesophageal cancer populations (with OAC diagnoses >50% of cases) that measured tumoural TP53 status and reported hazard ratios (HR), or adequate data for estimation of HR for survival for TP53-defined subgroups. Risk of bias for HR estimates was assessed using prespecified criteria for the appraisal of relevant domains as defined by the Cochrane Prognosis Methods Group including adherence to Grading of Recommendations, Assessment, Development and Evaluation and Reporting recommendations for tumor MARKer prognostic studies guidelines, as well as assay method used (direct TP53 mutation assessment vs immunohistochemistry) and adjustment for standard prognostic factors. A pooled HR and 95% CI were calculated using a random-effects model.

Results Sixteen eligible studies (11 with OAC only and 5 mixed histology cohorts) including 888 patients were identified. TP53 mutations were associated with reduced survival (HR 1.48, 95% CI 1.16 to 1.90, I²=33%). A greater prognostic effect was observed in a sensitivity analysis of those studies that reported survival for OAC-only cohorts and were assessed at low risk of bias (HR 2.11, 95% CI 1.35 to 3.31, I²=0%).

Conclusions Patients with OAC and TP53 gene mutations have reduced overall survival compared with patients without these mutations, and this effect is independent of tumour stage.

INTRODUCTION
The incidence of oesophageal adenocarcinoma (OAC) has increased faster than any other cancer since the 1970s in many Western countries with highest incidence rates found in Northern and Western Europe, Northern America and Oceania,1–4 with a greater than sixfold increase during the past three decades.1,5 Population-based studies have observed consistent changes across different age and tumour stage groups, indicating that the rise of OAC incidence is a true increase and not an artefact of enhanced surveillance programmes.1,6 Possible explanations for this increase include the increasing prevalence of obesity and Barrett’s oesophagus (BO),7,8 well-recognised risk factors for OAC.

Fewer than half of the patients with a new diagnosis of OAC are eligible for curative treatment, and OAC continues to have one of the highest cancer case-fatality rates with population-based 5-year survival rates typically around 15%.9,9 For those patients referred to curative treatment,
usually by neoadjuvant chemotherapy or chemoradiotherapy (CRTx) followed by oesophagectomy, 5-year survival rates are still generally <45%.10 In contrast to other malignancies, such as breast and colon cancer, where the incorporation of molecular information has become part of routine practice for therapeutic stratification,11 current treatment algorithms for OAC still depend on only imaging and histological assessments to determine disease stage and grade to guide treatment and help classify prognosis.

The tumour-suppressor gene tumour protein 53 (TP53) (National Centre for Biotechnology Information gene ID: 7157), which encodes the p53 protein and is sometimes called ‘the guardian of the genome’,12 is one of the most frequently mutated and studied genes in human cancers.13 The p53 protein plays multiple functions in regulating cell cycle progression and apoptosis, autophagy, differentiation and senescence, as well as DNA repair functions, and also exerts effects on cell metabolic pathways and cytokines.14 Substantial efforts have been made to study the effect of TP53 mutations on prognosis for patients with cancer. Furthermore, as most chemotherapeutic agents act by inducing DNA damage,15–17 the predictive effect of TP53 gene mutations on therapeutic response has also been explored.18

Recent large-scale whole-genome and whole-exome sequencing studies have shown that both OAC and dysplastic BO harbour a very high TP53 mutation rate of up to 70%,19–21 indicating a central role for this gene in OAC pathogenesis. This finding raises the question of whether the 30% of patients harbouring wild-type TP53 may thus have a different underlying tumour biology that may impact patient outcome.

We aimed to resolve the existing uncertainty regarding the prognostic value of TP53 for staging OAC by conducting a systematic review and meta-analysis of all published data with subgroup analysis of studies assessed as low risk of bias, and studies using direct TP53 gene mutation analysis techniques to determine TP53 mutation status, since these are the most accurate methods for determining tumoural TP53 mutations.

MATERIALS AND METHODS
Search strategy and study eligibility criteria
The electronic bibliographic databases MEDLINE, Embase, PubMed and Current Contents Connect were searched to identify eligible studies published between January 1990 and 8 February 2015 using MeSH terms and text words for adenocarcinoma*, oesophagus* or oesophageal*, TP53* or p53* or 17p13* or 17p*. In an attempt to minimize the risk of publication bias, conference abstracts and proceedings were searched through Web of Science, Embase and Scopus using the terms o/esophagus* and p53*. Further, the following major GI and oncological conferences were searched for relevant reports: Digestive Disease Week (DDW; by screening the DDW website and supplementary material of the journal Journal of Gastroenterology), American Society for Clinical Oncology (ASCO; by searching the ASCO library and Journal of Clinical Oncology supplementary material) and American Association for Cancer Research (AACR; searched through their webportal of all AACR conference proceedings) from 1990 to 2015.

Two reviewers (OMF and DF) scanned the search results (title and abstract) and retrieved full text publications using the criteria outlined below to identify eligible studies. Reference lists of relevant studies identified from the search including reviews were further screened to identify studies that may not have been identified by the strategy outlined above.

Study inclusion criteria were prospective or retrospective clinical studies of OAC populations that assessed TP53 mutation status and/or p53 expression in primary tumours, and compared overall survival for TP53 mutation versus TP53 non-mutation subgroups with calculation of HR and 95% CIs, or reported adequate data for their estimation. To include all available data, we also included studies of oesophageal cancer cohorts that included patients with squamous cell carcinoma if at least ≥50% of the patient cohort had a diagnosis of OAC.

Study exclusion criteria were studies of TP53 DNA germline mutations or autoantibody detection in blood; and reports available in abstract form only that did not report adequate information to determine study eligibility or to assess study methods for risk of bias.

If studies did not report sufficient data to calculate HRs or in case of missing/unclear data, the corresponding author was contacted by email to request this information. If the same research unit (identified from author names and institution) published multiple reports with overlapping patient recruitment time periods, HR estimates were extracted from the most recent publication with the largest patient numbers to avoid duplication of data.

Data extraction
Three investigators (OMF, DF and NJC) reviewed eligible studies and extracted the following variables into a standardised data extraction form: author’s name; publication year; country where study was conducted; number of patients included and general patient demographics; tumour histology (number and proportion of OAC tumours included); treatment modality (surgery alone, neoadjuvant or adjuvant CRTx followed by surgery); tissue specimen type (surgical specimen vs endoscopic biopsy); TP53 assay methods (TP53 gene sequencing, single-strand confirmation polymorphism (SSCP), immunohistochemistry (IHC), and type of antibody, dilution for IHC); criteria or cut-point used to define TP53 mutation status for the survival analysis; study prevalence of TP53 ‘mutation’ and ‘non-mutation’ subgroups; median survival of all patients and by TP53 mutation status; unadjusted and adjusted HR with 95% CI and corresponding p values where available. For consistency and to facilitate further quantitative analyses, the authors’ definitions for TP53 ‘mutations’ were used for studies performing only IHC as the respective studies did not use uniform staining classification criteria. As such, nuclear p53 protein overexpression was interpreted to represent TP53 mutations by all authors of the included studies, although loss of p53 protein expression has also been associated with tumoural TP53 gene mutations.22 23 This staining pattern was not reported and/or interpreted in such a manner in any of the included studies.

Risk of bias assessment, subgroup and sensitivity analyses
All studies were assessed for risk of bias for the study estimate of the impact of TP53 on survival by appraising six domains (study participation, biomarker measurement, outcome measurement, confounding measurement and account, participant attrition, analysis method) using prespecified criteria adapted from the Grading of Recommendations, Assessment, Development and Evaluation (http://www.gradeworkinggroup.org),24 REporting recommendations for tumor MARker prognostic studies (REMARK)25 and from Hayden et al’s26 (Cochrane Prognosis Methods Group) guidelines for quality appraisal for prognostic studies. Risk of bias for each domain was graded as high, low or unclear based on assessment of each criterion. The overall risk of bias for the study was assessed as high if one or
more of the domains was assessed as high risk of bias as recom-
mended by the Cochrane Collaboration.27

Assessment of the risk of bias for different methods of assess-
ing TP53 mutation status was informed by data regarding the
analytical validity of different methods reported in the Inter-
national Agency for Research on Cancer (IARC) TP53
mutation database (R17)13 (as summarised in online supplement-
ary file 1). Studies performing TP53 gene sequencing or direct
assessment of TP53 gene mutations were assessed as being at
low risk of biomarker measurement bias compared with studies
performing only IHC analysis, based on estimates from the
IARC TP53 mutation database that approximately 27% of all
TP53 mutations stain as false negatives in OAC using IHC (see
online supplementary figure S2). Studies that did not adjust for
tumour stage to assess the independent impact of TP53 muta-
tion status on patient survival were classified as high risk of bias.

Statistical analysis
To estimate the effect of TP53 mutation status on OAC survival,
we calculated a pooled HR and 95% CI using the generic
inverse variance method. If the HR was not reported, it was
estimated from the corresponding Kaplan–Meier curves using the
Parmar method.38 28 If the SE was not reported, it was esti-
mated from the 95% CI.

Because different TP53 mutation analysis methods were used
across studies, we expected heterogeneity in study estimates of
the TP53 mutation effect on survival, and thus we applied a
random-effects model to estimate the HR.29 Heterogeneity was
tested using Cochran’s Q statistic, with p<0.1 indicating hetero-
genicity. The degree of heterogeneity was quantified using the I2
statistic.31 Meta-regression was performed to inspect possible
sources of intersstudy heterogeneity.32 Study level factors that
may modify the prognostic effect of TP53 were included as cov-
ariates if they were present in ≥10 of the included studies.27

Sensitivity analysis was performed to assess the impact of
tumour histology and assay type on survival by repeating the
pooled HR analysis in the following subgroups: (i) OAC-only
versus OAC-mixed study populations; (ii) IHC versus direct
TP53 gene mutation analyses; and (iii) studies assessed as having
low versus high risk of bias. Differences between subgroups
were assessed with a test for interaction.33 In order to estimate
the prognostic value of TP53 independent of stage, a sensitivity
analysis was also performed of studies reporting HR adjusted by
stage. Publication bias was quantified using the Egger’s regres-
sion model and visualised using funnel plot analyses.34

Descriptive statistics as well as quantitative analysis of the
IARC TP53 mutation database to guide risk of bias assessment
were performed using R statistical software35 and the
metafor package.36 Meta-analyses of HR estimates were performed
in the following subgroups: (i) OAC-only cancers and had complete survival information on 22 patients, all with OAC cancers. One of the patients in this study was excluded from the survival analysis due to an early, post-operative mortality.42 The other study,44 included survival data for all 60 patients with OAC.

Four studies41–43 54 were assessed as being at a low risk of
bias, and 1219 40 44–53 studies were assessed as being at high risk of
bias (table 2). Funnel plot analyses did not reveal substantial
publication bias (see online supplementary figure S3).

Overall analyses
The meta-analysis of data from all 16 included studies showed that
TP53 mutation is associated with a statistically significant
effect on patient overall survival with an HR 1.48
(95% CI 1.16 to 1.90, p=0.002, n=888 patients; figure 2) with
low-moderate heterogeneity across studies that was not statistically
significant (I²=33%, p heterogeneity=0.1). The analysis of
studies including pure OAC patient cohorts showed similar
results with low heterogeneity (HR 1.46, 95% CI 1.17 to 1.83, p=0.0009, n=11 studies and 644 patients, I²=0%, p for heter-
genocity=0.53, p for interaction=0.78; figure 3).

Subgroup and sensitivity analyses
The effect of TP53 mutation status on survival appeared to be
smaller among studies performing IHC (pooled HR 1.28, 95%
CI 0.95 to 1.73, p=0.10, 8 studies, 417 patients, I²=0%) com-
pared with studies performing direct TP53 gene assessments
sequencing and SSCP) or LOH analyses (HR 1.68, 95% CI
1.14 to 2.47, p=0.009, 8 studies, 471 patients, I²=50%,
figure 4A). However, this difference was not statistically signi-
cant (p for interaction=0.28). This finding was similar in
studies including OAC only cohorts (figure 4B).

The effect of mutant TP53 on patient overall survival was
larger in studies that had adjusted their analyses for tumour

RESULTS

Study characteristics
The search strategy yielded 323 studies, of which 16 met our
eligibility criteria (figure 1). Study characteristics are sum-
murised in table 1. Eleven studies included pure OAC cohorts39–
49 and five studies included mixed histological cohorts50–54 in
which the percentage of OAC cases in the study data ranged
from 56%50 to 81%.53 Overall, the 16 studies totalled 1211
patients including 986 OAC, with survival data reported for
888 patients. The number of patients with survival data in each
study ranged from 16 to 142 (median 50) with a median
number of OAC tumours of 53 (range 20–142) per study.

Half of the studies (n=8) assessed TP53 mutation status by
IHC,39 40 44–46 48 51 53 one study assessed TP53 mutations
through 17p/17p.13 loss of heterozygosity (LOH).41 one
through SSCP39 and the remaining six studies,42 43 45 50 52 54
performed TP53 gene sequencing to determine the presence of
mutations. Using these methods, a median of 53% (range 33–79%)
of all tumours and 57% of all OACs (range 33–79%) were
classified as harbouring TP53 mutations.

The clinicopathological variables and survival times reported
in the included studies are summarised in online supplementary
table S1. Briefly, 10 studies provided information on pathological
T-stage,39 42 44–47 50 51 53 54 13 studies on pathological
N-stage,39 40 42 44–45 51 53 54 8 studies on metastatic
status42 44 46 47 49 50 53 54 and 10 described the final tumour
differentiation grade.39 42 43 45 47–50 52 54 Approximately half
the studies reported an overall staging of patients according to
Union for International Cancer Control (UICC) or American Joint
Committee on Cancer (AJCC) criteria (n=9),39 41 43–47 50 52 54
and 5 studies described resection margin status.41 43 45 51 54 In the 14
studies39–42 45–51 53 54 that reported survival time data for
biomarker-defined patient subgroups, the median survival time
for patients assessed as having mutated TP53 was 18.9 months
(n=488) compared with 26.2 months for patients with non-
mutated TP53 (n=423).

HRs were reported in nine studies and extrapolated from five
studies. In addition, individual patient data were available for
two studies to calculate tumour stage-adjusted HR and 95% CI
rs42 44 (table 1). One of these studies42 included six subcardia
cancers and had complete survival information on 22 patients,
all with OAC cancers. One of the patients in this study was
excluded from the survival analysis due to an early, post-
operative mortality.42 The other study,44 included survival data
for all 60 patients with OAC.

Four studies41–43 54 were assessed as being at a low risk of
bias, and 1219 40 44–53 studies were assessed as being at high risk of
bias (table 2). Funnel plot analyses did not reveal substantial
publication bias (see online supplementary figure S3).

Overall analyses
The meta-analysis of data from all 16 included studies showed
that TP53 mutation is associated with a statistically significant
negative effect on patient overall survival with an HR 1.48
(95% CI 1.16 to 1.90, p=0.002, n=888 patients; figure 2) with
low-moderate heterogeneity across studies that was not statistically
significant (I²=33%, p heterogeneity=0.1). The analysis of
studies including pure OAC patient cohorts showed similar
results with low heterogeneity (HR 1.46, 95% CI 1.17 to 1.83, p=0.0009, n=11 studies and 644 patients, I²=0%, p for heter-
genocity=0.53, p for interaction=0.78; figure 3).

Subgroup and sensitivity analyses
The effect of TP53 mutation status on survival appeared to be
smaller among studies performing IHC (pooled HR 1.28, 95%
CI 0.95 to 1.73, p=0.10, 8 studies, 417 patients, I²=0%) com-
pared with studies performing direct TP53 gene assessments
sequencing and SSCP) or LOH analyses (HR 1.68, 95% CI
1.14 to 2.47, p=0.009, 8 studies, 471 patients, I²=50%,
figure 4A). However, this difference was not statistically signi-
cant (p for interaction=0.28). This finding was similar in
studies including OAC only cohorts (figure 4B).

The effect of mutant TP53 on patient overall survival was
larger in studies that had adjusted their analyses for tumour

Fisher OM, et al. Gut 2017;66:399–410. doi:10.1136/gutjnl-2015-310888

401

Gut: first published as 10.1136/gutjnl-2015-310888 on 5 January 2016. Downloaded from http://gut.bmj.com/ on December 30, 2023 by guest. Protected by copyright.
stage (HR 1.95, 95% CI 1.41 to 2.66, \( p \leq 0.0001 \), 7 studies, 430 patients, \( I^2=0\% \); see online supplementary figure S4) compared with the estimates from studies that reported unadjusted risk estimates (HR 1.22, 95% CI 0.88 to 1.70, \( p=0.24 \), 9 studies, 458 patients, \( I^2=38\% \)). This difference was borderline statistically significant (\( p \) for interaction=0.05). A similar effect was seen in the subset of studies containing pure OAC cohorts (see online supplementary figure S4).

The prognostic effect of TP53 mutations was also significantly larger in the subset of four studies assessed as low risk of bias (HR 2.29, 95% CI 1.50 to 3.48, \( p=0.0001 \), 197 patients, \( I^2=0\% \); figure 5A), compared with those assessed as high risk of bias (HR 1.29, 95% CI 0.98 to 1.70, \( p=0.07 \), 691 patients, \( I^2=30\% \), \( p \) for interaction=0.03). This effect size was similar in the three studies with low risk of bias that contained pure OAC cohorts (HR 2.11, 95% CI 1.35 to 3.31, \( p=0.001 \), \( n=161 \) patients, \( I^2=0\% \); figure 5B).

Exploratory subgroup and sensitivity analyses

Subgroup analysis of studies that determined TP53 status using gene sequencing also showed a negative prognostic impact of mutant TP53 on patient survival (HR 1.80, 95% CI 1.50 to 3.08, \( p=0.002 \), 6 studies, 330 patients, \( I^2=62\% \); see online supplementary figure S5A), compared with those assessed as high risk of bias (HR 1.29, 95% CI 0.98 to 1.70, \( p=0.07 \), 691 patients, \( I^2=30\% \), \( p \) for interaction=0.03). This effect size was similar in the three studies with low risk of bias that contained pure OAC cohorts (HR 2.11, 95% CI 1.35 to 3.31, \( p=0.001 \), \( n=161 \) patients, \( I^2=0\% \); figure 5B).

Findings were consistent when the same subgroup analysis was performed for only those TP53 gene sequencing studies including pure OAC cohorts (HR 1.76, 95% CI 1.26 to 2.47; \( p=0.0009 \), 3 studies, 213 patients, \( I^2=0\% \); see online supplementary figure S5C). A summary of the results of other exploratory subgroup analyses can be found in online supplementary table S2.

Metaregression: potential sources of interstudy heterogeneity

Of the 14 inspected study covariates, only the adjustment of HR estimates for standard prognostic variables (change log HR 0.46, 95% CI 0.01 to 0.91; \( p=0.04 \)) and the study appraised as being at low risk of bias (change log HR 0.58, 95% CI 0.06 to 1.10; \( p=0.03 \)) were significant sources of heterogeneity (see online supplementary figure S6).

DISCUSSION

This study indicates that mutated TP53 negatively impacts overall survival in patients with OAC, independent of tumour stage. This effect is estimated as a relative increase in hazard of death of 48%, with up to a 211% increase when studies of mixed histology cohorts or high risk of bias are excluded. This corresponds to a reduced survival time of approximately 7 months based on median survival from the included studies. This effect size is similar to the difference in median survival of stage IIIA compared with stage IIIC OAC.55

To our knowledge, this is the first comprehensive systematic review and meta-analysis of the prognostic impact of TP53 mutations in patients with OAC that includes an assessment independent of tumour stage. Two previous systematic reviews discounted the association due to inconclusive evidence either due to small study size or non-inclusion of tumour stage. However, to our knowledge, this is the first study to use a systematic and rigorous method to assess the impact of TP53 mutations independent of tumour stage. This study uses a comprehensive set of covariates to account for potential sources of heterogeneity, which are significantly lower in studies of pure OAC cohorts compared with studies of mixed histology cohorts or those at high risk of bias. The study validates the impact of TP53 mutations with a larger cohort size and robust statistical analysis that is independent of tumour stage. This study is the first to evaluate the prognostic impact of TP53 mutations in patients with OAC for a comprehensive range of covariates and extend the evidence of TP53 as a biomarker for patient survival. The findings of this study are consistent with previous studies that reported a negative prognostic impact of TP53 mutations in patients with OAC.52,53,54,55

However, the biological mechanism of the prognostic impact of TP53 mutations in patients with OAC is unclear. It is possible that TP53 mutations may affect the biological behaviour of the tumour and its response to treatment. Further studies are required to explore the underlying mechanisms of the association of TP53 mutations with patient survival in patients with OAC.
Table 1 Baseline characteristics of included studies

| Author       | Year | Country | N in survival analysis | N in survival analysis (%) | Number of patients with OAC (% total) | Only curative surgery in survival analysis? | Specimen type | CRTx? | Specimen type | CRTx? | Analysis method | IHC antibody | Dilution | Percent 'mutated' | HR estimation method | Multivariable analysis performed and reported in original paper? |
|--------------|------|---------|------------------------|-----------------------------|--------------------------------------|---------------------------------------------|--------------|-------|--------------|-------|----------------|-------------|----------|-----------------|--------------------------------|---------------------------------------------------------------|
| Fléjou      | 1993 | France  | 62                     | 62 (100)                    | Yes, only curative surgery in survival analysis? | Surgery, NR, NR | DO7 (Dako) and PAb1801 (Oncogene Science) | NR | 66 | Extrapolated | No |
| Duhanlongsd  | 1995 | USA     | 42                     | 42 (100)                    | Yes, only curative surgery in survival analysis? | Surgery, Yes, Yes | PAb1801 (Oncogene Science) | NR | 79 | Extrapolated | No |
| Sauter      | 1995 | USA     | 24                     | 24 (100)                    | Yes, only curative surgery in survival analysis? | Biopsies, Yes, Yes | PAb1801 (Oncogene Science) | 1 μg/mL (7) | 50 | Extrapolated | No |
| Wu          | 1998 | USA     | 92                     | 92 (100)                    | Yes, only curative surgery in survival analysis? | Surgery, Yes, Yes | LOH+IHC, DO7 (Dako) | NR | 57 | Reported in text | Yes |
| Ribeiro     | 1998 | USA     | 42                     | 31 (74)                     | Yes, only curative surgery in survival analysis? | Surgery, Yes, Yes | Sequencing | NR | 40 | Reported in text | No |
| Soontrapornchar | 1999 | Australia | 135                   | 135 (100)                   | No, only curative surgery in survival analysis? | Biopsies, Yes, Yes | SSCP | -- | 36 | Reported in text | No |
| Schneider   | 2000 | Germany | 59                     | 59 (100)                    | Yes, only curative surgery in survival analysis? | Biopsies, Yes, No | Sequencing | -- | 44 | Reported in text | Yes |
| Ireland     | 2000 | USA     | 37                     | 37 (100)                    | Yes, only curative surgery in survival analysis? | N/A, No, No | Sequencing | -- | 49 | Calculated from raw data | No |
| Aloia       | 2001 | USA     | 61                     | 44 (72)                     | Yes, only curative surgery in survival analysis? | Surgery, No, No | IHC #1801 (Biogenex) | 200 | 67 | Reported in text | Yes |
| Gibson      | 2003 | USA     | 54                     | 41 (76)                     | Yes, only curative surgery in survival analysis? | Biopsies, Yes, Yes | Sequencing | -- | 63 | Reported in text | Yes |
| Falkenback  | 2008 | Sweden  | 54                     | 54 (100)                    | Yes, only curative surgery in survival analysis? | Surgery, No, No | IHC, DO7 | 300 | 60 | Reported in text | No |
| Madani      | 2009 | Canada  | 142                    | 142 (100)                   | Yes, only curative surgery in survival analysis? | Surgery, No, No | Sequencing + IHC, DO7 (Dako) | 50 | 47 | Reported in text | Yes |
| Cavazzola   | 2009 | Brazil  | 46                     | 46 (100)                    | Yes, only curative surgery in survival analysis? | Surgery, No, No | IHC, DO7, PAb1801 (Sigma) | 100 | 52 | Reported in text | Yes |
| Leirbach    | 2009 | Brazil  | 75                     | 75 (100)                    | Yes, only curative surgery in survival analysis? | Surgery, No, No | IHC, DO7 (Novocstra) | NR | 60 | Extrapolated | No |
| Fareed      | 2010 | UK      | 245                    | 83 (81)                     | Yes, only curative surgery in survival analysis? | Surgery, Yes, Yes | IHC, DO7? (Vector Labs) | 50 | 30 | Extrapolated | No |
| Kandodoier  | 2014 | Austria | 36                     | 20 (56)                     | No, only curative surgery in survival analysis? | Biopsies, Yes, Yes | Sequencing | -- | 50 | Reported in text | Yes |

CRTx, chemoradiotherapy; IHC, immunohistochemistry; LOH, loss of heterozygosity; NR, not reported; OAC, oesophageal adenocarcinoma; SSCP, single-strand confirmation polymorphism.
Table 2  Risk of bias assessment

| Reference                  | Patient inclusion/exclusion criteria clearly defined | Patient treatment clearly characterised | Specimen characteristics | Adequate detection method of TP53 mutation | Study design | Study/statistical methods | Presentation (explanation of dropouts, number of events) | Reporting of basic demographic characteristics | Comparison of marker to standard prognostic variable | Univariate and/or time-to-event data presentation | Multivariable analysis adjusting for standard prognostic factors and adequate reporting hereof | Other potential sources of bias | Risk of bias |
|----------------------------|------------------------------------------------------|----------------------------------------|---------------------------|--------------------------------------------|--------------|--------------------------|------------------------------------------------|------------------------------------------------|-----------------------------------------------|-----------------------------------------------|------------------------------------------------|-----------------------------------------------|----------------|
| Fléjou et al\(^{19}\)    | Low                                                  | Low                                    | Low                       | High                                       | Low          | Low                      | Low                                                  | Low                                                  | Low                                           | Low                                           | High                                           | Low                                           | High                      |
| Duhaylongsod et al\(^{40}\) | Low                                                  | Low                                    | Low                       | Unclear                                    | Low          | Low                      | Low                                                  | Low                                                  | Low                                           | Low                                           | High                                           | Low                                           | High                      |
| Sauter et al\(^{18}\)    | Low                                                  | Low                                    | Low                       | Unclear                                    | Low          | Low                      | Low                                                  | Low                                                  | Low                                           | Low                                           | Unclear                                        | Low                                           | High                      |
| Ribeiro et al\(^{60}\)   | Low                                                  | Low                                    | Low                       | Low                                        | Low          | Unclear                  | Low                                                  | Low                                                  | Low                                           | Low                                           | High                                           | Unclear                                        | High                      |
| Soonthrapornchai et al\(^{56}\) | Low                                                  | Low                                    | Low                       | Low                                        | Low          | Unclear                  | Low                                                  | Low                                                  | Low                                           | Low                                           | Low                                           | Low                                           | High                      |
| Wu et al\(^{81}\)        | Low                                                  | Low                                    | Low                       | Low                                        | Low          | Low                      | Low                                                  | Low                                                  | Low                                           | Low                                           | Low                                           | Low                                           | Low                      |
| Ireland et al\(^{42}\)   | Low                                                  | Low                                    | Low                       | Low                                        | Low          | Low                      | Low                                                  | Low                                                  | Low                                           | Low                                           | Low                                           | Low                                           | Low                      |
| Schneider et al\(^{43}\) | Low                                                  | Low                                    | Low                       | Low                                        | Low          | Low                      | Low                                                  | Low                                                  | Low                                           | Low                                           | Low                                           | Low                                           | Low                      |
| Aloia et al\(^{61}\)     | Low                                                  | Low                                    | Low                       | Low                                        | Low          | Low                      | Low                                                  | Low                                                  | Low                                           | Low                                           | Low                                           | Low                                           | Low                      |
| Gibson et al\(^{62}\)    | Low                                                  | Low                                    | Low                       | Low                                        | Low          | Low                      | Low                                                  | Low                                                  | High                                          | Low                                           | Low                                           | Low                                           | Low                      |
| Falkenbcket al\(^{44}\)  | Low                                                  | Low                                    | Low                       | Low                                        | Low          | Low                      | Low                                                  | Low                                                  | Low                                           | Low                                           | Low                                           | Low                                           | Low                      |
| Cavazzola et al\(^{47}\) | Low                                                  | Low                                    | Low                       | Low                                        | Low          | Low                      | Low                                                  | Low                                                  | Low                                           | Low                                           | Low                                           | Low                                           | Low                      |
| Lehrbach et al\(^{46}\)  | Low                                                  | Low                                    | Low                       | Low                                        | Low          | Low                      | Low                                                  | Low                                                  | Low                                           | Low                                           | Low                                           | Low                                           | Low                      |
| Madani et al\(^{55}\)    | Low                                                  | Low                                    | Low                       | Unclear                                    | Low          | Low                      | Low                                                  | Low                                                  | Low                                           | Low                                           | Low                                           | Unclear                                        | High                      |
| Fareed et al\(^{53}\)    | Low                                                  | Low                                    | Low                       | Unclear                                    | Low          | Low                      | Low                                                  | Low                                                  | Low                                           | Low                                           | Low                                           | Low                                           | Low                      |
| Kandioler et al\(^{44}\) | Low                                                  | Low                                    | Low                       | Low                                        | Low          | Low                      | Low                                                  | Low                                                  | Low                                           | Low                                           | Low                                           | Low                                           | Low                      |
and meta-analyses of various prognostic biomarkers in OAC have included an analysis of TP53 mutations.\textsuperscript{56, 57} Both studies reported similar significant negative effect estimates on patient survival with two\textsuperscript{57} and five\textsuperscript{56} primary studies in their respective analyses, but did not consider potential confounders such as tumour stage or TP53 mutation analysis methods. As such, with 16 studies and $>800$ patients, our study is the largest and most extensive analysis of the effect of TP53 mutations on OAC patient survival.

This study is timely because recent whole-genome sequencing studies have shown a high mutation rate of TP53 in OAC.\textsuperscript{19, 20} With whole-genome sequencing technologies still not used in clinical practice for personalised cancer treatment, targeted genomic approaches remain a valid area of investigation.\textsuperscript{58} One potential explanation for earlier conflicting results for the prognostic significance of TP53 status in OAC is the use of different and potentially less accurate assay methods for TP53 mutation detection. Our study identified considerable variability in IHC methods across the included studies, such as the use of different antibodies, antibody dilutions and variable scoring systems for immunopositivity. Further, none of the included studies used loss of p53 expression as a method for interpreting

### Figure 2
Forest plot of the effect of tumour protein 53 (TP53) mutation status on survival, all 16 included studies.

### Figure 3
Forest plot of the effect of tumour protein 53 (TP53) on survival stratified by tumour histology included in studies.

---

| Study or Subgroup | log[Hazard Ratio] | SE  | Weight | Hazard Ratio IV, Random, 95% CI | Year | Hazard Ratio IV, Random, 95% CI |
|-------------------|-------------------|-----|--------|-------------------------------|------|-------------------------------|
| Fléjou            | -0.351            | 0.378 | 7.2%  | 0.70 [0.34, 1.48]              | 1994 |                               |
| Sauter            | 1.08              | 0.781 | 2.3%  | 2.94 [0.64, 13.61]             | 1995 |                               |
| Duhaylongsod      | 0.077             | 0.555 | 4.1%  | 1.08 [0.36, 3.21]              | 1995 |                               |
| Wu                | 0.588             | 0.331 | 8.5%  | 1.80 [0.94, 3.44]              | 1998 |                               |
| Ribeiro           | 1.024             | 0.479 | 5.2%  | 2.78 [1.09, 7.12]              | 1998 |                               |
| Soontarporchhai    | 0.139             | 0.4   | 6.7%  | 1.15 [0.52, 2.52]              | 1999 |                               |
| Schneider         | 0.904             | 0.396 | 6.8%  | 2.47 [1.14, 5.37]              | 2000 |                               |
| Ireland           | 0.88              | 0.53  | 4.4%  | 2.41 [0.85, 6.81]              | 2000 |                               |
| Aloia             | 0.88              | 0.389 | 6.9%  | 2.41 [1.12, 5.17]              | 2001 |                               |
| Gibson            | -0.792            | 0.452 | 5.6%  | 0.45 [0.19, 1.10]              | 2003 |                               |
| Falkenback        | 0.06              | 0.377 | 7.2%  | 1.06 [0.51, 2.22]              | 2008 |                               |
| Madani            | 0.432             | 0.204 | 13.2% | 1.54 [1.03, 2.30]              | 2009 |                               |
| Cavazzola         | 0.357             | 0.612 | 3.5%  | 1.43 [0.43, 4.74]              | 2009 |                               |
| Lehrbach          | 0.443             | 0.406 | 6.6%  | 1.56 [0.70, 3.45]              | 2009 |                               |
| Fareed            | 0.14              | 0.342 | 8.2%  | 1.15 [0.59, 2.25]              | 2010 |                               |
| Kandider          | 1.4               | 0.62  | 3.5%  | 4.06 [1.20, 13.67]             | 2014 |                               |

**Total (95% CI)** 100.0% 1.48 [1.16, 1.90]

Heterogeneity: Tau² = 0.08; Chi² = 22.43, df = 15 (P = 0.10); I² = 33%

Test for overall effect: Z = 3.12 (P = 0.002)

---

| Study or Subgroup | log[Hazard Ratio] | SE  | Weight | Hazard Ratio IV, Random, 95% CI | Year | Hazard Ratio IV, Random, 95% CI |
|-------------------|-------------------|-----|--------|-------------------------------|------|-------------------------------|
| Ribeiro           | 1.024             | 0.479 | 5.2%  | 2.78 [1.09, 7.12]              | 1998 |                               |
| Aloia             | 0.88              | 0.389 | 6.9%  | 2.41 [1.12, 5.17]              | 2001 |                               |
| Gibson            | -0.792            | 0.452 | 5.6%  | 0.45 [0.19, 1.10]              | 2003 |                               |
| Fareed            | 0.14              | 0.342 | 8.2%  | 1.15 [0.59, 2.25]              | 2010 |                               |
| Kandider          | 1.4               | 0.62  | 3.5%  | 4.06 [1.20, 13.67]             | 2014 |                               |

**Subtotal (95% CI)** 29.4% 1.62 [0.80, 3.29]

Heterogeneity: Tau² = 0.45; Chi² = 13.35, df = 4 (P = 0.010); I² = 70%

Test for overall effect: Z = 1.34 (P = 0.18)

---

| Study or Subgroup | log[Hazard Ratio] | SE  | Weight | Hazard Ratio IV, Random, 95% CI | Year | Hazard Ratio IV, Random, 95% CI |
|-------------------|-------------------|-----|--------|-------------------------------|------|-------------------------------|
| Fléjou            | -0.351            | 0.378 | 7.2%  | 0.70 [0.34, 1.48]              | 1994 |                               |
| Sauter            | 1.08              | 0.781 | 2.3%  | 2.94 [0.64, 13.61]             | 1995 |                               |
| Duhaylongsod      | 0.077             | 0.555 | 4.1%  | 1.08 [0.36, 3.21]              | 1995 |                               |
| Wu                | 0.588             | 0.331 | 8.5%  | 1.80 [0.94, 3.44]              | 1998 |                               |
| Soontarporchhai    | 0.139             | 0.4   | 6.7%  | 1.15 [0.52, 2.52]              | 1999 |                               |
| Schneider         | 0.904             | 0.396 | 6.8%  | 2.47 [1.14, 5.37]              | 2000 |                               |
| Ireland           | 0.88              | 0.53  | 4.4%  | 2.41 [0.85, 6.81]              | 2000 |                               |
| Lehrbach          | 0.443             | 0.406 | 6.6%  | 1.56 [0.70, 3.45]              | 2009 |                               |
| Madani            | 0.432             | 0.204 | 13.2% | 1.54 [1.03, 2.30]              | 2009 |                               |
| Cavazzola         | 0.357             | 0.612 | 3.5%  | 1.43 [0.43, 4.74]              | 2009 |                               |

**Subtotal (95% CI)** 70.6% 1.46 [1.17, 1.83]

Heterogeneity: Tau² = 0.00; Chi² = 9.05, df = 10 (P = 0.53); I² = 0%

Test for overall effect: Z = 3.32 (P = 0.0009)

---

### Figure 2
Forest plot of the effect of tumour protein 53 (TP53) mutation status on survival, all 16 included studies.

### Figure 3
Forest plot of the effect of tumour protein 53 (TP53) on survival stratified by tumour histology included in studies.

---

Fisher OM, et al. Gut 2017;66:399–410. doi:10.1136/gutjnl-2015-310888

Oesophagus

Fisher OM, et al. Gut 2017;66:399–410. doi:10.1136/gutjnl-2015-310888
The presence of TP53 gene mutations. These technical variations, combined with the knowledge that not all TP53 mutations lead to accumulation of mutant protein, and that wild-type p53 protein can be overexpressed, leading to either false negative or false positive staining results, may explain why IHC has been regarded as a less adequate analysis method in other cancers. In our subgroup analysis by assay type, the prognostic effect of TP53 appeared to be smaller in studies that used IHC versus other methods, particularly for mixed histology populations; however, a test for interaction was not statistically significant. We were also not able to demonstrate that the use of IHC contributed to interstudy heterogeneity in our meta-regression analysis. Nevertheless, with increased knowledge about TP53 mutation variants and more standardised

Figure 4 Forest plot of the effect of tumour protein 53 (TP53) on patient survival stratified by TP53 analysis methodology, including all studies (A) and only those studies with pure oesophageal adenocarcinoma cohorts (B).
and economical targeted gene sequencing technologies available, our results support the validity of using direct TP53 mutation status as a prognostic biomarker in OAC and raise the possibility that this provides improved prognostic classification than IHC consistent with findings in other cancers. Recent genomic studies have found that TP53 gene mutations remain the most common genetic alteration in both OAC and its precursor lesion BO with dysplasia, with wild-type TP53 estimated to be present in only approximately 30% of OAC tumours. These findings suggest that the mutation...
frequency of TP53 may be underestimated in all studies included here, even those performing TP53 gene sequencing. Most studies only sequenced exons 5–8, whereas whole-genome sequencing (which includes the sequencing of intronic and intergenic regions) has been shown to be more accurate at detecting mutations, even within exonic regions. Further research using current sequencing technologies will be needed to assess whether the overall prognostic impact of TP53 mutation status is larger or smaller than the present study estimate, and to determine whether the prognostic effect of TP53 mutations varies by type of gene mutation. While our subgroup analysis of only TP53 gene sequencing studies showed a pronounced effect of TP53 mutations on OAC patient survival, this pooled analysis displayed substantial heterogeneity. This was largely caused by one study that reported a non-significant survival advantage for TP53 mutations. Although this study included sequencing data from pretreatment biopsies of mainly patients with OAC, no information is provided on patients’ pathological AJCC/UICC tumour stage, which to date remains the most accurate predictor of patient prognosis. Study patients were enrolled from two experimental, prospective studies assessing the effect of two similar neoadjuvant CRTx regimens. Patients were only staged clinically and the HR estimation could not be adjusted for pathological tumour stage, because of which we assessed the estimate as being at high risk of bias. A sensitivity analysis performed to exclude this study suggests that gene sequencing-defined TP53 mutations are associated with an almost twofold mortality risk (HR 1.95, 95% CI 1.44 to 2.65).

The two main limitations of this study that may affect the validity of our findings are the quality of the primary studies and data limitations to explore potential confounders. First, more than half of the studies were assessed as high risk of bias based on criteria defined by Hayden et al., as a significant contributor to interstudy heterogeneity. Equally, the lack of adjustment of HR estimates for standard prognostic criteria, as recommended in current biomarker reporting guidelines, was also a significant contributor to heterogeneity. These methodological flaws may lead to an underestimate of the actual effect size, as suggested by our subgroup analyses. Second, we were unable to conduct analyses to investigate the impact of potentially relevant factors (such as the inclusion of Siewert Type III or proximal gastric cancers or patient smoking status) because of the lack of data reported in the original publications. Other concerns may be how the inclusion of mixed patient populations, different staging systems and treatment regimens may impact the present findings. However, our meta-regression did not identify these factors as potential sources impacting overall effect estimates. Further, the key finding of TP53 negatively affecting patient prognosis persisted across all pooled and subgroup analyses, despite our comprehensive methodological approach, which accounted for sources of confounding, bias and interstudy heterogeneity. But most of the studies only included patients from surgical series, potentially limiting the generalisability of our findings. Despite almost 70% of the studies (n = 11) having AJCC stage IV patients in their analysis, only nine of these studies also included such patients in their p53-stratified survival analysis. With the median percentage of stage IV patients in such studies being limited to 5.5%, the generalisability of our findings to patients with OAC with more advanced stages of disease that do not permit curative treatment is limited. Finally, despite conducting an extensive search strategy including searching conference abstracts and presenting a funnel plot that excludes major asymmetry, we cannot eliminate publication bias as a possible explanation of our results.

Further studies are warranted to better estimate the size of the prognostic effect independent of tumour, node, metastasis staging. Given the substantial amount of heterogeneity identified, adherence to REMARK guidelines and adjustment of the survival analysis for known prognostic factors in future studies is recommended. Further, preregistration of such prognostic studies is important to help avoid the issues of reporting and publication bias. Using assumptions based on the findings of the present meta-analysis and a mutation frequency of 70%, we estimate a minimum of 433 patients with OAC (303 TP53 mutated and 130 TP53 wild-type) would be required to determine the effect of TP53 mutations on patient overall survival.

One approach to collect high-quality data is to include TP53 mutation analysis using targeted gene sequencing in the baseline analysis of trials of OAC therapies. If validated, TP53 analysis could be used to stratify patients in future trials. Stratifying patients based on TP53 mutation status may also have a role in clinical practice to guide treatment selection. For example, data suggest that TP53 mutation status may predict response to standard chemotherapeutic regimens such as fluorouracil or cisplatin, which has also recently been demonstrated in OAC. A prospective randomised trial in oesophageal cancer, aiming at determining this predictive effect of TP53 gene mutations (p53-Adjusted Neoadjuvant Chemotherapy for Potentially Resectable Esophageal Cancer; http://www.clinicalTrials.gov identifier NCT00525200; http://www.p53.at) has recently completed recruitment. Moreover, multiple therapeutic options directly targeting the TP53 gene are either currently in clinical trials or are already clinically available. For example, the first-in-class mutant p53 reactivator APR-246 has recently been shown to have significant antitumourigenic activity in OAC and synergises with DNA damaging agents such as cisplatin and 5-fluorouracil. As recent next-generation sequencing studies have demonstrated a high frequency of TP53 mutations in OAC of 70%, it seems likely that a TP53-directed therapeutic approach would be worthwhile for patients with this highly fatal cancer.

In summary, OAC remains one of the few GI malignancies for which molecular information is still not used to guide patient management. This study suggests that TP53 gene mutations have a clinically important negative prognostic impact on patients with OAC, which is relevant in light of recent genomic findings highlighting a central role of this gene in OAC pathogenesis and drugs currently in development and testing that directly target this gene. High-quality studies with large patient cohorts using modern sequencing technologies for TP53 mutation analysis are needed to confirm the independent prognostic effect of this frequent gene mutation.

Author affiliations
1Gastroesophageal Cancer Program, St Vincent’s Centre for Applied Medical Research University of New South Wales, Sydney, New South Wales, Australia
2NHMRC Clinical Trials Centre University of Sydney, Sydney, New South Wales, Australia
3Department of Epidemiology and Medical Statistics, School of Medicine, University of Notre Dame, Sydney, New South Wales, Australia
4Department of Surgery, Lund University Hospital (Skåne University Hospital) and Lund University, Lund, Sweden
5Cancer Biology and Surgical Oncology Research Laboratory, Peter MacCallum Cancer Centre, Melbourne, Victoria, Australia
6Sir Peter MacCallum Department of Oncology, University of Melbourne, Melbourne, Victoria, Australia
REFERENCES

1. Pohl H, Welch HG. The role of overdiagnosis and reclassification in the marked increase of esophageal adenocarcinoma incidence. J Natl Cancer Inst 2013;105:1714-22.

2. Clemens N, Phillips W, Lord RV. Signaling pathways in the molecular pathogenesis of adenocarcinomas of the esophagus and gastroesophageal junction. Cancer Biol Ther 2013;14.

3. Arnold M, Sjoerdmamara I, Ferlay J, et al. Global incidence of esophageal cancer by histological subtype in 2012. Gut 2015;64:381-7.

4. Edgren G, Adami HO, Wikerskas E, et al. A global assessment of the esophageal adenocarcinoma epidemic. Gut 2013;62:1406-14.

5. Ehemann C, Henley SJ, Ballard-Barbash R, et al. Annual Report to the Nation on the status of cancer, 1975–2008, featuring cancers associated with excess weight and lack of sufficient physical activity. Cancer 2012;118:2338-66.

6. Brown LM, Devesa SS, Chow WH. Incidence of adenocarcinomas of the esophagus and gastroesophageal junction: relationship between clinicopathological data and p53, cyclin D1 and B-2 immunoeexpression. Ann Oncol 2008;46:837–43.

7. Young KH, Leroy K, Muller MB, et al. Structural profiles of TP53 gene mutations predict clinical outcome in diffuse large B-cell lymphoma: an international collaborative study. Blood 2008;112:3088–98.

8. Olivier M, Hanaitou P, Barreens-Dale AL. Prognostic and predictive value of TP53 mutations in human cancer. In: Hanaitou P, Wiman K, eds. 25 years of p53 research. Dordrecht, The Netherlands: Springer, 2005:320–8.

9. Dalak AM, Stojanov P, Peng S, et al. Exome and whole-genome sequencing of esophageal adenocarcinoma identifies recurrent driver events and mutational landscape. Nat Genet 2013;45:478–86.

10. Nones K, Waddell N, Wayne N, et al. Genomic catastrophes frequently arise in esophageal adenocarcinoma and drive tumorigenesis. Nat Commun 2014;5:5224.

11. Weaver JM, Ross-Innes CS, Shannon N, et al. Ordering of mutations in preneoplastic disease stages of esophageal carcinogenesis. Nat Genet 2014;46:837–43.

12. Soussi T, Berouz C. Assessing TP53 status in human tumours to evaluate clinical outcome. Nat Rev Cancer 2001;1:233–40.

13. Kaye PV, Haider SA, James PD, et al. Novel staining pattern of p53 in Barrett’s dysplasia–the absent pattern. Histopathology 2010;57:933–5.

14. Guyett GH, Osman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008;336:942–6.

15. McShane LM, Altman DG, Sauerbrei W, et al. Reporting recommendations for tumor MARKer prognostic studies (REMARK). Nat Clin Pract Oncol 2005;2:416–22.

16. Hayden JA, van der Windt DA, Cartwright JL, et al. Assessing bias in studies of prognostic factors. Ann Intern Med 2013;158:280–6.

17. Higgins JPT GSe. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (Updated March 2011), 2011.

18. Parmar MK, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. Stat Med 1999;18:2815–34.

19. Tierney JF, Stewart LA, Ghersi D, et al. Practical methods for incorporating summary time-to-event data into meta-analysis. Trials 2007;8:16.

20. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:77–88.

21. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. BMJ 2003;327:557–60.

22. Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. Stat Med 1999;18:693–708.

23. Altman DG, Bland JM. Interaction revisited: the difference between two estimates. BMJ 2003;326:219.

24. Egger M, Davey Smith G, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:629–34.

25. R: A language and environment for statistical computing[program], Vienna, Austria: R Foundation for Statistical Computing, 2013.

26. Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer, 2009.

27. Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package. J Stat Softw 2010;36:1–48.

28. Max Gordon TL. forestplot: Advanced Forest Plot Using ggplot2: elegant graphics for data analysis. J Stat Softw 2011.[updated March 2011], 2011.

29. Flejou JF, Paraf F, Potet F, et al. Immunohistochemical evaluation for P53 mutation status improves tumor marker prediction of clinical outcome in patients with curatively resected adenocarcinoma. Eur J Cancer 2002;38:287–94.

30. Ireland AF, Shibata DK, Chandrasoma P, et al. Clinical significance of p53 mutations in adenocarcinoma of the esophagus and cardia. Ann Surg 2000;231:179–87.

31. Schneider PM, Stoelethig O, Roth JA, et al. TP53 mutational status improves estimation of prognosis in patients with curatively resected adenocarcinoma in Barrett’s esophagus. Clin Cancer Res 2006;12:6315–8.

32. Falkenbach D, Nilbert M, Oberg S, et al. Prognostic value of cell adhesion in esophageal adenocarcinoma. Eur J Cancer 2008;44:487–9.

33. Duhaylongsod FG, Gottfried MR, Iglehart JD, et al. Immunohistochemical evaluation for P53 mutation status improves tumor marker prediction of clinical outcome in patients with curatively resected adenocarcinoma. Eur J Cancer 2000;36:287–94.

34. Leibltham BM, Coccoitno I, Ribeiro U, et al. Adenocarcinoma of the esophagogastric junction: relationship between clinicopathological data and p53, cyclin D1 and B-2 immunoeexpression. Arg Gastroenterol 2009;46:315–20.

35. Cavazzola LT, Rosa AR, Shirmer CC, et al. Immunohistochemical evaluation for p53 and VEGF (Vascular Endothelial Growth Factor) is not prognostic for long term survival in end stage esophageal adenocarcinoma. Revista do Colegio Brasileiro de Cirurgios 2009;36:24–34.

36. Sauter ER, Keller SM, Emer SM, et al. TP53 gene mutation status in pretreatment biopsy of oesophageal adenocarcinoma has no prognostic value. Eur J Cancer 1999;35:1683–7.
50 Ribeiro U Jr, Finkelstein SD, Safatle-Ribeiro AV, et al. p53 sequence analysis predicts treatment response and outcome of patients with esophageal carcinoma. Cancer 1998;83:7–18.

51 Aloia TA, Harpole DH, Jr., Reed CE, et al. Tumor marker expression is predictive of survival in patients with esophageal cancer. Ann Thorac Surg 2001;72:859–66.

52 Gibson MK, Abraham SC, Wu TT, et al. Epidermal growth factor receptor, p53 mutation, and pathological response predict survival in patients with locally advanced esophageal cancer treated with preoperative chemoradiotherapy. Clin Cancer Res 2003;9:6461–8.

53 Fareed KR, Al-Attar A, Soomro IN, et al. Tumour regression and ERCC1 nuclear protein expression predict clinical outcome in patients with gastro-oesophageal cancer treated with neoadjuvant chemotherapy. Br J Cancer 2010;102:1600–7.

54 Kandioler D, Schoppmann SF, Zwrtek R, et al. The biomarker TP53 divides patients with neoadjuvantly treated esophageal cancer into 2 subgroups with markedly different outcomes. A p53 Research Group study. J Thorac Cardiovasc Surg 2014;148:2280–6.

55 Rice TW, Blackstone EH, Rusch VW. 7th edition of the AJCC Cancer Staging Manual: esophagus and esophagogastric junction. Ann Surg Oncol 2010;17:1721–4.

56 Findlay JM, Middleton MR, Tomlinson I. A systematic review and meta-analysis of somatic and germline DNA sequence biomarkers of esophageal cancer survival, therapy response and stage. Ann Oncol 2015;26:624–44.

57 Chen M, Huang J, Zhu Z, et al. Systematic review and meta-analysis of tumor biomarkers in predicting prognosis in esophageal cancer. BMC Cancer 2013;13:539.

58 Berg JS, Khoury MJ, Evans JP. Deploying whole genome sequencing in clinical practice and public health: meeting the challenge one bin at a time. Genet Med 2011;13:499–504.

59 Olivier M, Taniere P. Somatic mutations in cancer prognosis and prediction: lessons from TP53 and EGFR genes. Curr Opin Oncol 2011;23:88–92.

60 Edlund K, Larsson O, Ameer A, et al. Data-driven unbiased curation of the TP53 tumor suppressor gene mutation database and validation by ultradeep sequencing of human tumors. Proc Natl Acad Sci USA 2012;109:9551–6.

61 Akiyama J, Alexandre L, Banah A, et al. Strategy for prevention of cancers of the esophagus. Ann N Y Acad Sci 2014;1325:108–26.

62 Selkadi A, Bolze A, Itan Y. Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. Proc Natl Acad Sci USA 2015;112:5473–8.

63 Hayden IA, Cole P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. Ann Intern Med 2006;144:427–37.

64 Altman DG. The time has come to register diagnostic and prognostic research. Clin Chem 2014;60:580–2.

65 Weller M. Predicting response to cancer chemotherapy: the role of p53. Cell Tissue Res 1998;292:435–45.

66 Heeren PA, Klokpenberg FW, Hollema H, et al. Predictive effect of p53 and p21 alteration on chemotherapy response and survival in locally advanced adenocarcinoma of the esophagus. Anticancer Res 2004;24:2579–83.

67 Beardsmore DM, Verbeke CS, Davies CL, et al. Apoptotic and proliferative indexes in esophageal cancer: predictors of response to neoadjuvant therapy[corrected]. J Gastrointest Surg 2003;7:77–86; discussion 86–7.

68 van Olphen, SH, Biermann K, wijnhoven bP, et al. Sa1926: SOX2 and p53 Protein Expression Predicts Response to Preoperative Chemoradiotherapy in Patients With Esophageal Adenocarcinoma. Gastroenterology 2015;148(Suppl1):S-357.

69 Liu DS, Read M, Cullinane C, et al. APR-246 potently inhibits tumour growth and overcomes chemoresistance in preclinical models of oesophageal adenocarcinoma. Gut 2015;64:1506–16.
SUPPLEMENTARY FILE 1

Summary Findings of IARC TP53 Database Analysis

The most current R17 version of the IARC TP53 database contains three studies\(^{42, 51, 66}\) that report on the prognostic effect of TP53 mutations in esophageal adenocarcinoma cohorts, all of which are included in our meta-analysis.

In 345 EACs compiled in the IARC TP53 database, the most frequently occurring mutations were G:C to A:T transitions at CpG sites (43.5% of tumors) in exons 5, 7, 8 (Supplementary Figure 1A) with a mean mutation frequency in any single nucleotide at this site of 1.27 (SD 0.03). The most frequent effect of this type of mutation was missense mutations (found in 78% of p53 mutations; Supplementary Figure 1B).

Of all the tumors included in the database, 245 had information on their immunohistochemistry staining pattern. Missense mutations most frequently caused positive immuno-staining and occurred most commonly within the L2/L3, L1/S/H2 and NDBL/beta-sheet protein domains (Supplementary Figure 1C and 1D). However, approximately 27% of TP53 mutant tumors showed negative immuno-staining patterns (false negatives), as these are frequently deletion mutations (Supplementary Figure 2).

SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure 1. Analysis of EAC patients (n = 345) in the IARC TP53 mutation database. Supplementary Figure 1A depicts the frequency of types of TP53 gene mutations as well as their genetic location. Supplementary Figure 1B shows the resulting effects of these mutations on protein isotypes. Supplementary Figure 1C shows the corresponding IHC staining patterns of the type of TP53 gene mutations, where 1D shows how the TP53 mutation effect affects IHC staining patterns (EAC n = 245).
**Supplementary Figure 2.** Analysis of IARC TP53 database to determine frequency of interpretations of immunohistochemistry staining patterns in the presence of TP53 gene mutations.

**Supplementary Figure 3.** Funnel plot of all studies included in the present meta-analysis for assessment of possible publication bias.

**Supplementary Figure 4.** Forest plot of the effect of TP53 on survival stratified by histology and adjustment for standard prognostic variables, all studies.

**Supplementary Figure 5.** Forest plot of effect of TP53 on survival only including studies performing TP53 gene sequencing before (supplementary figure 5A) and after sensitivity analysis (removal of Gibson et al.; 5B). Supplementary figure 5C depicts the forest plot of studies with pure EAC cohorts that performed TP53 gene sequencing.

**Supplementary Figure 6.** Forest plot of meta-regression analysis of study factors associated inter-study heterogeneity. Depicted are the effect estimates (solid squares) of change in log HR and corresponding 95% confidence interval (95%CI, solid line). Total (solid diamond) represents the overall effect estimate (log HR) of mutant TP53 and corresponding 95%CI as calculated from all 16 included studies.
Supplementary Figure 2.

p53 immunohistochemistry for TP53 mutations in EAC
Supplementary Figure 3.
Supplementary Figure 4.

| Study or Subgroup | log[Hazard Ratio] | SE | Weight | Hazard Ratio IV, Random, 95% CI |
|-------------------|------------------|----|--------|---------------------------------|
| **3.27.1 Unadjusted HR estimates and mixed histology**
| Fareed            | 0.14             | 0.342 | 8.2% | 1.15 [0.59, 2.25]              |
| Gibson            | -0.792           | 0.452 | 5.6% | 0.45 [0.19, 1.10]              |
| Ribeiro           | 1.024            | 0.479 | 5.2% | 2.78 [1.09, 7.12]             |
| **Subtotal (95% CI)** |                |      |       | 19.0% | 1.12 [0.44, 2.85]               |
| Heterogeneity: $\tau^2 = 0.50$, $\chi^2 = 7.63$, $df = 2$ ($P = 0.02$); $I^2 = 74\%$
| Test for overall effect: $Z = 0.24$ ($P = 0.81$) |

| **3.27.2 Unadjusted HR estimates and EAC only cohorts**
| Duhabylongsod     | 0.077            | 0.555 | 4.1% | 1.08 [0.36, 3.21]              |
| Fléjou            | -0.351           | 0.378 | 7.2% | 0.70 [0.34, 1.48]              |
| Lehrbach          | 0.443            | 0.406 | 6.6% | 1.56 [0.70, 3.45]              |
| Madani            | 0.432            | 0.204 | 13.2%| 1.54 [1.03, 2.30]             |
| Sauter            | 1.08             | 0.781 | 2.3% | 2.94 [0.64, 13.61]            |
| Soonthapornchaimi | 0.339            | 0.4   | 6.7% | 1.15 [0.52, 2.52]             |
| **Subtotal (95% CI)** |                |      |       | 40.2% | 1.32 [1.00, 1.75]               |
| Heterogeneity: $\tau^2 = 0.00$, $\chi^2 = 4.81$, $df = 5$ ($P = 0.44$); $I^2 = 0\%$
| Test for overall effect: $Z = 1.94$ ($P = 0.05$) |

| **3.27.3 Adjusted HR estimates and mixed histology**
| Aloia             | 0.88             | 0.389 | 6.9% | 2.41 [1.12, 5.17]              |
| Kandioer          | 1.4              | 0.62  | 3.5% | 4.06 [1.20, 13.67]             |
| **Subtotal (95% CI)** |                |      |       | 10.4% | 2.79 [1.46, 5.33]               |
| Heterogeneity: $\tau^2 = 0.00$, $\chi^2 = 0.50$, $df = 1$ ($P = 0.48$); $I^2 = 0\%$
| Test for overall effect: $Z = 3.12$ ($P = 0.002$) |

| **3.27.4 Adjusted HR estimates and EAC only cohorts**
| Cavazzola         | 0.357            | 0.612 | 3.5% | 1.43 [0.43, 4.74]              |
| Falkenbach        | 0.06             | 0.377 | 7.2% | 1.06 [0.51, 2.22]              |
| Ireland           | 0.88             | 0.53  | 4.4% | 2.41 [0.85, 6.81]              |
| Schneider          | 0.904            | 0.396 | 6.8% | 2.47 [1.14, 5.37]              |
| Wu                 | 0.588            | 0.331 | 8.5% | 1.80 [0.94, 3.44]              |
| **Subtotal (95% CI)** |                |      |       | 30.5% | 1.72 [1.20, 2.48]               |
| Heterogeneity: $\tau^2 = 0.00$, $\chi^2 = 2.99$, $df = 4$ ($P = 0.56$); $I^2 = 0\%$
| Test for overall effect: $Z = 2.92$ ($P = 0.004$) |

| **Total (95% CI)** | 100.0% | 1.48 [1.16, 1.90] |
| Heterogeneity: $\tau^2 = 0.08$, $\chi^2 = 22.43$, $df = 15$ ($P = 0.10$); $I^2 = 33\%$
| Test for overall effect: $Z = 3.12$ ($P = 0.002$) |
| Test for subgroup differences: $\chi^2 = 5.18$, $df = 3$ ($P = 0.16$), $I^2 = 42.1\%$

0.01 0.1 1 10 100
Favors wild-type TP53
Favors mutant TP53
### Supplementary Figure 5.

#### A

| Study or Subgroup | log[Hazard Ratio] | SE   | Weight | Hazard Ratio IV, Random, 95% CI | Year |
|-------------------|------------------|------|--------|---------------------------------|------|
| Ribeiro           | 1.024            | 0.479| 15.4%  | 2.78 [1.09, 7.12]               | 1998 |
| Schneider         | 0.904            | 0.396| 18.0%  | 2.47 [1.14, 5.37]               | 2000 |
| Ireland           | 0.88             | 0.53 | 13.9%  | 2.41 [0.85, 6.81]               | 2000 |
| Gibson            | -0.792           | 0.452| 16.2%  | 0.45 [0.19, 1.10]               | 2003 |
| Madani            | 0.432            | 0.204| 24.8%  | 1.54 [1.03, 2.30]               | 2009 |
| Kandilor          | 1.4              | 0.62 | 11.7%  | 4.06 [1.20, 13.67]              | 2014 |
| **Total (95% CI)** |                 |      | 100.0% | 1.80 [1.05, 3.08]              |      |

Heterogeneity: $\tau^2 = 0.26$; $\chi^2 = 13.13$, df = 5 ($P = 0.02$); $I^2 = 62$

Test for overall effect: $Z = 2.13$ ($P = 0.03$)

#### B

| Study or Subgroup | log[Hazard Ratio] | SE   | Weight | Hazard Ratio IV, Random, 95% CI | Year |
|-------------------|------------------|------|--------|---------------------------------|------|
| Ribeiro           | 1.024            | 0.479| 10.6%  | 2.78 [1.09, 7.12]               | 1998 |
| Ireland           | 0.88             | 0.53 | 8.7%   | 2.41 [0.85, 6.81]               | 2000 |
| Schneider         | 0.904            | 0.396| 15.6%  | 2.47 [1.14, 5.37]               | 2000 |
| Madani            | 0.432            | 0.204| 58.7%  | 1.54 [1.03, 2.30]               | 2009 |
| Kandilor          | 1.4              | 0.62 | 6.4%   | 4.06 [1.20, 13.67]              | 2014 |
| **Total (95% CI)** |                 |      | 100.0% | 1.95 [1.44, 2.65]              |      |

Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 3.80$, df = 4 ($P = 0.43$); $I^2 = 0$

Test for overall effect: $Z = 4.28$ ($P < 0.0001$)

#### C

| Study or Subgroup | log[Hazard Ratio] | SE   | Weight | Hazard Ratio IV, Random, 95% CI | Year |
|-------------------|------------------|------|--------|---------------------------------|------|
| Ireland           | 0.88             | 0.53 | 10.5%  | 2.41 [0.85, 6.81]               | 2000 |
| Schneider         | 0.904            | 0.396| 18.8%  | 2.47 [1.14, 5.37]               | 2000 |
| Madani            | 0.432            | 0.204| 70.7%  | 1.54 [1.03, 2.30]               | 2009 |
| **Total (95% CI)** |                 |      | 100.0% | 1.76 [1.26, 2.47]              |      |

Heterogeneity: $\tau^2 = 0.00$; $\chi^2 = 1.51$, df = 2 ($P = 0.47$); $I^2 = 0$

Test for overall effect: $Z = 3.31$ ($P = 0.0009$)
Supplementary Figure 6.

| Study variable                              | Number of studies | logHR change estimate | 95% CI         |
|---------------------------------------------|-------------------|-----------------------|----------------|
| Pure EAC cohort                             | 16                | -0.0670               | -0.629–0.496   |
| No immunohistochemistry                     | 16                | 0.2440                | -0.248–0.735   |
| Adjusted HR estimates                       | 16                | 0.4590                | 0.011–0.908    |
| Risk of bias = low                          | 16                | 0.5800                | 0.058–1.102    |
| Number of patients                          | 16                | -0.0020               | -0.009–0.007   |
| Study cohort age (mean/median)              | 14                | -0.0960               | -0.183–0.009   |
| Surgical specimens analysed                 | 15                | -0.0420               | -0.643–0.56    |
| Only surgically treated patients            | 16                | -0.2010               | -1.044–0.643   |
| CRTx included                               | 15                | -0.0590               | -0.541–0.424   |
| Neo-adjuvant CRTx included                  | 15                | -0.1900               | -0.649–0.27    |
| Percentage p53 mutated                      | 16                | -0.0180               | -0.039–0.003   |
| Stage 4 included in analysis                | 11                | -0.4670               | -1.217–0.284   |
| p53 mutated EAC more LN                     | 12                | 0.1970                | -0.438–0.831   |
| p53 mutated EAC higher stage                | 12                | 0.1140                | -0.494–0.723   |
| **Total**                                   | **16**            | **0.3941**            | **0.151–0.638**|
Supplementary Table 1. Clinicopathological and survival data of all studies included in final meta-analysis.

| Reference       | N   | T1  | T1a | T1b | T2  | T3  | T4  | N0 | N+ | N1  | N2  | N+ | N1  | N2  | Mx | M1 | M2 | G1 | G2 | G3 | Stage 0 | Stage I | Stage II | Stage III | Stage IV | TNM Stage Edition | R0 | R1 | R2 | RX | Median Survival Total (months) | Median Survival TP53 "wild-type" | Median Survival TP53 "mutated" |
|-----------------|-----|-----|-----|-----|-----|-----|-----|----|----|-----|-----|----|-----|-----|----|----|----|----|----|-----|--------|--------|----------|-----------|----------|---------------------|-----|-----|-----|-----|--------------------------------|-----------------------------|-----------------------------|
| Flijou et al.   | 62  | 20  | 8   | 12  | 7   | 35  | -   | 30 | 32 | 32  |    |    |    |    |    |    |    |    |    |    | 25 | 20 | 17 |    | 19 | 14 | 23 |    |    | 6    |    |    |    |    | 28.0 | 15 | 15 |
| Dukhayling et al. | 42  | -   | -   | -   | -   | -   | 28  | 14 | -  | -   | -   | -  | -   | -   | -   | -   | -   | -   | -   | -   |    | 24 | 42 | 15 |    | 4    |    |    |    |    | 19.8 | 18.2 |
| Sauter et al.   | 24  | -   | -   | -   | -   | 8   | 8   | -  | -  | -   | 15 | -   | 8   | -   | -   | -   | -   | -   | -   | -   |    |    |    |    | 11   | 24 | 42 | 15 |    | 28.0 | 13.0 |
| We et al.       | 92  | -   | -   | -   | -   | -   | -   | -  | -  | -   | -   | -  | -   | -   | -   | -   | -   | -   | -   | -   |    | 11 | 24 | 42 | 15 |    |    |    |    | 28.0 | 13.0 |
| Ribeiro et al.  | 42  | 7   | 8   | -   | -   | 8   | 12  | -  | 19 | 18 | 18  | -   | 1   | 4   | 15 | 13 | 7   | 4   | 17 | 8   | 1   | 4   | -   | -   | -   | -   | 16.3 | 10 | 23 | 10 | 23 | 21.6 | 40.2 |
| Sonntraporkul et al. | 13  | 5   | -   | -   | -   | -   | -   | 20 | 34 | -   |    | 21 | 5   | 34 | 33 | -   | -   | -   | -   | -   | 6   |    | 33 | 15 |    |    |    |    |    | 12.0 | 14.0 |
| Schmidt et al.  | 92  | -   | -   | -   | -   | -   | -   | -  | -  | -   | -   | -  | -   | -   | -   | -   | -   | -   | -   | -   |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| Ireland et al.  | 37  | 5   | 3   | 2   | 7   | 19  | -   | 7  | 24 | 10 | 14  | -   | 2   | 4   | 13 | 14 | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 20.0 | 16 | 23 | 10 | 28 | 41.7 | 39.8 |
| Akola et al.    | 61  | 3   | 11  | -   | -   | 14  | 16  | -  | 61 | -   | -   | -  | -   | -   | -   | -   | -   | -   | -   | -   |    | 31 | 30 | 0  | 0   | 5   | 61 | 0   | 0   | 8   | 38.3 | 18.0 | 49.0 |
| Gibson et al.   | 44  | -   | -   | -   | -   | -   | -   | -  | -  | -   | -   | -  | -   | -   | -   | -   | -   | -   | -   | -   |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| Falkenhuber et al. | 54  | 4   | 16  | -   | -   | 7   | 32  | -  | 36 | 23 | 23  | -   | 1   | 4   | 16 | 17 | 21 | 1   | 6   | -   | -   | -   |    | 41.7 | 41.7 | 39.8 |
| Madaan et al.   | 14  | 2   | -   | -   | -   | 6   | 13  | 13 | 18 | 20 | 20  | -   | 1   | 3   | 6   | 12 | 20  | -   | 5   | 9   | 20 | 4   | -   | 38 | 0   | 0   | 0   | 70.4 | 58.1 | 63.2 |
| Cavazzola et al. | 46  | -   | 6   | -   | -   | 6   | 13  | 13 | 18 | 20 | 20  | -   | 1   | 3   | 6   | 12 | 20  | -   | 5   | 9   | 20 | 4   | -   | 38 | 0   | 0   | 0   | 70.4 | 58.1 | 63.2 |
| Lehrbach et al. | 75  | 4   | -   | -   | 27  | 44  | 0   | 21 | 54 | 54  | -   | 5   | -   | -   | -   | -   | -   | -   | -   | 14 | 15 | 30 | 16 | 6   | -   | -   | -   | -   | -   | 21.5 | 29.5 |
| Fareed et al.   | 24  | 5   | 2   | 18  | -   | -   | 72  | 139 | 14 | 62 | 183 | -   | -   | 2   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | 26.7 | 53.2 |
| Kandider et al. | 26  | 3   | 5   | -   | -   | 7   | 13  | 8  | 10 | 17 | 17  | -   | 3   | 5   | 13 | 8   | -   | -   | -   | -   | -   | -   | -   | 24 | 3   | -   | -   | 13.9 | 8.6 | 26.2 |
| Subgroup                                      | Number of studies | Number of patients | Pooled HR (95% CI) | p-value | Heterogeneity | $I^2$ statistic (p-value) |
|-----------------------------------------------|-------------------|--------------------|---------------------|---------|---------------|--------------------------|
| **Assay and histology type**                  |                   |                    |                     |         |               |                          |
| Immunohistochemistry only                     | 8                 | 417                | 1.28 (0.95 – 1.73)  | 0.10    |               | 0% (0.43)                |
| Immunohistochemistry only and pure EAC cohorts| 6                 | 290                | 1.14 (0.79 – 1.66)  | 0.49    |               | 0% (0.57)                |
| No immunohistochemistry                       | 8                 | 471                | 1.68 (1.14 – 2.47)  | 0.009   |               | 50% (0.05)               |
| No immunohistochemistry after sensitivity analysis | 7           | 425                | 1.82 (1.40 – 2.36)  | <0.0001 |               | 0% (0.50)                |
| No immunohistochemistry and pure EAC cohorts  | 5                 | 354                | 1.68 (1.27 – 2.22)  | 0.0003  |               | 0% (0.64)                |
| Sequencing only studies                       | 6                 | 330                | 1.80 (1.05 – 3.08)  | 0.03    |               | 62% (0.02)               |
| Sequencing only studies after sensitivity analysis | 5           | 284                | 1.95 (1.44 – 2.65)  | <0.0001 |               | 0% (0.43)                |
| Sequencing only studies and pure EAC cohorts  | 3                 | 213                | 1.76 (1.26 – 2.47)  | 0.0009  |               | 0% (0.47)                |
| **Adjustment for tumor stage**                |                   |                    |                     |         |               |                          |
| Studies including HRs adjusting for standard prognostic variables | 7               | 327                | 1.94 (1.41 – 2.66)  | <0.0001 |               | 0% (0.53)                |
| Studies including HRs adjusting for standard prognostic variables and pure EAC only cohorts | 5               | 258                | 1.72 (1.20 – 2.48)  | 0.004   |               | 0% (0.56)                |
| Studies with unadjusted HRs for standard prognostic variables | 9               | 533                | 1.22 (0.88 – 1.70)  | 0.24    |               | 38% (0.12)               |
| Studies with unadjusted HRs for standard prognostic variables, but pure EAC cohorts | 6               | 386                | 1.32 (1.00 – 1.75)  | 0.05    |               | 0% (0.44)                |
| **Risk of bias and tumor type**               |                   |                    |                     |         |               |                          |
| Low risk of bias                              | 4                 | 197                | 2.29 (1.50 – 3.48)  | 0.0001  |               | 0% (0.70)                |
| Low risk of bias and EAC only cohorts         | 3                 | 161                | 2.11 (1.35 – 3.31)  | 0.001   |               | 0% (0.80)                |
| High and unclear risk of bias                 | 12                | 691                | 1.29 (0.98 – 1.70)  | 0.07    |               | 30% (0.15)               |
| Subgroup                                                                 | Number of studies | Number of patients | Pooled HR (95% CI) | p-value  | Heterogeneity I² statistic (p-value) |
|------------------------------------------------------------------------|-------------------|--------------------|--------------------|----------|------------------------------------|
| High risk of bias but pure EAC cohort                                 | 8                 | 488                | 1.29 (1.00 – 1.67) | 0.05     | 0% (0.64)                         |
| **Table 3 continued.**                                                 |                   |                    |                    |          |                                    |
| **Other exploratory subgroups of interest**                            |                   |                    |                    |          |                                    |
| Only chemo-radiotherapy naïve patients (regardless of risk of bias)    | 6                 | 397                | 1.60 (1.21 – 2.11) | 0.0009   | 0% (0.71)                         |
| Only chemo-radiotherapy naïve patients (regardless of risk of bias) *pure EAC only cohorts* | 5                 | 336                | 1.50 (1.11 – 2.02) | 0.0008   | 0% (0.80)                         |
| Only studies with no neo-adjuvant chemo-radiotherapy (regardless of risk of bias) | 7                 | 446                | 1.68 (1.29 – 2.18) | 0.0001   | 0% (0.67)                         |
| Only studies with no neo-adjuvant chemo-radiotherapy *pure EAC cohorts* (regardless of risk of bias) | 6                 | 385                | 1.60 (1.21 – 2.11) | 0.0009   | 0% (0.69)                         |