# Trends in conditional overall survival of esophageal cancer: a population-based study

Wei Deng[1#^], Rong Yu[1#], Zhao Yang[2#], Xin Dong[1], Weihu Wang[1]

[1]Department of Radiation Oncology, Peking University Cancer Hospital & Institute, Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Beijing, China; [2]School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Sassoon Road, Hong Kong, China

*Contributions:* (I) Conception and design: Z Yang, W Wang; (II) Administrative support: W Wang, R Yu; (III) Provision of study materials or patients: W Deng, Z Yang; (IV) Collection and assembly of data: W Deng, Z Yang; (V) Data analysis and interpretation: W Deng, Z Yang, R Yu, W Wang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

*Correspondence to:* Weihu Wang, MD; Rong Yu, MD. Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Department of Radiation Oncology, Peking University Cancer Hospital & Institute, 52 Fucheng Road, Beijing 100142, China. Email: wangweihu88@163.com; yurong311@aliyun.com.

**Background:** This study aims to investigate the causes of death in patients with esophageal cancer (EC) and report reliable and accurate estimates of adjusted conditional overall survival (COS).

**Methods:** We retrieved data on patients aged 18 years or older who were diagnosed with EC between 1975 and 2016 from the Surveillance Epidemiology End Results (SEER) registry. We estimated COS by using an inverse probability weighting method to adjust for the available covariates. Linear trends were analyzed via a weighted linear regression.

**Results:** A total of 40,142 confirmed patients were included in the final analysis. Of these, 20,971 were diagnosed with esophageal squamous cell carcinoma, and 19,171 were diagnosed with esophageal adenocarcinoma. EC was the leading cause of death, followed by heart disease (5.2%), stomach cancer (3.2%), and other diseases (2.9%). Five-year adjusted COS, given that patients had already survived 36 months was improved from 0.50 to 0.87 ($P_{Difference}$<0.001) in comparison with the 5-year conventional overall survival. As the survival time increased from 0 to 48 months, the 5-year adjusted COS improved significantly. The adjusted conditional EC-specific survival also showed similar trends.

**Conclusions:** For patients with EC who had survived over time, the 5-year adjusted COS and the conditional EC-specific survival improved dramatically. Other causes of death in EC patients should also be considered.

**Keywords:** Causes of death; adjusted conditional survival; esophageal cancer (EC)

## Introduction

Esophageal cancer (EC) is the ninth most common cancer and the sixth leading cause of cancer death globally, with 572,000 new cases [95% uncertainty intervals (UI): 552,100–592,700) and 508,600 deaths (95% UI: 492,500–252,200) in 2015, respectively (1). The global incidence rate has increased remarkably in the past few decades, even though there is a significant regional variation (2). Esophageal squamous cell cancer (ESCC) remains the

---

^ ORCID: 0000-0003-1778-2013.

Page 2 of 13

Deng et al. Conditional overall survival of EC

most frequent histological type, particularly in the "Asian Esophageal Cancer Belt" (which extends from northern Iran, east to China, and north to Russia) (3,4); whereas esophageal adenocarcinoma (EAC) is the most prevalent histological type in developed countries (e.g., UK, North America, Western Europe, and Australia) (5,6). The etiology of EC remains unknown, and obesity and gastroesophageal reflux disease (which are associated with EAC), alcohol consumption, and tobacco use (which are associated with ESCC) are thought to be the major risk factors (5,6).

However, the 5-year overall survival in patients with EC is poor, ranging from 9% to 22% (4,7-10). It is evident that these estimates are rather pessimistic because overall survival derived from the Kaplan-Meier estimator becomes less relevant over time after diagnosis (11,12). In such a case, a useful quantity that addresses this issue is the conditional overall survival (COS), which is defined as the survival probability calculated after a pre-specified length of survival (11). In general, COS estimates increase as the number of years survived increases, and such a relationship is usually even more striking for patients with advanced-stage cancers (12).

It has been widely acknowledged that studies of COS may provide a more reliable and accurate measurement of overall survival in patients, which may benefit those surviving patients and their physicians (11,12). For example, as a surviving patient, knowing about his/her life expectancy if he/she had survived certain years may have an impact on reducing the level of anxiety and its effects on the quality of life. This information will also benefit physicians for choosing appropriate treatment options, subsequent surveillance, and advanced management strategies, playing a vital role in monitoring trends in prognosis and commissioning appropriate services at the population level. Moreover, such a method has already been used for various cancer types, including brain tumor (13), head and neck cancer (14), colon cancer (15), glioblastoma (16), primary retroperitoneal sarcoma (17), renal cell carcinoma (18), and other cancers (19).

The purpose of this study was (I) to provide insights into the causes of death for patients with EC, (II) to evaluate the adjusted COS in patients with EC, and (III) to explore linear trends in COS for EC patients over time by using the population-based Surveillance, Epidemiology, and End Results (SEER) Cancer Statistics Review (CSR) 1975–2016 data set. We present the following article in accordance with the STROBE reporting checklist (available at http://dx.doi.org/10.21037/atm-20-2798).

## Methods

### Design and setting

We performed a population-based study using SEER CSR 1975–2016 data set released by the National Cancer Institute in April 2019 (20). The SEER CSR 1975–2016 reports the most recent cancer incidence, mortality, survival, prevalence, and lifetime risk statistics from the population-based registries, which covers ~26% of the total population in the United States from 14 regions (greater California, greater Georgia, Metropolitan Atlanta, Metropolitan Detroit, Connecticut, Hawaii, Iowa, New Mexico, Settle (Puget Sound), Utah, Alaska, Kentucky, Louisiana, New Jersey) (20). We retrieved data on patient demographics, clinical observations, and prescriptions from the SEER program. All data used in this study are publicly available (https://seer.cancer.gov/).

### Study population

Patients with either squamous cell carcinoma (ICD-O-3 codes 8050-8082) or adenocarcinoma (ICD-O-3 codes 8140-8573) of the esophagus (ICD-O-3 for topography codes: C150-C155, C158-C159) diagnosed between January 1975 and December 2016 were included in this study (*Figure 1*). Other inclusion criteria included (I) age at diagnosis being 18 years or older, and (II) with a diagnostic confirmation (based on either positive histology or positive cytology) of EC in the primary record. Exclusion criteria included (I) age at diagnosis younger than 18 years; (II) unspecified ethnicity code; (III) lack of either primary site code, histological code, or the pathological confirmation of positive histology or cytology; (IV) the misspecified SEER stage of localized/regional/distant; and (V) the absence of overall survival information.

We extracted demographical characteristics including age at diagnosis, sex, ethnicity, marital status at diagnosis, clinical features, including tumor grade and stage, histological type, year at diagnosis, and treatment modalities (e.g., radiation therapy, chemotherapy, and surgery) from SEER CSR 1975–2016 data set. Furthermore, the causes of death among EC patients were defined as the SEER recode, including both cancer- and non-cancer-related deaths before December 2016.

### Case definition

We defined a patient with EC based on the pathological confirmation (either histology or cytology) of squamous
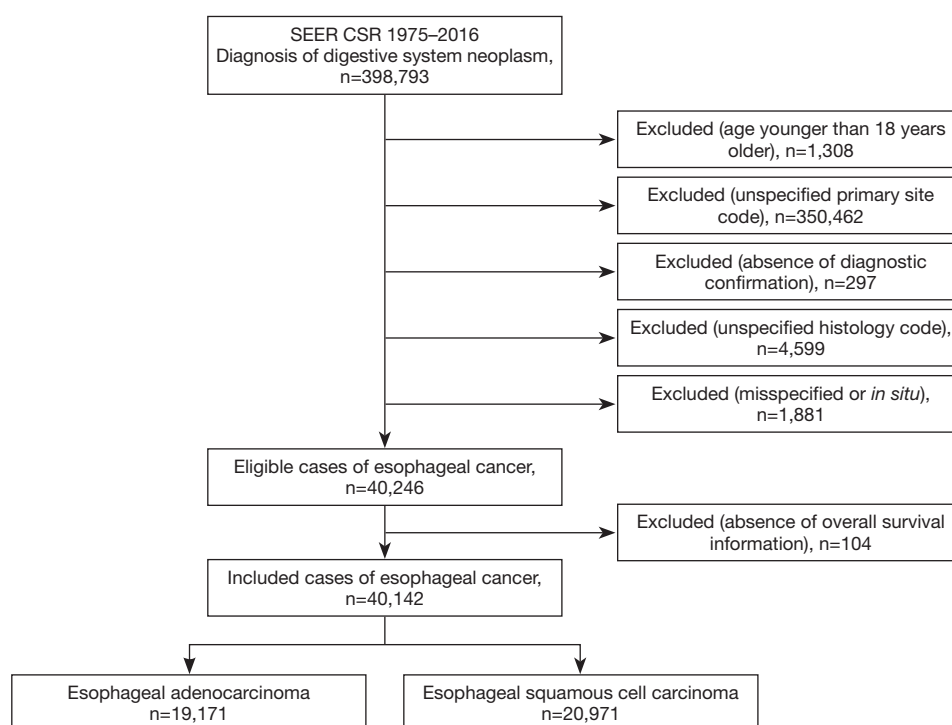
**Figure 1** Study flowchart. SEER, Surveillance Epidemiology End Results; CSR, cancer statistics review.

cell carcinoma or adenocarcinoma of the esophagus in the primary record.

### Outcomes

The primary outcome was all causes of death. The secondary outcome was death due to EC. Adjusted COS according to the histological type (i.e., ESCC and EAC) were reported separately.

### Statistical analysis

Data on EC patients available in SEER CSR 1975–2016 data set were used in this study. Overall survival was measured from the date of diagnosis to any causes of death. A patient was recorded as censored if he/she was alive at the end of December 2016. The median follow-up time was calculated using the reverse Kaplan-Meier method (21).

Hazard ratios (HRs) and 95% confidence interval (CI) were estimated using Cox proportional-hazards model to assess the effects of these publicly available covariates, including age at diagnosis, sex, ethnicity, marital status, tumor location, year of diagnosis, tumor grade, stage, and treatment modalities, whose performance was also evaluated

by the Harrell concordance index (C-index) and 95% CI (22). A concordance index of 0.5 represents a random chance, whereas the larger the C-index, the better performance of the model. We further evaluated its performance using the calibration curve with Efron's bootstrap method (b=500 resamples) and at least 4,500 patients per interval, based on which patients were grouped by the model-predicted 5-year overall survival. The proportional-hazards assumption was verified graphically using the Schoenfeld residual (23). No statistically significant linear trends over time were observed for those covariates included in the model (Figures S1,S2). In addition, adjusted sub-distributed hazard ratios (SDHRs) and 95% CI were obtained from the Fine-Gary model (24), to assess the hazards for dying from a specific cause in patients with ESCC in comparison with those with EAC (25-27). Lastly, the linear trend test was conducted for these estimated SDHRs via a weighted linear regression model, in which the weights were proportional to the inverse of the variance of the SDHRs.

COS represents the probability of surviving an additional number of years (y), given that a patient had already survived certain years (x) (11,12). For example, the y-year COS for patients who had survived x years can be estimated using the (x + y)-year survival divided by the x-year survival.

Hence, when x =0, the COS is exactly the conventional survival derived from the Kaplan-Meier estimator (28). The 95% CI of COS can also be calculated using a variation of the standard Greenwood formula, as described by Davis *et al.* (13). Adjusted COS was estimated using the inverse-probability (IP) weighting method (29,30). The IP weighting removes confounding by creating a pseudo-population, in which the probability of death from EC is predicted via a logistic regression given age at diagnosis, sex, ethnicity, marital status, tumor grade and stage, location, year at diagnosis, and treatment modalities, referred as to $\widehat{\Pr}[EC=1|L]$. Thus, the IP weighting for the patients with death from EC was $1/\widehat{\Pr}[EC=1|L]$, and that for others was $1/\left(1-\widehat{\Pr}[EC=1|L]\right)$. We further investigated potentially linear trends in COS over time by employing a weighted linear regression of COS to the year of diagnosis. The weights were inversely proportional to the variance of COS. Moreover, covariate-specific adjusted COS was reported with additional adjustment for other available covariates in the SEER CSR 1975–2016 data set.

All statistical analyses were carried out in R software (version 3.6.1) (31). Results were considered significant if a two-sided P<0.05 was obtained. Lastly, the current study was a secondary analysis of SEER CSR 1975–2016 data set and was exempt from review by the institutional review board.

### Patient and public involvement

No participants were involved in either the design or the implementation of the study, nor were participants engaged in analyzing or describing the results.

## Results

A total of 398,793 patients were diagnosed with digestive system neoplasm between January 1975 and December 2016. After stepwise case exclusion, as detailed in *Figure 1*, a total of 40,142 patients, including 19,171 patients who were diagnosed with adenocarcinoma and 20,971 patients who were diagnosed with squamous cell carcinoma, were analyzed. Patients were grouped into 18–39 years (n=389), 40–49 years (n=2,398), 50–59 years (n=7,881), 60–69 years (n=12,580), 70–79 years (n=10,973), and 80+ years (n=5,921) based on the age at diagnosis. The mean age was 66.89 (standard deviation =11.06) years. The median follow-up time was 139 months (interquartile range: 74–222 months), 12,176 (30.3%) patients had survived over 12 months, 6,145 (15.3%) patients had survived over 24 months, 3,961

(9.87%) had survived over 36 months, and 2,866 (7.14%) had survived over 48 months.

*Table 1* presents baseline characteristics, clinical observations, and treatment modalities of the study population according to the histological type of EC. The bootstrap-corrected concordance index was 0.698 (95% CI: 0.695–0.701), indicating a well-fitted Cox model. The calibration curve also suggested a good agreement between the predicted and the observed 5-year overall survival, as shown in Figure S3. Men (85.3% *vs.* 66.1%), white patients (95.2% *vs.* 64.8%), and married patients (62.9% *vs.* 48.3%) were more likely to be diagnosed with EAC, with poorly differentiated (43.5% *vs.* 34.7) and distant stage (35.5% *vs.* 26.0%) in comparison with ESCC patients. The proportion of EAC increased over the year of diagnosis, whereas the proportion of ESCC slightly decreased. In addition, almost half of the enrolled patients (48.8%) received chemotherapy. Patients with EAC more frequently received chemotherapy than those with ESCC (54.0% *vs.* 44.0%). In contrast, patients with ESCC were more likely to receive radiation therapy (64.9% *vs.* 52.3%). Of note, ~14.7% of patients underwent surgery, and most of them were EAC patients (76.8% *vs.* 23.2%).

### Causes of death

*Figure 2* shows the causes of death in patients with EC. As expected, EC (75.2%, n=27,678) was the leading cause of death, followed by heart disease (5.2%, n=1,918), stomach cancer (3.2%, n=1,191), and other diseases (2.9%, n=881). Compared with patients diagnosed with EAC, patients with ESCC had almost 11-fold (SDHR =12.55, 95% CI: 7.60–20.74) increased hazards for dying from nephropathy, and more than 6-fold (SDHR =7.47, 95% CI: 4.61–12.09) increased hazards for dying from other non-cancer diseases after adjustment for age at diagnosis, sex, and treatment modalities. Figure S4 presents the change of the SDHR and 95% CI over the period of diagnosis from 1995–1999 to 2015–2016. During this time, the hazards for dying from other respiratory system cancer ($P_{Trend}$=0.030) significantly increased in patients with ESCC in comparison with those with EAC, whereas the hazards for dying from nephropathy decreased ($P_{Trend}$=0.094). Furthermore, linear increased hazards for dying from heart disease ($P_{Trend}$=0.139), oral/pharynx cancer ($P_{Trend}$=0.326), and suicide/injury ($P_{Trend}$=0.835) were observed, although these linear trends did not reach a statistically significant level of 0.05. The reason may lie in competing risks due to other causes of death; that is, the occurrence of other deaths precludes the

**Table 1** Baseline characteristics, clinical observations, and treatment modalities of the study population according to the histology type of esophageal cancer

| Characteristic | Overall | Esophageal cancer | | Adjusted HR* (95% CI) | P_Trend |
| --- | --- | --- | --- | --- | --- |
| | | Adenocarcinoma | Squamous cell cancer | | |
| Total number | 40,142 | 19,171 | 20,971 | 1.06 (1.03–1.09) | |
| Sex [female, n (%)] | 9,924 (24.7) | 2,816 (14.7) | 7,108 (33.9) | 0.88 (0.86–0.90) | – |
| Age at diagnosis [years, mean (SD)] | 66.89 (11.60) | 67.06 (12.04) | 66.74 (11.18) | | – |
| Age at diagnosis group, n (%) | | | | | 0.006 |
| 18–39 | 389 (1.0) | 270 (1.4) | 119 (0.6) | 1 | |
| 40–49 | 2,398 (6.0) | 1,209 (6.3) | 1,189 (5.7) | 1.40 (1.25–1.58) | |
| 50–59 | 7,881 (19.6) | 3,633 (19.0) | 4,248 (20.3) | 1.49 (1.33–1.67) | |
| 60–69 | 12,580 (31.3) | 5,765 (30.1) | 6,815 (32.5) | 1.67 (1.49–1.87) | |
| 70–79 | 10,973 (27.3) | 5,189 (27.1) | 5,784 (27.6) | 1.92 (1.71–2.16) | |
| 80+ | 5,921 (14.8) | 3,105 (16.2) | 2,816 (13.4) | 2.45 (2.18–2.75) | |
| Ethnic group, n (%) | | | | | |
| White | 31,826 (79.3) | 18,246 (95.2) | 13,580 (64.8) | 1 | – |
| Black | 5,928 (14.8) | 440 (2.3) | 5,488 (26.2) | 1.15 (1.11–1.18) | |
| Other | 2,388 (5.9) | 485 (2.5) | 1,903 (9.1) | 0.95 (0.91–0.99) | |
| Marital status, n (%) | | | | | 0.138 |
| Single (never married) | 5,215 (13.0) | 2,173 (11.3) | 3,042 (14.5) | 1 | |
| Married (including common law) | 22,203 (55.3) | 12,064 (62.9) | 10,139 (48.3) | 0.82 (0.79–0.84) | |
| Other/unknown | 1,694 (4.2) | 746 (3.9) | 948 (4.5) | 0.84 (0.80–0.90) | |
| Separated/divorced/widowed | 11,030 (27.5) | 4,188 (21.8) | 6,842 (32.6) | 0.96 (0.92–0.99) | |
| Tumor grade, n (%) | | | | | 0.546 |
| Well differentiated | 2,207 (5.5) | 949 (5.0) | 1,258 (6.0) | 1 | |
| Moderately differentiated | 13,223 (32.9) | 5,759 (30.0) | 7,464 (35.6) | 1.22 (1.16–1.28) | |
| Poorly differentiated | 15,615 (38.9) | 8,332 (43.5) | 7,283 (34.7) | 1.37 (1.31–1.44) | |
| Undifferentiated | 688 (1.7) | 370 (1.9) | 318 (1.5) | 1.28 (1.17–1.40) | |
| Cell type not determined | 8,409 (20.9) | 3,761 (19.6) | 4,648 (22.2) | 1.15 (1.09–1.21) | |
| Tumor stage, n (%) | | | | | 0.037[†] |
| Localized | 10,285 (25.6) | 4,531 (23.6) | 5,754 (27.4) | 1 | |
| Regional | 11,636 (29.0) | 5,760 (30.0) | 5,876 (28.0) | 1.52 (1.48–1.57) | |
| Distant | 12,266 (30.6) | 6,813 (35.5) | 5,453 (26.0) | 2.77 (2.68–2.85) | |
| Unknown stage | 5,955 (14.8) | 2,067 (10.8) | 3,888 (18.5) | 1.52 (1.47–1.57) | |

**Table 1** *(continued)*

Page 6 of 13

Deng et al. Conditional overall survival of EC

**Table 1** *(continued)*

| Characteristic | Overall | Esophageal cancer | | Adjusted HR* (95% CI) | P$_{Trend}$ |
| --- | --- | --- | --- | --- | --- |
| | | Adenocarcinoma | Squamous cell cancer | | |
| Year at diagnosis group, n (%) | | | | | 0.027 |
| 1975–1979 | 3,127 (7.8) | 422 (2.2) | 2,705 (12.9) | 1 | |
| 1980–1984 | 3,561 (8.9) | 630 (3.3) | 2,931 (14.0) | 0.92 (0.88–0.96) | |
| 1985–1989 | 4,146 (10.3) | 1,154 (6.0) | 2,992 (14.3) | 0.86 (0.82–0.90) | |
| 1990–1994 | 4,601 (11.5) | 1,783 (9.3) | 2,818 (13.4) | 0.81 (0.78–0.85) | |
| 1995–1999 | 5,127 (12.8) | 2,593 (13.5) | 2,534 (12.1) | 0.87 (0.83–0.91) | |
| 2000–2004 | 5,595 (13.9) | 3,300 (17.2) | 2,295 (10.9) | 0.90 (0.86–0.95) | |
| 2005–2009 | 6,038 (15.0) | 3,907 (20.4) | 2,131 (10.2) | 0.79 (0.76–0.83) | |
| 2010–2014 | 6,496 (16.2) | 4,390 (22.9) | 2,106 (10.0) | 0.73 (0.70–0.77) | |
| 2015–2016 | 1,451 (3.6) | 992 (5.2) | 459 (2.2) | 0.69 (0.64–0.75) | |
| Tumor location [primary site, n (%)] | | | | | |
| C150 (cervical esophagus) | 1,230 (3.1) | 57 (0.3) | 1,173 (5.6) | 1 | 0.688 |
| C151 (thoracic esophagus) | 1,393 (3.5) | 221 (1.2) | 1,172 (5.6) | 1.10 (1.02–1.19) | |
| C152 (abdominal esophagus) | 334 (0.8) | 231 (1.2) | 103 (0.5) | 0.99 (0.88–1.13) | |
| C153 (upper 3rd of esophagus) | 2,769 (6.9) | 244 (1.3) | 2,525 (12.0) | 1.14 (1.06–1.22) | |
| C154 (middle 3rd of esophagus) | 8,835 (22.0) | 1,417 (7.4) | 7,418 (35.4) | 1.17 (1.10–1.24) | |
| C155 (lower 3rd of esophagus) | 20,202 (50.3) | 14,736 (76.9) | 5,466 (26.1) | 1.04 (0.97–1.10) | |
| C158 (overlapping lesion of esophagus) | 1,850 (4.6) | 702 (3.7) | 1,148 (5.5) | 1.23 (1.14–1.33) | |
| C159 (esophagus, NOS) | 3,529 (8.8) | 1,563 (8.2) | 1,966 (9.4) | 1.11 (1.04–1.19) | |
| Treatment modalities [yes, n (%)] | | | | | |
| Chemotherapy | 19,578 (48.8) | 10,361 (54.0) | 9,217 (44.0) | 0.66 (0.65–0.68) | – |
| Radiation therapy | 23,630 (58.9) | 10,019 (52.3) | 13,611 (64.9) | 0.86 (0.84–0.89) | – |
| Surgery | 5,919 (14.7) | 4,543 (23.7) | 1,376 (6.6) | 0.42 (0.41–0.44) | – |

*, adjusted hazard ratio (HR) and 95% confidence interval (CI); †, linear trend test was performed after excluding those patients with unknown tumor stage.

occurrence of EC and vice versa. However, the mechanistic link between EC-related death and other causes of death is far from certain. Last but not least, the hazards for dying from ESCC decreased over the year of diagnosis in comparison with EAC, as detailed in Figure S4.

## COS

The median survival time to all causes of death was 9 months (95% CI: 8–9 months). The crude and adjusted

COS according to the histological type for patients given that patients had already survived 0 to 48 months are shown in *Figure 3*. COS varied for those patients who had already survived certain years, and similar patterns were observed in patients with ESCC and EAC. After adjustment for age at diagnosis, sex, ethnicity, marital status, tumor location, histological type, tumor grade and stage, year of diagnosis, and treatment modalities, the COS given that patients had already survived 36 months was significantly improved at 1 year (0.95, 95% CI: 0.95–0.95 *vs.* 0.65, 95% CI: 0.65–0.65,
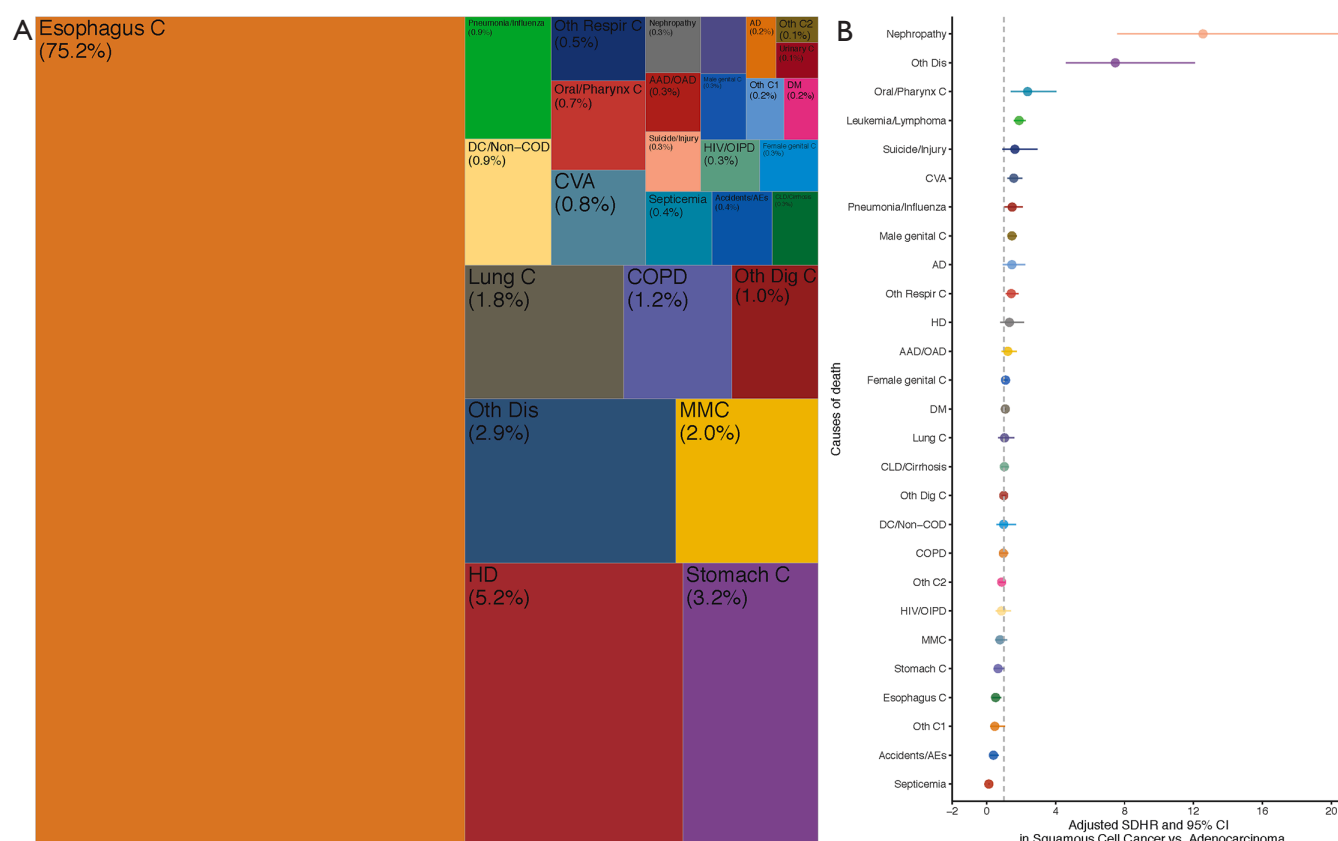
**Figure 2** Causes of death for patients with EC. (A) Tree plot of causes of death with their proportion; (B) the adjusted SDHR and 95% CI of esophageal squamous cell cancer in comparison with esophageal adenocarcinoma. EC, esophageal cancer; SDHR, subdistribution hazard ratio; Esophagus C, squamous cell carcinoma or adenocarcinoma of the esophagus; HD, diseases of heart; Stomach C, stomach cancer; Oth Dis, other diseases (non-cancer) of death, including unknown behavior neoplasm, tuberculosis, syphilis, hypertension without heart disease, stomach and duodenal ulcers, complications of pregnancy, childbirth, puerperium, congenital anomalies, certain conditions originating in perinatal period, symptoms, signs and ill-defined conditions, homicide and legal intervention; MMC, miscellaneous malignant cancer, including mesothelioma and kaposi sarcoma; Lung C, lung cancer; COPD, chronic obstructive pulmonary disease; Oth Dig C, other digestive system cancer, including small intestine, colon, rectum, anus, anal canal, anorectum and others; DC/non-COD, state DC not available or state DC available but no cause of death; CVA, cerebrovascular diseases; Accidents/AEs, accidents and adverse effects; HIV/OIPD, human immunodeficiency virus (HIV) and other infectious and parasitic diseases; Female genital C, female genital cancer; Oral/Pharynx C, oral/pharynx cancer; AAD/OAD, atherosclerosis, aortic aneurysm and dissection, and other diseases of arteries, arterioles, capillaries; Male genital C, male genital cancer; Oth C1, bones and joints, soft tissue (including heart), skin system cancer and myeloma; DM, diabetes mellitus; other Respir C, other respiratory system cancer, including larynx and others; CLD/Cirrhosis, chronic liver disease and cirrhosis; AD, Alzheimer's disease; Oth C2, eye, orbit, endocrine brain and other nervous system cancer; and Urinary C, urinary cancer.

$P_{Difference}<0.001$), 3 years (0.90, 95% CI: 0.90–0.90 *vs.* 0.50, 95% CI: 0.49–0.50, $P_{Difference}<0.001$), and 5 years (0.87, 95% CI: 0.86–0.87 *vs.* 0.46, 95% CI: 0.45–0.46, $P_{Difference}<0.001$) in comparison with the conventional adjusted overall survival. Furthermore, it is worth noting that patients with EAC had a better prognosis than those with ESCC, whereas this phenomenon disappeared and even reversed for the

adjusted COS, especially for those patients who had already survived at least 12 months (*Figure 3*).

Furthermore, the adjusted 1-, 3-, and 5-year COS, given that patients had survived 36 months, were improved by 27.7% ($P_{Trend}=0.241$, $P_{Difference}<0.001$), 39.1% ($P_{Trend}=0.135$, $P_{Difference}<0.001$), and 112.0% ($P_{Trend}=0.093$, $P_{Difference}<0.001$) from 1975–1979 to 2005–2009, respectively. Figures S5-S12

**Page 8 of 13**

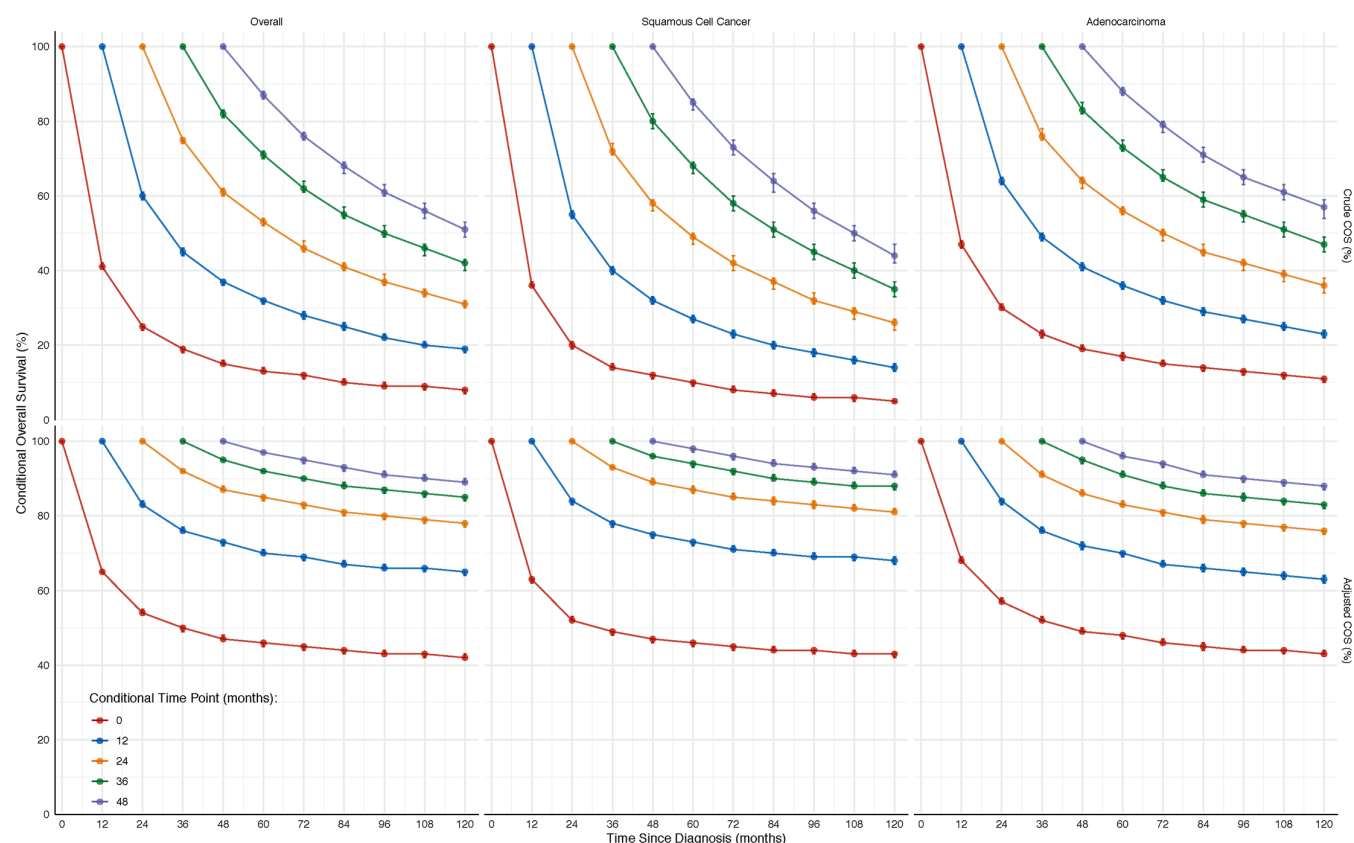Deng et al. Conditional overall survival of EC

**Figure 3** Crude (upper panel) and adjusted COS (bottom panel) stratified by histology for patients given that patients had already survived 0 to 48 months. COS, conditional overall survival.

presents the covariate-specific COS given that patients had survived 36 months with additional adjustment of all other available covariates. Notably, age at diagnosis (Figure S5), sex (especially for ESCC, Figure S6), and treatment modalities (especially for surgery, Figure S7) were significant determinants for the adjusted COS. In contrast, ethnicity (Figure S8), marital status (Figure S9), tumor grade (Figure S10) and stage (Figure S11), and year of diagnosis (Figure S12) were not statistically significant determinants for COS after adjustment for other available covariates in the SEER CSR 1975–2016 data set.

### Conditional EC-specific survival

The median survival time to death related to EC with other causes of death being censored was 11 months (95% CI: 11–11 months), and the median follow-up period was 57 months (95% CI: 55–58 months). *Figure 4* shows the crude and adjusted conditional EC-specific survival,

given that patients had already survived 0, 12, 24, 36, and 48 months. The conditional EC-specific survival significantly improved in comparison with that of the conventional EC-specific survival. This phenomenon still existed after adjustment for age at diagnosis, sex, ethnicity, marital status, tumor location, histological type, tumor grade and stage, year of diagnosis, and treatment modalities. Moreover, compared with ESCC patients, patients with EAC had improved prognosis, but this pattern reversed, especially for the adjusted COS given that patients had already survived at least 12 months, as shown in *Figure 4*. Finally, the adjusted 1-, 3-, and 5-year conditional EC-specific survival, as shown in Figure S13, was improved by 6.0% ($P_{Trend}$=0.222, $P_{Difference}$<0.001), 2.6% ($P_{Trend}$=0.806, $P_{Difference}$=0.003), and 51.0% ($P_{Trend}$=0.421, $P_{Difference}$<0.001) from 1975–1979 to 2005–2009, respectively. The age at diagnosis-, sex-, tumor grade- and stage-, surgery-, and conditional EC-specific survival given patients that had survived 36 months are presented
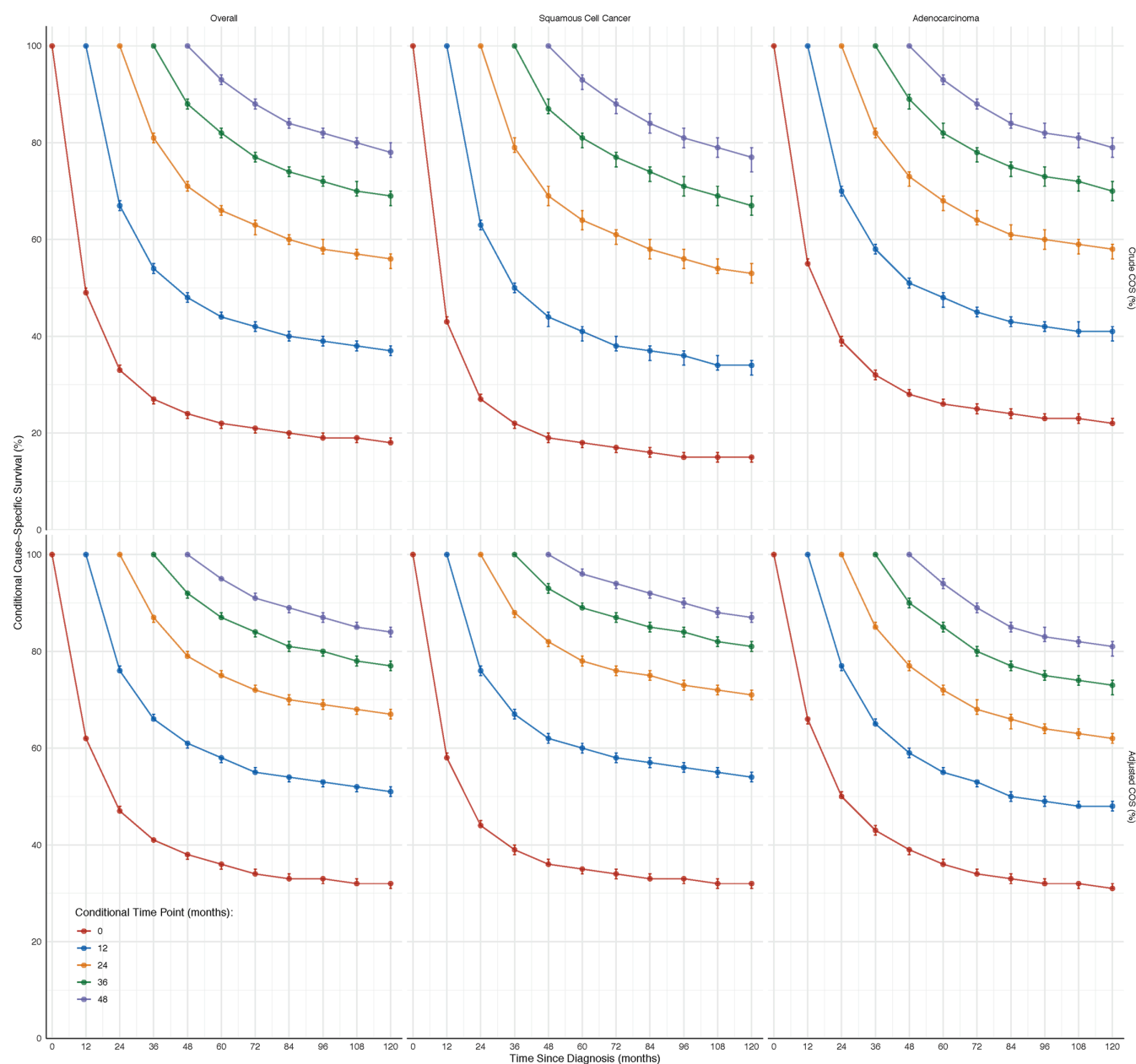
**Figure 4** Crude (upper panel) and adjusted conditional EC-specific survival (bottom panel) stratified by histology for patients given that patients had already survived 0 to 48 months. EC, esophageal cancer.

## Discussion

### *Principal findings*

This population-level based study investigates the causes of death in patients with EC aged 18 years or older in SEER CSR 1975–2016, among which EC is the leading cause of death, followed by heart disease, stomach cancer, and other diseases. Moreover, the adjusted COS for patients improved significantly over time, given that patients had survived certain years since diagnosis (e.g., 36 months). We did observe linear trends in adjusted COS over the year of diagnosis.

### Comparison with other studies

Traditionally, death from EC is considered as the primary outcome with other causes of death being censored, and the Kaplan-Meier estimator or cause-specific proportional-hazards model are used to perform the subsequent analysis (4,32-34). However, these analyses should be used with caution. The results derived from these conventional methods are often subjected to competing bias, which may further reverse the effects (25-27,34). For example, a study of causes of death in patients with EC in a population-based study in Sweden showed that the hazards of dying from EC decreased over time with the consideration of competing events. Of all causes of death, EC accounted for 79.5% of all deaths, with non-EC accounting for 9.8%, ischemic heart disease or cerebrovascular disease accounting for 4.2%, and respiratory diseases accounting for 1.3% (35). This conclusion was consistent with our results, suggesting other causes of death among EC patients should also be considered.

Furthermore, there are considerable variations in previous EC survival estimates. For example, a previous study (10) using SEER 1973–2009 data set reported that the overall 5-year survival rate was 9–22% in all patients. In contrast, for patients who underwent surgery, the 5-year overall survival rate significantly increased with an estimated cure rate ranging from 25% to 73%, depending on the tumor stage. Similar conclusions were also observed in the nationwide Swedish cohort study (36) and the population-based study in Ontario, Canada (37). Ethnic disparities in EC-related survival were also reported, in which black patients with ESCC were found to have a poor prognosis in comparison with white and Asian/Pacific Islander race patients, even though these disparities decreased over time (38). The United States Cancer Statistics in 2018 reported that the 5-year overall relative survival of EC was 19% (2008 to 2014), varying considerably from 5% to 45% depending on the tumor stage (39). In addition, a hospital-based pooled analysis (2000 to 2018) in China reported that the 5-year overall survival was around 40%, and increased over time from 2000 to 2018 (40). However, these conclusions were drawn without an additional adjustment for the possibly critical prognostic factors. Therefore, it is not surprising that these estimates were inconsistent with our results.

### Strengths and limitations of study

Based on the large, population-based SEER CSR 1975–2016 database, we identified 40,142 patients diagnosed with EC between January 1975 and December 2016, which allowed for reliable analyses of subgroups and trends in survival after diagnosis.

A unique feature of this study is the use of adjusted COS based on IP weighting method. The IP weighting accounts for these important prognostic factors, such as age at diagnosis, sex, ethnicity, marital status, tumor location, grade, stage, year of diagnosis, and treatment modalities, providing more reliable and accurate estimates. Furthermore, adjusted COS, providing valuable information to physicians and health practitioners, may serve a meaningful role in cancer survivorship research.

Nevertheless, there remain limitations to this study. First, the calculation of COS requires a relative long-term follow-up because probabilities are based on expected survival after a pre-specified period of survival. Consequently, the long-term COS in patients with EC diagnosed after the year of 2010 in this study could not be estimated. Second, the adjusted COS may reach a plateau after 96 months follow-up (*Figure 3* left-bottom), which suggests that there may exist a small proportion of patients being cured (41). For example, the plateau occurs after 72 months of follow-up given that patients have survived at least 12 months, but this time point varies. The cure rate provides crucial information for making a personalized surveillance strategy and has been advocated as a potentially important endpoint in clinical trial designs (42).

Third, the treatment modalities (e.g., chemotherapy, radiation therapy, and surgery), inheriting the drawbacks of the retrospective or surveillance study, suffer the missing value issues. For example, only ~14.7% of EC patients received surgical treatment. In such a case, non-misclassification may happen. However, its effect would be possibly minimal, as the missing values are more likely to occur in older cases due to the administrative decision (e.g., cases were diagnosed before 1995). Furthermore, immortal-time bias (which refers to the period of followed time during which death cannot occur) tends to overestimate COS because only those survival patients would receive the treatment and thus could be included in the final analysis. However, its magnitude remains unknown. Moreover, treatment options often change over time. In such a case, the inclusion of the time-varying effects of treatment modalities might improve the accuracy of our results. Another issue here is that different treatment options may cause other long-term adverse effects, such as cardiovascular events, which should also be considered.

Finally, other important prognostic factors (4-6,43), such as cancer screening, tobacco use, and alcohol consumption, were thought to play critical roles in EC patients' prognosis and were not included in the final analysis due to the limited data in SEER CSR 1975–2016. Ignoring these factors may further result in biased COS estimates. For example, smoking and alcohol consumption are positively associated with EC and death. Failure to adjust for them might lead to an overestimation of COS and the causal-specific hazards. Furthermore, consider smoking rates have been reduced in the United States since the 1960s, from 42.4% to 14.0% in adults. However, we could consider these in our analysis. The inclusion of these effects may improve the accuracy of COS and cause-specific hazards. Finally, our results were derived from the SEER CSR 1975–2016 data set, in which cause-of-death coding or other recording errors may occur. External validation of these estimates is thus necessary before they are uniformly accepted into clinical practice.

With the steady increase in the number of long-term cancer survivors, there is an increasing need for insightful understandings of the causes of death and reliable and accurate estimates of survival. In this study, we found other causes of death (except EC) in patients diagnosed with EC should also be considered to reduce the potential competing bias. We observed an improvement in adjusted COS over time, which is encouraging. However, the outlook of EC remains poor, particularly for these patients without appropriate treatment modalities. Primary care-led research (e.g., cancer screening) is also required to develop and construct good management guideline for patients with EC, to achieve a better prognosis for patients.

## Acknowledgments

## Footnote

*Reporting Checklist:* The authors have completed the STROBE reporting checklist. Available at http://dx.doi.org/10.21037/atm-20-2798

*Conflicts of Interest:* All authors have completed the ICMJE

uniform disclosure form (available at http://dx.doi.org/10.21037/atm-20-2798). The authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. Because the study used preexisting data with no personal identifiers, this study was exempt from review by the institutional review board.

## References

1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394-424.
2. Zhang GH, Fujita H, Yamana H, et al. A prediction of hospital mortality after surgical treatment for esophageal cancer. Surg Today 1994;24:122-7.
3. Zhang Y. Epidemiology of esophageal cancer. World J Gastroenterol 2013;19:5598-606.
4. Wei WQ, Chen ZF, He YT, et al. Long-Term Follow-Up of a Community Assignment, One-Time Endoscopic Screening Study of Esophageal Cancer in China. J Clin Oncol 2015;33:1951-7.
5. Vingeliene S, Chan DSM, Vieira AR, et al. An update of the WCRF/AICR systematic literature review and meta-analysis on dietary and anthropometric factors and esophageal cancer risk. Ann Oncol 2017;28:2409-19.
6. Thrumurthy SG, Chaudry MA, Thrumurthy SSD, et al. Oesophageal cancer: risks, prevention, and diagnosis. BMJ 2019;366:l4373.
7. Hiripi E, Jansen L, Gondos A, et al. Survival of stomach and esophagus cancer patients in Germany in the early 21st century. Acta Oncol 2012;51:906-14.
8. Rubenstein JH, Shaheen NJ. Epidemiology, Diagnosis, and Management of Esophageal Adenocarcinoma.

Gastroenterology 2015;149:302-17.e1.

9.   Njei B, McCarty TR, Birk JW. Trends in esophageal cancer survival in United States adults from 1973 to 2009: A SEER database analysis. J Gastroenterol Hepatol 2016;31:1141-6.

10.  Dubecz A, Gall I, Solymosi N, et al. Temporal trends in long-term survival and cure rates in esophageal cancer: a SEER database analysis. J Thorac Oncol 2012;7:443-7.

11.  Zabor EC, Gonen M, Chapman PB, et al. Dynamic prognostication using conditional survival estimates. Cancer 2013;119:3589-92.

12.  Hieke S, Kleber M, Konig C, et al. Conditional Survival: A Useful Concept to Provide Information on How Prognosis Evolves over Time. Clin Cancer Res 2015;21:1530-6.

13.  Davis FG, McCarthy BJ, Freels S, et al. The conditional probability of survival of patients with primary malignant brain tumors: surveillance, epidemiology, and end results (SEER) data. Cancer 1999;85:485-91.

14.  Fuller CD, Wang SJ, Thomas CR, Jr., et al. Conditional survival in head and neck squamous cell carcinoma: results from the SEER dataset 1973-1998. Cancer 2007;109:1331-43.

15.  Chang GJ, Hu CY, Eng C, et al. Practical application of a calculator for conditional survival in colon cancer. J Clin Oncol 2009;27:5938-43.

16.  Polley MY, Lamborn KR, Chang SM, et al. Conditional probability of survival in patients with newly diagnosed glioblastoma. J Clin Oncol 2011;29:4175-80.

17.  Abbott AM, Habermann EB, Parsons HM, et al. Prognosis for primary retroperitoneal sarcoma survivors: a conditional survival analysis. Cancer 2012;118:3321-9.

18.  Harshman LC, Xie W, Bjarnason GA, et al. Conditional survival of patients with metastatic renal-cell carcinoma treated with VEGF-targeted therapy: a population-based study. Lancet Oncol 2012;13:927-35.

19.  Janssen-Heijnen ML, Gondos A, Bray F, et al. Clinical relevance of conditional survival of cancer patients in europe: age-specific analyses of 13 cancers. J Clin Oncol 2010;28:2520-8.

20.  Tsujimoto H, Takahata R, Nomura S, et al. Predictive value of pleural and serum interleukin-6 levels for pneumonia and hypo-oxygenations after esophagectomy. J Surg Res 2013;182:e61-7.

21.  Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. Controlled clinical trials 1996;17:343-6.

22.  Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating

assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15:361-87.

23.  Schoenfeld D. Partial residuals for the proportional hazards regression model. Biometrika 1982;69:239-41.

24.  Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 1999;94:496-509.

25.  Austin PC, Lee DS, Fine JP. Introduction to the Analysis of Survival Data in the Presence of Competing Risks. Circulation 2016;133:601-9.

26.  Schumacher M, Ohneberg K, Beyersmann J. Competing risk bias was common in a prominent medical journal. Journal of Clinical Epidemiology 2016;80:135-6.

27.  van Walraven C, Hawken S. Competing risk bias in Kaplan-Meier risk estimates can be corrected. J Clin Epidemiol 2016;70:101-5.

28.  Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc 1958;53:457-81.

29.  Cole SR, Hernán MA. Adjusted survival curves with inverse probability weights. Comput Methods Programs Biomed 2004;75:45-9.

30.  Therneau TM, Crowson CS, Atkinson EJ. Adjusted survival curves. January; 2015.

31.  R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, 2019.

32.  Li JY, Taylor PR, Li B, et al. Nutrition intervention trials in Linxian, China: multiple vitamin/mineral supplementation, cancer incidence, and disease-specific mortality among adults with esophageal dysplasia. J Natl Cancer Inst 1993;85:1492-8.

33.  Qiao YL, Dawsey SM, Kamangar F, et al. Total and cancer mortality after supplementation with vitamins and minerals: follow-up of the Linxian General Population Nutrition Intervention Trial. J Natl Cancer Inst 2009;101:507-18.

34.  Wang SM, Taylor PR, Fan JH, et al. Effects of Nutrition Intervention on Total and Cancer Mortality: 25-Year Post-trial Follow-up of the 5.25-Year Linxian Nutrition Intervention Trial. J Natl Cancer Inst 2018;110:1229-38.

35.  Xie SH, Wahlin K, Lagergren J. Cause of death in patients diagnosed with esophageal cancer in Sweden: a population-based study. Oncotarget 2017;8:51800-9.

36.  Kauppila JH, Mattsson F, Brusselaers N, et al. Prognosis of oesophageal adenocarcinoma and squamous cell carcinoma following surgery and no surgery in a nationwide Swedish cohort study. BMJ Open 2018;8:e021495.

37.  Akhtar-Danesh N, Shakeel S, Seow H, et al. Trends in

survival based on treatment modality for esophageal cancer: a population-based study. Eur J Gastroenterol Hepatol 2019;31:1192-9.

38. Kim A, Ashman P, Ward-Peterson M, et al. Racial disparities in cancer-related survival in patients with squamous cell carcinoma of the esophagus in the US between 1973 and 2013. PLoS One 2017;12:e0183782.

39. Tepper J, Krasna MJ, Niedzwiecki D, et al. Phase III trial of trimodality therapy with cisplatin, fluorouracil, radiotherapy, and surgery compared with surgery alone for esophageal cancer: CALGB 9781. J Clin Oncol 2008;26:1086-92.

40. Hou H, Meng Z, Zhao X, et al. Survival of Esophageal Cancer in China: A Pooled Analysis on Hospital-Based Studies From 2000 to 2018. Front Oncol 2019;9:548.

41. Yilmaz YE, Lawless JF, Andrulis IL, et al. Insights From Mixture Cure Modeling of Molecular Markers for Prognosis in Breast Cancer. J Clin Oncol 2013;31:2047-54.

42. Wang SJ, Fuller CD, Kim JS, et al. Prediction model for estimating the survival benefit of adjuvant radiotherapy for gallbladder cancer. J Clin Oncol 2008;26:2112-7.

43. Wang GQ, Jiao GG, Chang FB, et al. Long-term results of operation for 420 patients with early squamous cell esophageal carcinoma discovered by screening. Ann Thorac Surg 2004;77:1740-4.
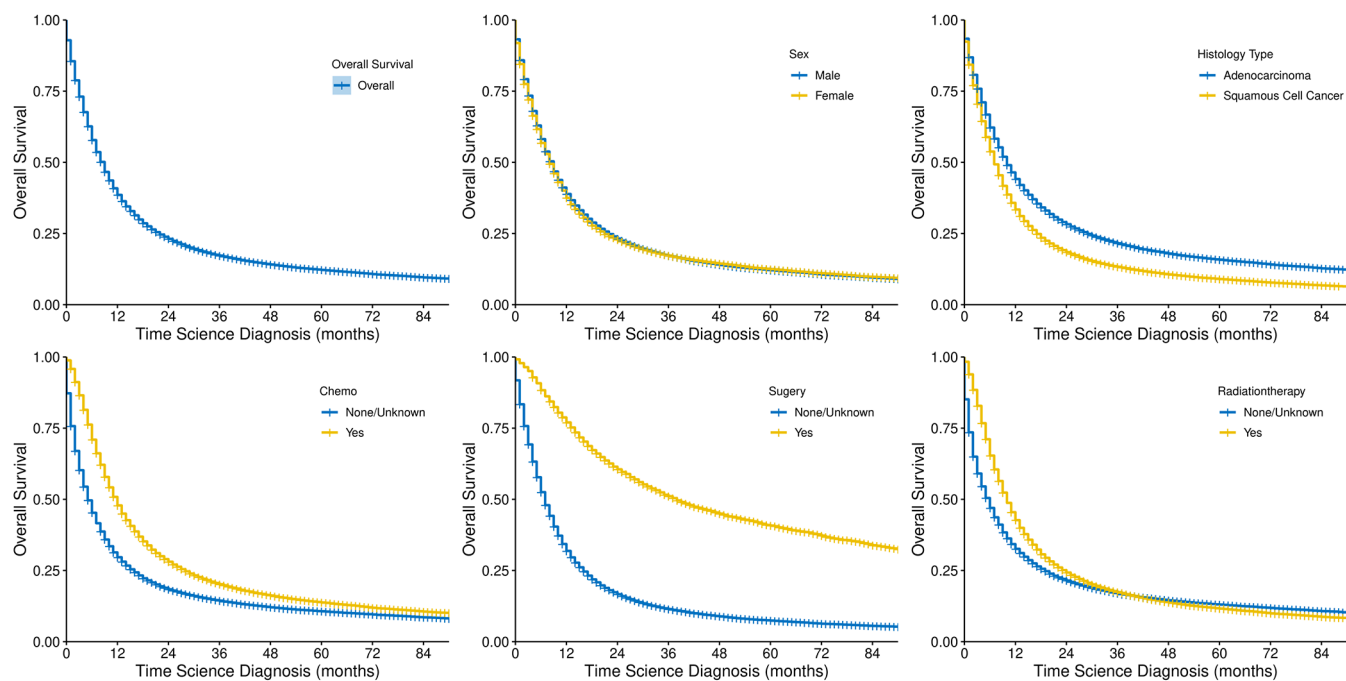
**Figure S1** Kaplan-Meier curves without adjustment for other covariates.
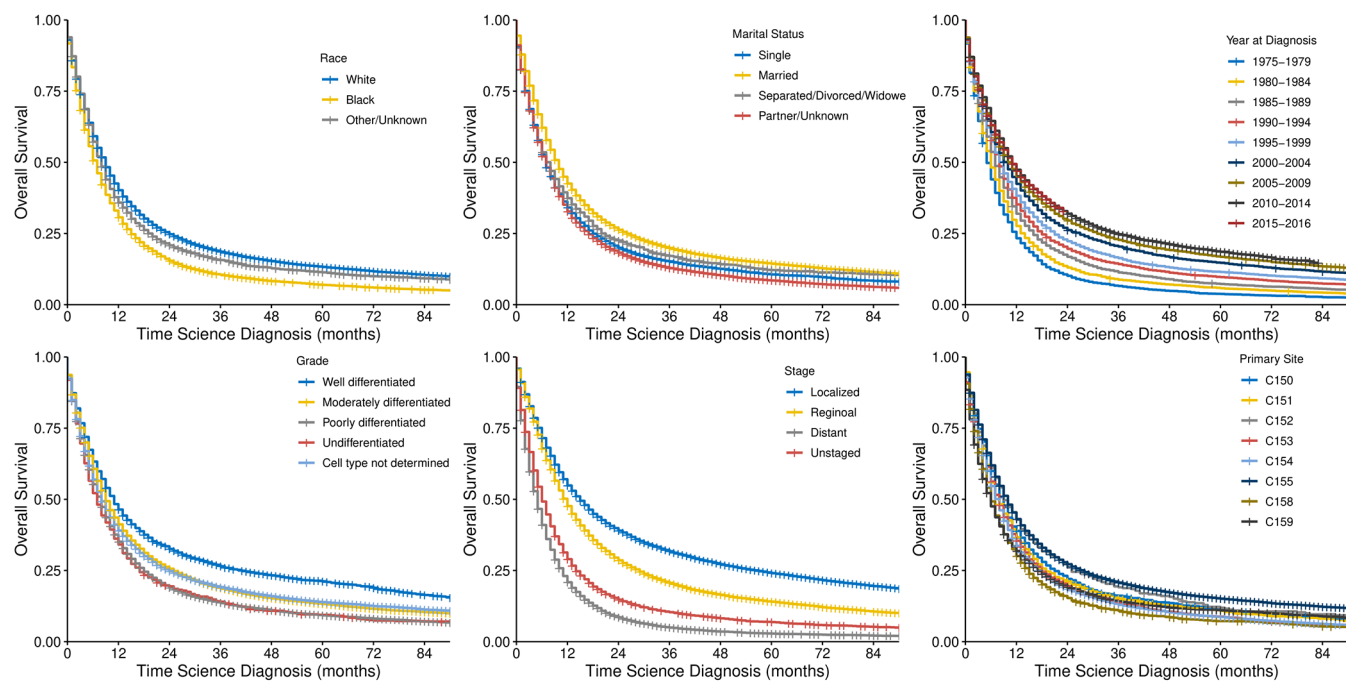


**Figure S2** (Continue) Kaplan-Meier curves without adjustment for other covariates.

http://dx.doi.org/10.21037/atm-20-2798
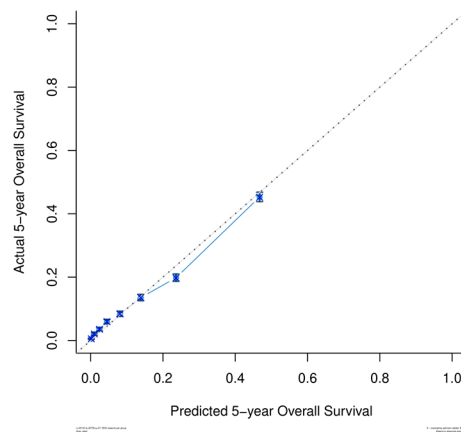
**Figure S3** Overfitting-corrected calibration plot with Efron's bootstrap and with 500 resamples and at least 4,500 patients per interval. Solid blue line represents bias-corrected calibration. Dashed dot line represents the perfect calibration in which predicted outcome corresponds perfectly with actual ones.
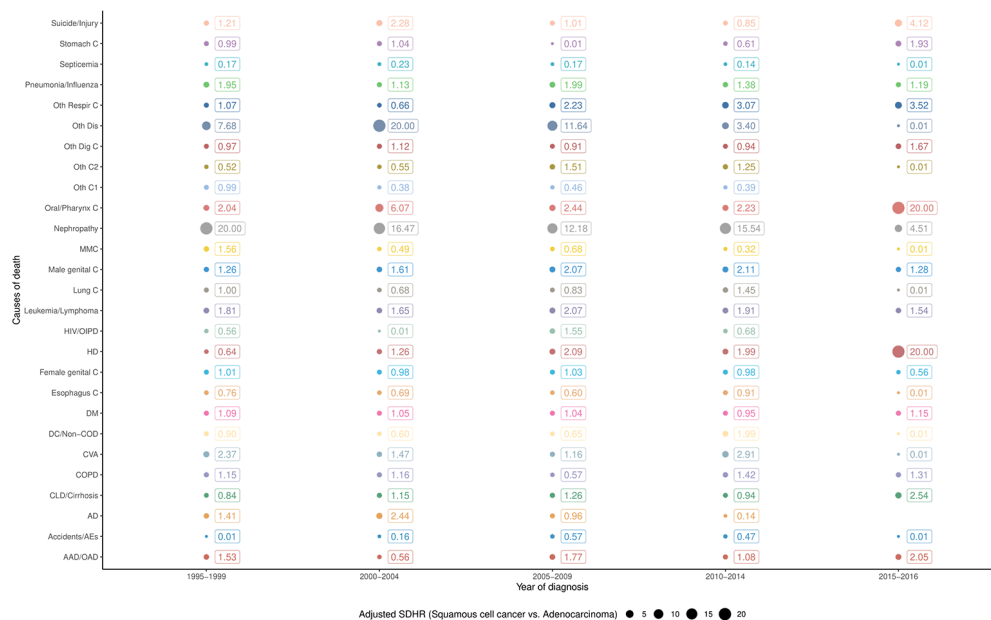


**Figure S4** Change of the adjusted subdistribution hazard ratio (SDHR) over time to year of diagnosis. Esophagua C, Esophageal cancer; HD, diseases of heart; Stomach C, stomach cancer; Oth Dis, other diseases (non-cancer) of death, including unknown behavior neoplasm, tuberculosis, syphilis, hypertension without heart disease, stomach and duodenal ulcers, complications of pregnancy, childbirth, puerperium, congenital anomalies, certain conditions originating in perinatal period, symptoms, signs and ill-defined conditions, homicide and legal intervention; MMC, miscellaneous malignant cancer, including mesothelioma and kaposi sarcoma; Lung C, lung cancer; COPD, chronic obstructive pulmonary disease; Oth Dig C, other digestive system cancer, including small intestine, colon, rectum, anus, anal canal, anorectum and others; DC/non-COD, state DC not available or state DC available but no cause of death; CVA, cerebrovascular diseases; Accidents/AEs, accidents and adverse effects; HIV/OIPD, human immunodeficiency virus (HIV) and other infectious and parasitic diseases; Female genital C, female genital cancer; Oral/Pharynx C, oral/pharynx cancer; AAD/OAD, atherosclerosis, aortic aneurysm and dissection, and other diseases of arteries, arterioles, capillaries; Male genital C, male genital cancer; Oth C1, bones and joints, soft tissue (including heart), skin system cancer and myeloma; DM, diabetes mellitus; Other Respir C, other respiratory system cancer, including larynx and others; CLD/Cirrhosis, chronic liver disease and cirrhosis; AD, Alzheimer's disease; Oth C2, eye, orbit, endocrine brain and other nervous system cancer; and Urinary C, urinary cancer.
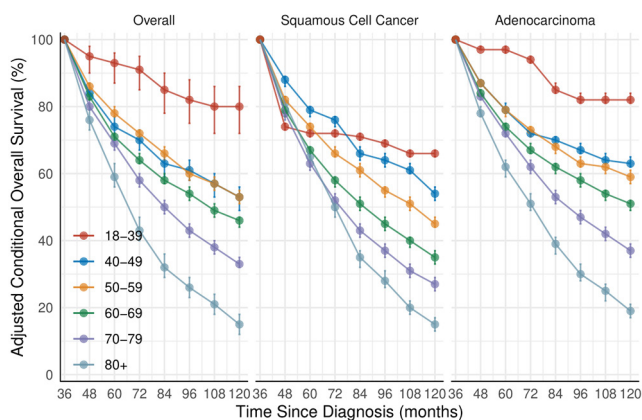
**Figure S5** The age at diagnosis-specific COS according to histology type given patients that have survived 36 months, with an adjustment for sex, ethnicity, marital status, tumor grade and stage, location, year of diagnosis, and treatment modalities.
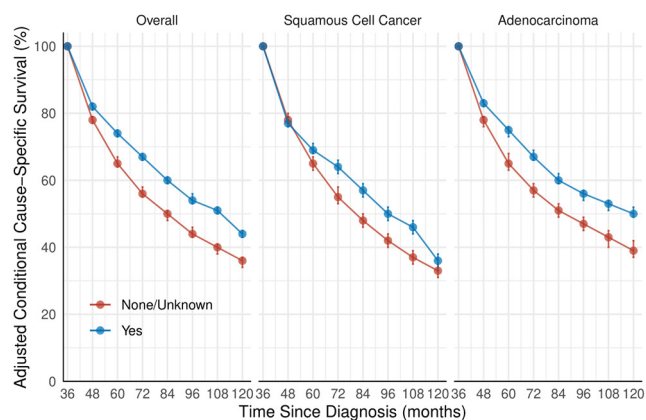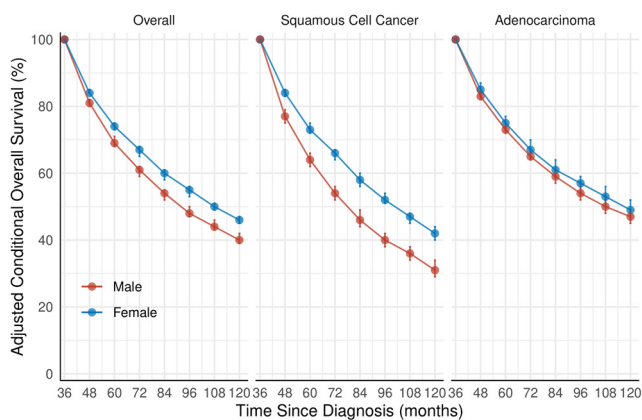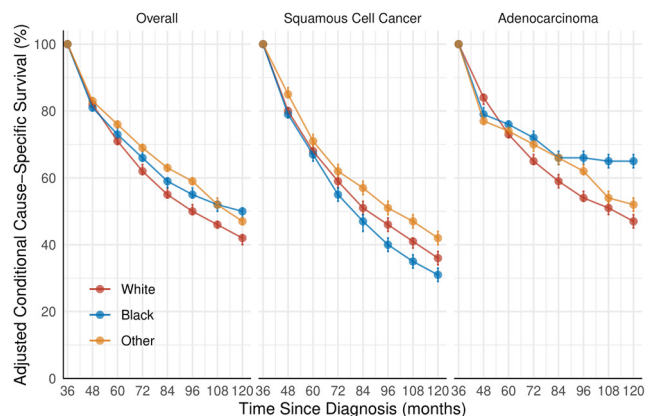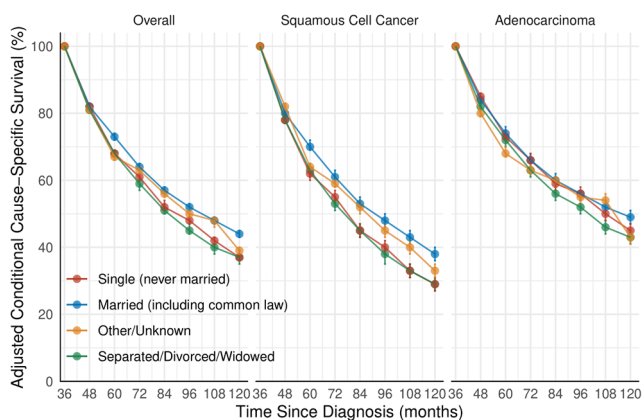


**Figure S7** The Surgery-specific COS according to histology type given patients that have survived 36 months, with an adjustment for age at diagnosis, sex, ethnicity, marital status, tumor grade and stage, location, year of diagnosis, and other treatment modalities, including chemotherapy and radiation therapy.



**Figure S6** The sex-specific COS according to histology type given patients that have survived 36 months, with an adjustment for age at diagnosis, ethnicity, marital status, tumor grade and stage, location, year of diagnosis, and treatment modalities.



**Figure S8** The ethnicity-specific COS according to histology type given patients that have survived 36 months, with an adjustment for age at diagnosis, sex, marital status, tumor grade and stage, location, year of diagnosis, and treatment modalities.

**Figure S9** The marital status-specific COS according to histology type given patients that have survived 36 months, with an adjustment for age at diagnosis, sex, ethnicity, tumor grade and stage, location, year of diagnosis, and treatment modalities.
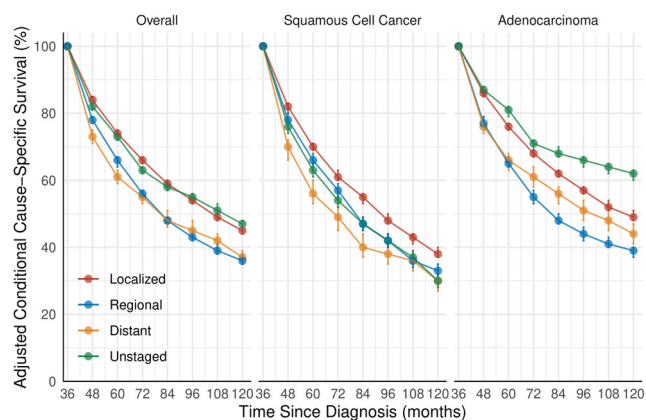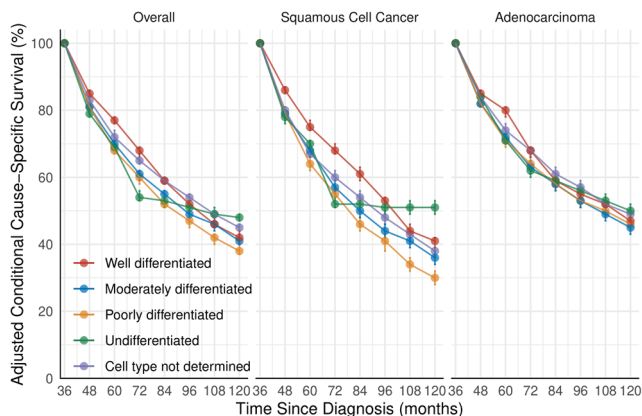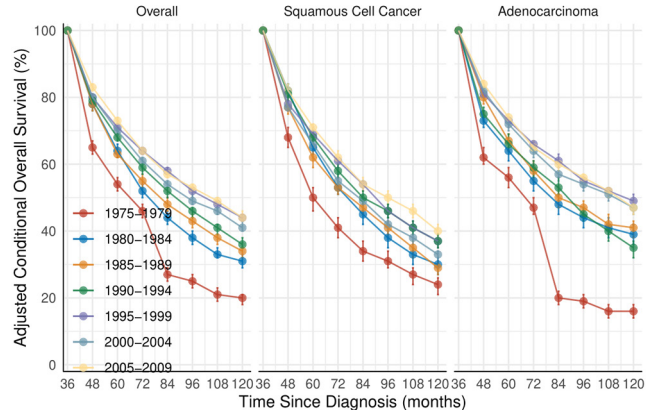


**Figure S11** The tumor stage-specific COS according to histology type given patients that have survived 36 months, with an adjustment for age at diagnosis, sex, ethnicity, marital status, tumor grade and location, year of diagnosis, and treatment modalities.



**Figure S10** The tumor grade-specific COS according to histology type given patients that have survived 36 months, with an adjustment for age at diagnosis, sex, ethnicity, marital status, tumor stage and location, year of diagnosis, and treatment modalities.



**Figure S12** The year of diagnosis-specific COS according to histology type given patients that have survived 36 months, with an adjustment for age at diagnosis, sex ethnicity, marital status, tumor grade and stage, location, and treatment modalities.
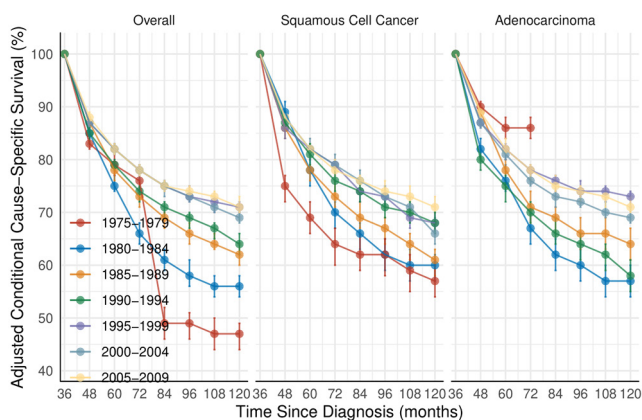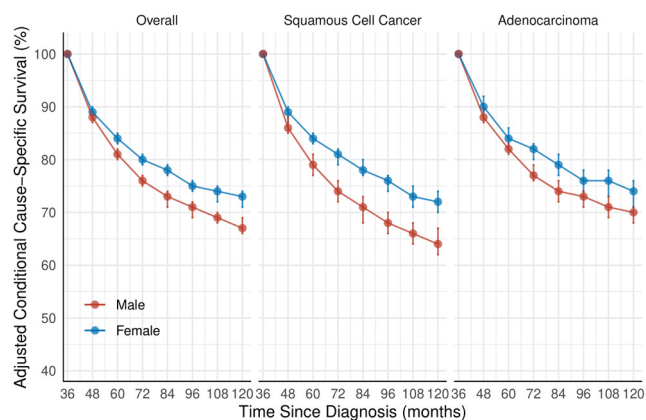
**Figure S13** The year of diagnosis- and esophageal cancer-specific conditional survival according to histology type given patients that have survived 36 months, with an adjustment for age at diagnosis, sex, ethnicity, marital status, tumor grade and stage, location, and treatment modalities.



**Figure S15** The sex- and esophageal cancer-specific conditional survival according to histology type given patients that have survived 36 months, with an adjustment for age at diagnosis, ethnicity, marital status, tumor grade and stage, location, year of diagnosis, and treatment modalities.
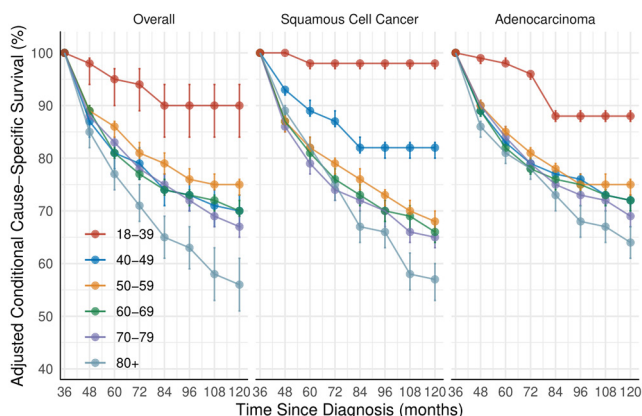


**Figure S14** The age at diagnosis- and esophageal cancer-specific conditional survival according to histology type given patients that have survived 36 months, with an adjustment for sex, ethnicity, marital status, tumor grade and stage, location, year of diagnosis, and treatment modalities.
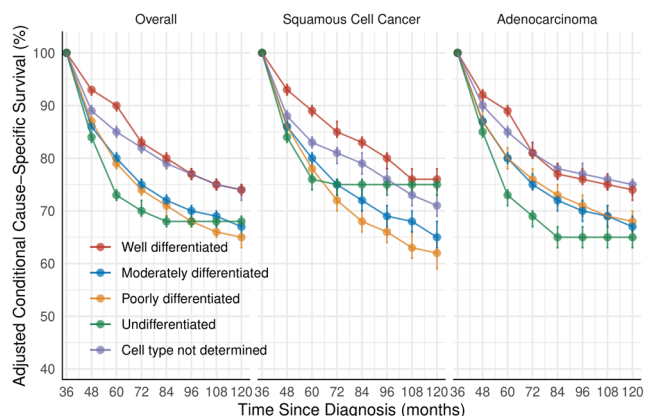


**Figure S16** The tumor grade- and esophageal cancer-specific conditional survival according to histology type given patients that have survived 36 months, with an adjustment for age at diagnosis, sex, ethnicity, marital status, tumor stage and location, year of diagnosis, and treatment modalities.
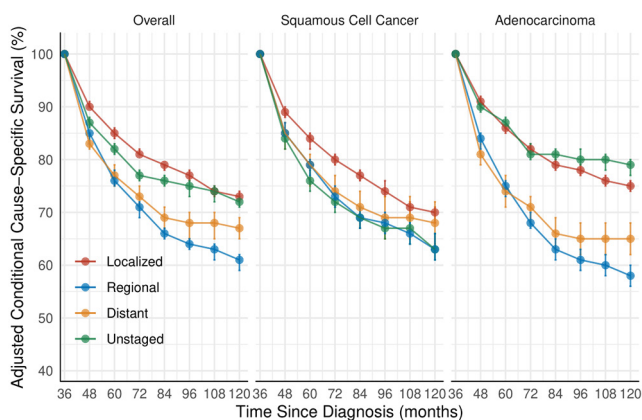
**Figure 17** The tumor stage- and esophageal cancer-specific conditional survival according to histology type given patients that have survived 36 months, with an adjustment for age at diagnosis, sex, ethnicity, marital status, tumor grade and location, year of diagnosis, and treatment modalities.
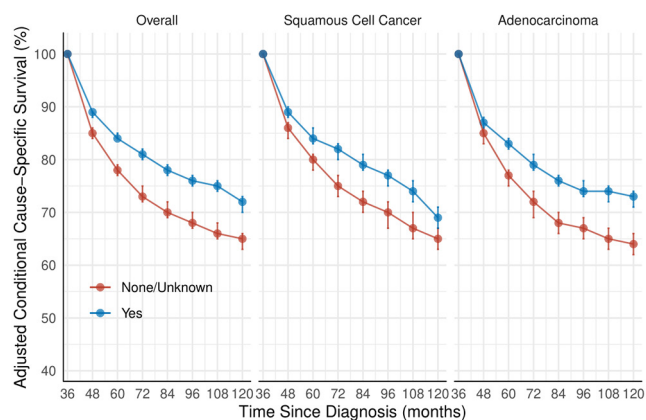
**Figure 18** The surgery- and esophageal cancer-specific conditional survival according to histology type given patients that have survived 36 months, with an adjustment for age at diagnosis, sex, ethnicity, marital status, tumor grade and stage, location, year of diagnosis, and other treatment modalities, including chemotherapy and radiation therapy.