SUPPLEMENTAL DIGITAL CONTENT

Supplemental Digital Content Figure 1. Example of how range of plausible extent of lymphadenectomy was determined for 2 patients; the first (left panel) had a ypN0 esophageal adenocarcinoma and the second (right panel) a ypN+ cancer. For 0 to 50 resected nodes, the probability of each resected number was predicted. Vertical axis depicts the cumulative frequency distribution of these values for both patients. For example, the value on the horizontal axis that intersects the cumulative frequency of 0.5 is the median number of nodes, which is much smaller for patient #1 (yoN0) than patient #2 (ypN+). Plausible values for extent of lymphadenectomy are 2 to 22 for patient #1 and 1 to 37 for patient #2 (see square brackets; note that for patient #2, the lower range is restricted to being greater than 0). Red dashed line indicates the upper cutoff for plausible extent of lymphadenectomy at 99%.

Supplemental Digital Content Figure 2. Restricted mean survival time in months for ypN0M0 patients according to ypT category, actual range of resected nodes, and potential “what if” number of resected nodes. Patients with ypT3 and ypT4 cancers have been combined. Black solid line is a loess fit to “what if” predictions shown by dots, each of which represents a minimum of 10 patients.

Supplemental Digital Content Figure 3. Restricted mean survival time (RMST) in months for ypN+M0 patients according to ypT category, actual range of lymph nodes resected, and potential “what if” number of resected nodes. Each dot represents a minimum of 10 patients. Green and red dashed lines show loess fits of values for

patients with 1-2 or 3+ positive lymph nodes (determined by pathology). Lines are displayed only when there are enough patients for reasonable approximation (10 or more required). Patients with ypT3 and ypT4 cancers have been combined.

Supplemental Digital Content Table 1. Extensive (≥30 lymph nodes) lymphadenectomy by site

Supplemental Digital Content Table 2. Thirty-day mortality according to number of lymph nodes resected

Supplemental Digital Content Appendix 1. Variables Used in Random Forest Analysis

Supplemental Digital Content Appendix 2: Method for Identifying Plausible Extents of Lymphadenectomy

Supplemental Digital Content Appendix 3: Method for Survival Analysis Using Random Survival Forests, Virtual Twin, with Interactions (RSF-VT-I)

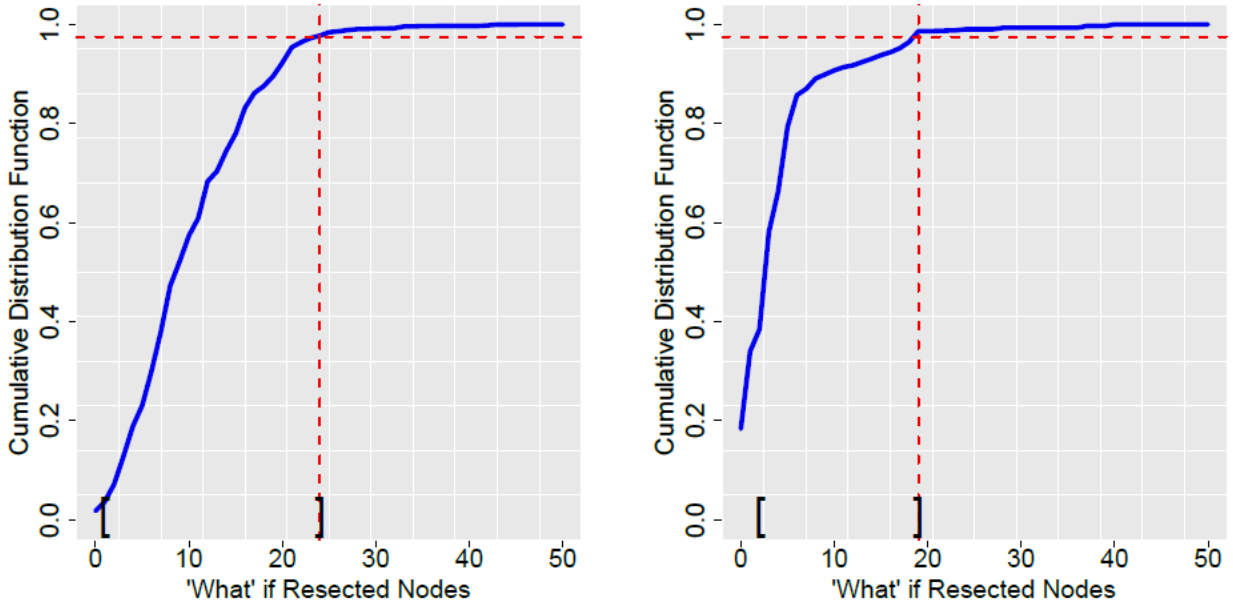
Supplemental Digital Content Table 1. Extensive (≥30 lymph nodes) lymphadenectomy by site

Site	No. of patients with ≥30 nodes resected No. (%)	No. of patients
1	130 (15)	879
2	72 (13)	542
3	78 (21)	379
4	55 (16)	333
5	116 (35)	332
6	35 (12)	287
7	2 (0.92)	217
8	11 (6.1)	180
9	12 (7.2)	166
10	13 (11)	118
11	9 (8.6)	105
12	13 (15)	84
13	19 (23)	82
14	1 (2.0)	50
15	0 (0)	39
16	2 (6.1)	33
17	0 (0)	13
18	3 (25)	12
19	1 (33)	3
20	0 (0)	2
21	0 (0)	2
22	0 (0)	1

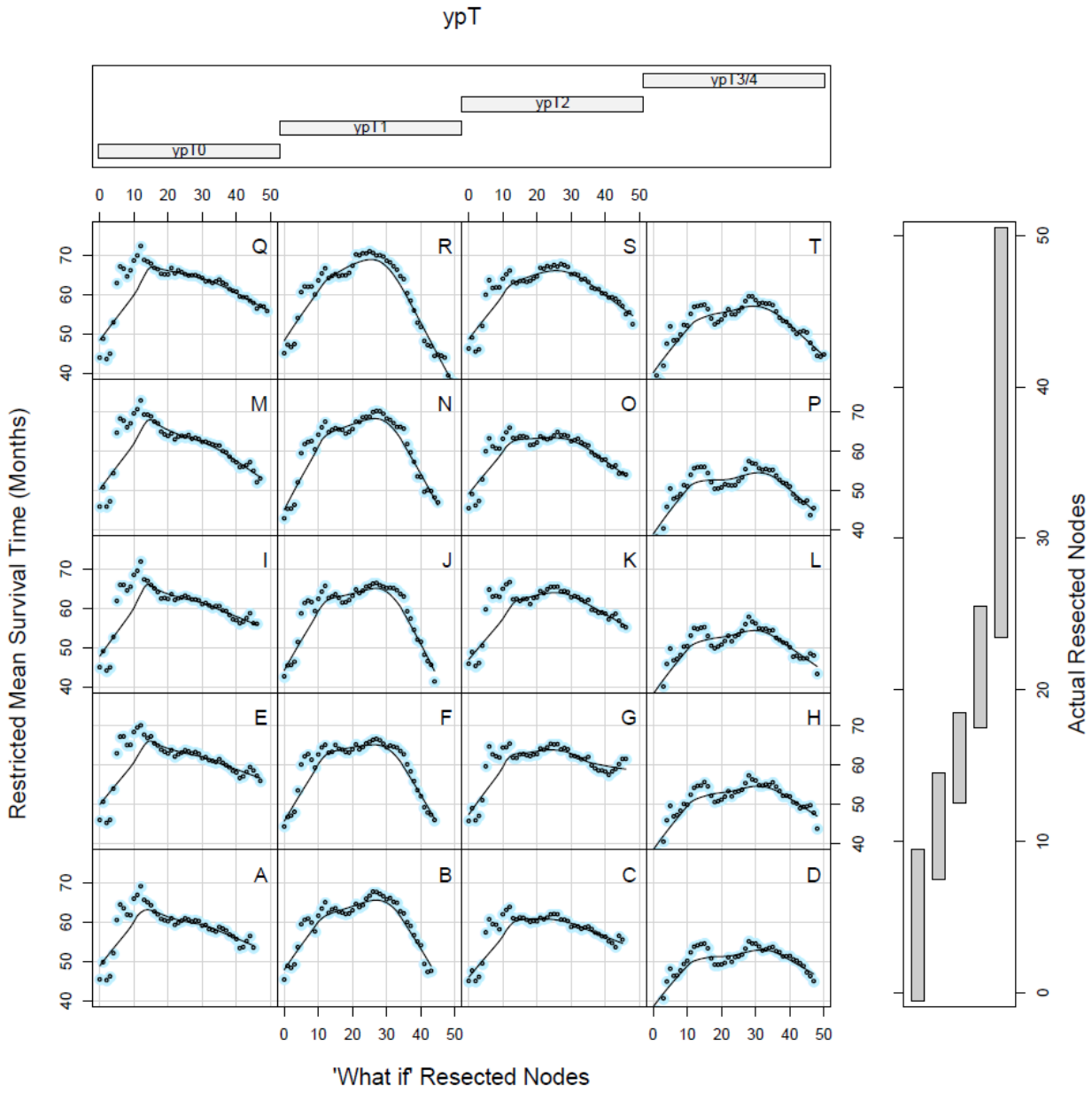
Note: Sixteen sites performed this extensive a lymphadenectomy in 20 or more patients for adenocarcinoma of the esophagus or esophagogastric junction.

Supplemental Digital Content Table 2. Thirty-day mortality according to number of lymph nodes resected

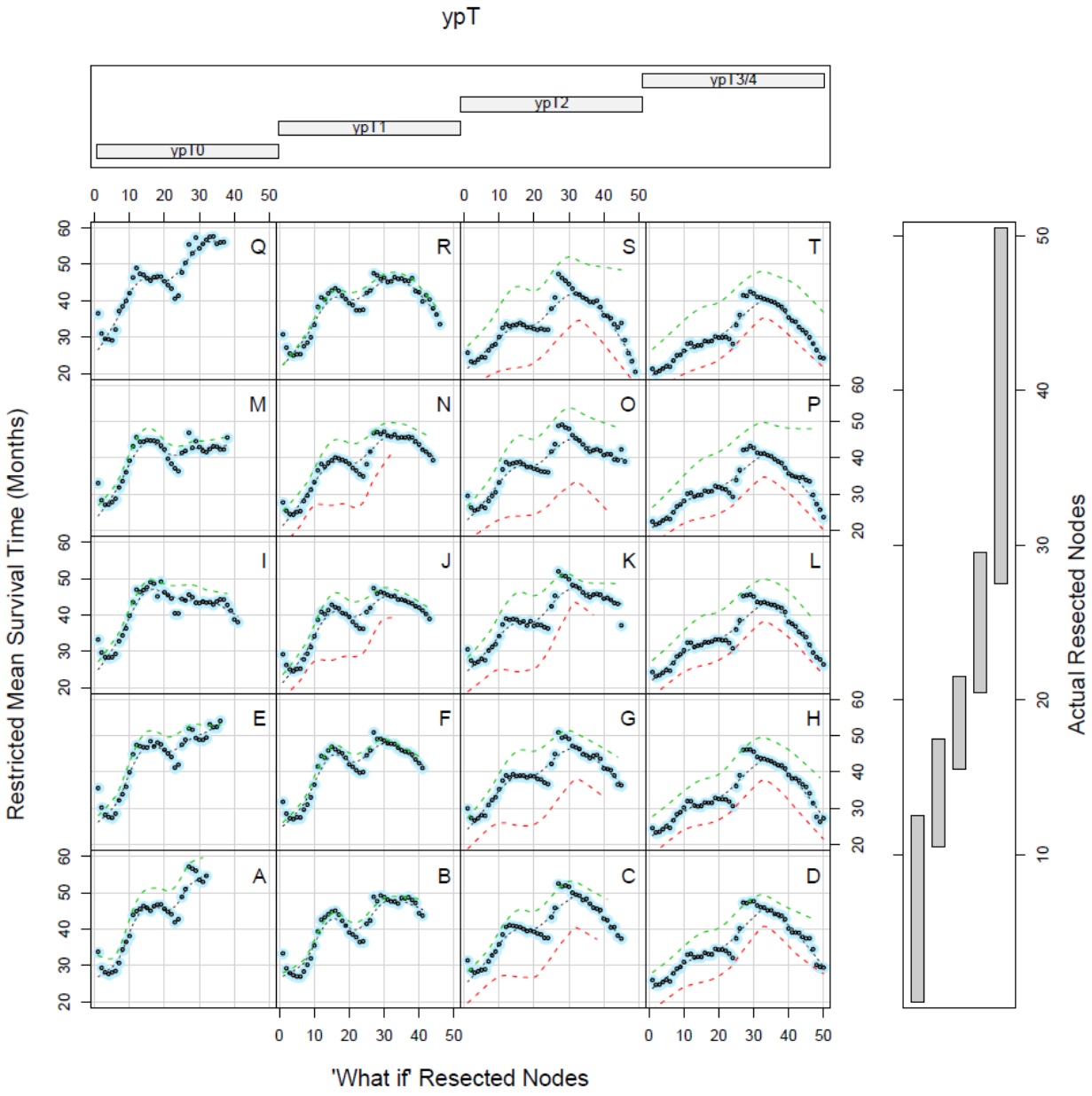
No. of lymph nodes resected	30-day mortality (% of patients)
0	9.0
1-9	2.9
10-19	1.5
20-29	1.2
30-49	1.2
≥50	0



Supplemental Digital Content Figure 1. Example of how range of plausible extent of lymphadenectomy was determined for 2 patients; the first (left panel) had a ypN0 esophageal adenocarcinoma and the second (right panel) a ypN+ cancer. For 0 to 50 resected nodes, the probability of each resected number was predicted. Vertical axis depicts the cumulative frequency distribution of these values for both patients. For example, the value on the horizontal axis that intersects the cumulative frequency of 0.5 is the median number of nodes, which is larger for patient #1 (ypN0) than patient #2 (ypN+). Plausible values for extent of lymphadenectomy are 1 to 24 for patient #1 and 2 to 19 for patient #2 (see square brackets; note that for patient #2, the lower range is restricted to being greater than 0). Red dashed line indicates the upper cutoff for plausible extent of lymphadenectomy at 0.975.



Supplemental Digital Content Figure 2. Restricted mean survival time in months for ypN0M0 patients according to ypT category, actual range of resected nodes, and potential “what if” number of resected nodes. Patients with ypT3 and ypT4 cancers have been combined. Black solid line is a loess fit to “what if” predictions shown by dots, each of which represents a minimum of 10 patients.



Supplemental Digital Content Figure 3. Restricted mean survival time (RMST) in months for ypN+M0 patients according to ypT category, actual range of lymph nodes resected, and potential "what if" number of resected nodes. Each dot represents a minimum of 10 patients. Green and red dashed lines show loess fits of values for patients with 1-2 or 3+ positive lymph nodes (determined by pathology). Lines are displayed only when there are enough patients for reasonable approximation (10 or more required). Patients with ypT3 and ypT4 cancers have been combined.

Supplemental Digital Content Appendix 1. Variables Used in Random Forest Analysis

Demographics

Age (y), sex, race

Patient characteristics

Body mass index (kg/m^2), weight loss, Eastern Cooperative Oncology Group performance status

Comorbidities

Diabetes, coronary artery disease, peripheral arterial disease, arrhythmia, hypertension, smoking, other cancers

Laboratory studies

Forced expiratory volume in 1 second, creatinine (mg/dL), bilirubin (mg/dL)

Pre-treatment cancer characteristics

Barrett esophagus, cT, cN, number of positive regional nodes, cM, histologic grade, location, distance from incisors to top of cancer, length of cancer

Pathologic cancer characteristics

pT, pN, number of positive regional lymph nodes, extracapsular lymph node invasion, lymphovascular invasion, pM, histologic grade

Resection

Number of lymph nodes resected, completeness of surgical resection (coded as R0, R1, or R2), margin positive

Supplemental Digital Content Appendix 2: Method for Identifying Plausible Extents of Lymphadenectomy

Quantile random forests regression[1] was used to estimate the association of patient, clinical, and cancer variables with number of regional lymph nodes resected and to subsequently predict the probability of each whole number of lymph nodes to be resected. Analysis was performed using the quantileReg function in the randomForestSRC R package.[2] The goal of quantile random forests is to estimate the conditional distribution function of the outcome (number of nodes resected). This value can then be used to estimate the conditional density.

Let Y be a discrete real-valued outcome (in our study, Y is the number of resected nodes). A quantile random forest is a collection of random regression trees such that for a given patient covariate $X = x$, it provides an estimate for the conditional distribution function of Y given $X = x$.

$$F_{Y|X}(y) = P\,[Y \leq y \,|\, X = x].$$

The conditional density is defined as

$$\begin{aligned} f_{Y|X}(y) &= P\,[Y = y \,|\, X = x] \\ &= P\,[Y \leq y \,|\, X = x] - P\,[Y \leq y - 1 \,|\, X = x] \\ &= F_{Y|X}(y) - F_{Y|X}(y - 1). \end{aligned}$$

Therefore, once we are given $F_{Y|X}$, we can always obtain the density $f_{Y|X}$. This is the strategy used by quantile random forests.

Note that quantile regression, which estimates the conditional distribution function, differs from usual regression, where the goal is to estimate the conditional expectation of Y.

Random forests regression[3] used 1,000 bootstrap regression trees split using mean-squared-error. Node size of a tree was set at 5 observations; random feature selection was set to 20 (i.e., 20 of the 34 variables were randomly selected at each tree node for splitting).

Estimated density for number of resected nodes revealed that density function values were strongly related to actual number of resected nodes. To determine plausible "what if" values for a patient, we found the first instance when the patient's density for number of resected nodes exceeded the Q_L-quantile for population density values for that number of resected nodes. We set Q_L to 0.025. The highest "what if" value was determined by identifying the last instance when the patient's cumulative distribution function was no larger than a preset value of Q_u, which we set to Q_u = 0.975. This defined the region of plausible "what if" values.

Typical cumulative distribution functions are illustrated in Supplemental Digital Content Figure 1, which shows the range of plausible extent of lymphadenectomy. However, for ypN+, the lowest possible "what if" value was required to be larger than 0.

---

1. Meinshausen N. Quantile regression forests. *J Machine Learning Res.* 2006;7:983-999.

2. Ishwaran H, Kogalur UB. RandomForestSRC: random forests for survival, regression and classification (RF-SRC). R package version 2.5.0. http://cran.r-project.org, 2017.

3. Breiman L. Random forests. *Machine Learning.* 2001;45:5-32.

Supplemental Digital Content Appendix 3: Method for Survival Analysis Using Random Survival Forests, Virtual Twin, with Interactions (RSF-VT-I)

We applied a novel causal inferential approach based on an extension of random survival forests (RSF), RSF-VT-I, to nonparametrically model the survival function of a patient.[1] This required expanding the number of variables by including all interactions with extent of lymphadenectomy. Variables with 2 or more levels, such as cT, were converted to 0/1 values for each level before forming the interactions. The resulting data set contained 133 variables. This induced considerable flexibility to better estimate the survival function in the face of possibly complex interactions with extent of lymphadenectomy as well as patient and cancer characteristics.

Random forest for survival (RSF) was fit using the dataset, with all-cause mortality as the outcome. Computations were implemented using randomForestSRC R software.[2] For this, 1,000 bootstrap survival trees were grown under log-rank splitting, using a tree node size of 50, with 50 features selected at random for candidates to split each tree node. (This is larger than the default setting equal to the square root of the number of features; a larger value was used due to the large number of features.)

Using the forest so constructed, RSF-VT-I yields an out-of-sample estimate of the survival function for each patient. However, each patient has only the single actual number of lymph nodes resected, but individual treatment effect (ITE) estimation requires predicted values of survival for all plausible number of nodes resected. Thus, for each patient, a sequence of survival curves was predicted by substituting for actual number of nodes resected counterfactual numbers for that same patient—a series of

"what if" lymphadenectomies.

Under the assumption of weak unconfoundedness, RSF-VT-I–estimated survival curves yield unbiased estimators of the ITE for a patient. However, just as when using propensity scores to provide unbiased estimates of an average treatment effect (ATE), it is still necessary to restrict "what if" resection values to plausible values.

---

1. Lu M, Sadiq S, Feaster DJ, et al. Estimating individual treatment effect in observational data using random forest methods. *J Comput Graph Stat.* 2018;27:209-219.

2. Ishwaran H, Kogalur UB. RandomForestSRC: Random forests for survival, regression and classification (RF-SRC). R package version 2.5.0. http://cran.r-project.org, 2017.