Assessing Treatment Effects with Adjusted Restricted Mean Time Lost in Observational Competing Risks Data

Web Appendix

Web Appendix 1 Derivation process of aRMTLd variance

Web Appendix 2 Data Collection from the SEER

Supplemental Figure 1 The flowchart of patient selection for SCCC analysis in SEER

Supplemental Table 1 The true value of marginal RMTLd under large samples in three situations in simulation study

Supplemental Table 2 Results of the simulation study comparing the aRMTLd-Conner and aRMTLd-proposed estimators

Supplemental Table 3 Baseline characteristics (Surgical groups)

Supplemental Table 4 Baseline characteristics (Radiotherapy groups)

Supplemental Table 5 Baseline characteristics (Treatment groups: combined radiotherapy with chemotherapy)

Web Appendix 1: Derivation process of aRMTLd variance

To obtain the large sample properties of aRMTLd to derive the variance, we need to assume a vector of covariates $\mathbf{Z}_i$ is bounded almost surely, and cumulative hazard rate function $\Lambda_{jk}(t) < \infty$. In addition, for time $t \in [0, \tau]$ and β, $y_k^w(t, \beta)$ is bounded away from 0, where $y_k^w(t, \beta)$ is the limit of $Y_k^w(t, \beta)$. At last, the positive-definiteness of the matrix $\Omega(\beta)$ in logistic regression must be satisfied, where $\Omega(\beta) = E\left\{p_i(\beta)\left(1, \mathbf{Z}_i^T\right)^T\left(1, \mathbf{Z}_i^T\right)\left[1 - p_i(\beta)\right]\right\}$.

First, for the large-sample nature of the derivation of the k-th group $\widehat{\mu}_k^w\left(t, \widehat{\beta}\right)$, decompose $\sqrt{n}\left[\widehat{\mu}_k^w\left(t, \widehat{\beta}\right) - \mu_k(t)\right]$,

$$\begin{aligned}&\sqrt{n}\left[\widehat{\mu}_k^w\left(t, \widehat{\beta}\right) - \mu_k(t)\right]\\&= \sqrt{n}\int_0^t\left\{\int_0^s S_k(u)\, d\left[\widehat{\Lambda}_{1k}^w\left(u, \widehat{\beta}\right) - \Lambda_{1k}(u)\right] - \int_0^s\left[\sum_{j=1}^2\left(\widehat{\Lambda}_{jk}^w\left(u, \widehat{\beta}\right) - \Lambda_{jk}(u)\right)\right] dF_{1k}(u)\right\} ds\end{aligned} \quad (1)$$

For $\sqrt{n}\left[\widehat{\Lambda}_{jk}^w\left(t, \widehat{\beta}\right) - \Lambda_{jk}(t)\right]$, it can be decomposed into two parts,

$$\begin{aligned}&\sqrt{n}\left[\widehat{\Lambda}_{jk}^w\left(t, \widehat{\beta}\right) - \Lambda_{jk}(t)\right]\\&= \sqrt{n}\left[\widehat{\Lambda}_{jk}^w\left(t, \widehat{\beta}\right) - \widehat{\Lambda}_{jk}^w(t, \beta)\right] + \sqrt{n}\left[\widehat{\Lambda}_{jk}^w(t, \beta) - \Lambda_{jk}(t)\right]\end{aligned} \quad (2)$$

Using Taylor series expansion, the first part of formula (2) can be written as

$$\sqrt{n}\left[\widehat{\Lambda}_{jk}^w\left(t, \widehat{\beta}\right) - \widehat{\Lambda}_{jk}^w(t, \beta)\right] = \hat{h}_{jk}(t)^T \sqrt{n}\left(\widehat{\beta} - \beta\right) + o_p(1), \quad (3)$$

where

$$\hat{h}_{jk}(t) = \frac{1}{n}\sum_{i=1}^n \int_0^t \left\{\frac{a_{ik}(\beta)}{Y_k^w\left(s, \widehat{\beta}\right)} - \frac{\left[n^{-1}\sum_{i=1}^n a_{ik}(\beta) Y_{ik}(s)\right] w_i(\beta)}{Y_k^w(s, \beta)^2}\right\} dN_{ijk}(s).$$

$$a_{ik}(\beta) = \frac{\partial w_i(\beta)}{\partial \beta} = (-1)^{K_i}\left(w_i(\beta) - 1\right)\left(1, \mathbf{Z}_i^T\right)^T$$

By the Strong Law of Large Numbers theory, $Y_k^w\left(t, \widehat{\beta}\right)$ and $Y_k^w(t, \beta)$ converges to $y_k^w(t, \beta)$, $n^{-1}\sum_{i=1}^n a_{ik}(\beta) Y_{ik}(t)$ converges to $E\left[a_{ik}(\beta) Y_{ik}(t)\right]$. Therefore, $\hat{h}_{jk}(t)$

converges to $h_{jk}(t)$, where

$$h_{jk}(t)=E\left[\int_0^t\left\{\frac{a_{ik}(\beta)}{y_k^w(s,\beta)}-\frac{E\left[a_{ik}(\beta)Y_{ik}(s)\right]w_i(\beta)}{y_k^w(s,\beta)^2}\right\}dN_{ijk}(s)\right].$$

The estimated value $\widehat{\beta}$ of β is obtained by the maximum likelihood method, and its score function is defined as $U(\beta)=\sum_{i=1}^{n}\left(1,\mathbf{Z}_i^T\right)^T\left[K_i-p_i(\beta)\right]$, so using Taylor expansion, we can get
$$\begin{aligned}&\frac{1}{\sqrt{n}}U(\beta)=\widehat{\Omega}(\beta)\sqrt{n}\left(\widehat{\beta}-\beta\right)+o_p(1)\\&\widehat{\Omega}(\beta)=\frac{1}{n}\sum_{i=1}^{n}p_i(\beta)\left(1,\mathbf{Z}_i^T\right)^T\left(1,\mathbf{Z}_i^T\right)\left[1-p_i(\beta)\right]\end{aligned}.$$

By the Strong Law of Large Numbers theory, $\widehat{\Omega}(\beta)$ converges to $\Omega(\beta)$. Adding $U(\beta)=\sum_{i=1}^{n}\left(1,\mathbf{Z}_i^T\right)^T\left[K_i-p_i(\beta)\right]$, we can get

$$\sqrt{n}\left(\widehat{\beta}-\beta\right)=\Omega(\beta)^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(1,\mathbf{Z}_i^T\right)^T\left[K_i-p_i(\beta)\right]+o_p(1). \quad (4)$$

According to formulas (3) and (4), we can get the second part of equation (2) as

$$\sqrt{n}\left[\widehat{\Lambda}_{jk}^w\left(t,\widehat{\beta}\right)-\widehat{\Lambda}_{jk}^w(t,\beta)\right]=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}h_{jk}(t)^T\,\Omega(\beta)^{-1}\left(1,\mathbf{Z}_i^T\right)^T\left[K_i-p_i(\beta)\right]. \quad (5)$$

Next, according to $Y_k(t,\beta)$ convergence to $y_k(t,\beta)$, we get the second part of formula (2) as

$$\sqrt{n}\left[\widehat{\Lambda}_{jk}^w(t,\beta)-\Lambda_{jk}(t)\right]=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\int_0^t\frac{w_i(\beta)}{y_k^w(s,\beta)}dM_{ijk}(s)+o_p(1) \quad , (6)$$

where $dM_{ijk}(t)=dN_{ijk}(t)-Y_{ik}(t)\,d\Lambda_{jk}(t)$.

Finally, putting formulas (5) and (6) into formula (1), we get

$$\begin{aligned}&\sqrt{n}\left[\widehat{\mu}_k^w\left(t,\widehat{\beta}\right)-\mu_k(t)\right]\\&=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\int_0^t\left\{\int_0^s S(u)\,d\varphi_{i1k}(u,\beta)-\int_0^s\left[\varphi_{i1k}(u,\beta)+\varphi_{i2k}(u,\beta)\right]dF_{1k}(u)\right\}ds\right]+o_p(1)\end{aligned},$$

$$\varphi_{ijk}(t,\beta)=h_{jk}(t)\Omega^{-1}(\beta)\left(1,\mathbf{Z}_i^T\right)^T\left[K_i-p_i(\beta)\right]+\int_0^t\frac{w_i(\beta)}{y_k^w(s,\beta)}dM_{ijk}(s)$$

where

$$dM_{ijk}(t)=dN_{ijk}(t)-Y_{ik}(t)d\Lambda_{jk}(t)$$

$$h_{jk}(t)=E\left[\int_0^t\left\{\frac{a_{ik}(\beta)}{y_k^w(s,\beta)}-\frac{E\left[a_{ik}(\beta)Y_{ik}(s)\right]w_i(\beta)}{y_k^w(s,\beta)^2}\right\}dN_{ijk}(s)\right].$$

$$a_{ik}(\beta)=(-1)^{K_i}\left(w_i(\beta)-1\right)\left(1,\mathbf{Z}_i^T\right)^T$$

Since for j=0,1 and k=0,1, $E\left[\varphi_{ijk}(t,\beta)\right]=0$, $\sqrt{n}\left[\widehat{\mu}_k^w\left(t,\widehat{\beta}\right)-\mu_k(t)\right]$ is the sum of n independent and identically distributed random variables with a mean value of 0, and therefore approximately obeys a Gaussian process with a mean value of 0. In addition, since $\sqrt{n}\left[\widehat{\mu}_k^w\left(t,\widehat{\beta}\right)-\mu_k(t)\right]$ is tight, $\sqrt{n}\left[\widehat{\Delta}^w\left(t,\widehat{\beta}\right)-\Delta(t)\right]$ approximately obeys a Gaussian process with mean 0 and covariance function $E\left[\xi_{i1}^w(t,\beta)\xi_{i0}^w(s,\beta)\right]$ (for a pair (t, s)), where

$$\xi_{ik}^w(t,\beta)=\int_0^t\left\{\int_0^s S(u)d\varphi_{i1k}(u,\beta)-\int_0^s\left[\varphi_{i1k}(u,\beta)+\varphi_{i2k}(u,\beta)\right]dF_{1k}(u)\right\}ds. \quad (7)$$

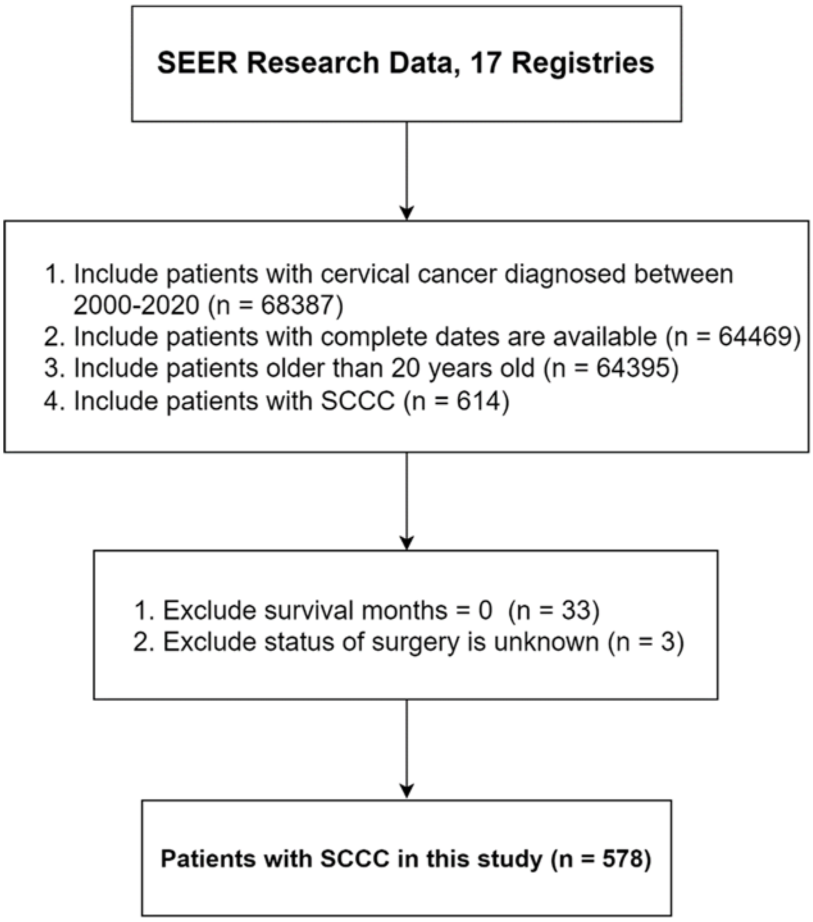
For fixed time point τ, $\sqrt{n}\left[\widehat{\Delta}^w\left(\tau,\widehat{\beta}\right)-\Delta(\tau)\right]$ converges to a mean 0 Normal distribution with a variance of $\sigma^w(\tau)^2=E\left[\left(\xi_{i1}^w(\tau,\beta)-\xi_{i0}^w(\tau,\beta)\right)^2\right]$. The estimated value of its variance is $\widehat{\sigma}^w(\tau)^2=n^{-1}\sum_{i=1}^n\left(\widehat{\xi}_{i1}^w\left(\tau,\widehat{\beta}\right)-\widehat{\xi}_{i0}^w\left(\tau,\widehat{\beta}\right)\right)^2$, which is obtained by substituting $\widehat{\xi}_{ik}^w\left(\tau,\widehat{\beta}\right)$ of the corresponding empirical estimator.

Web Appendix 2: Data Collection from the SEER

The additional selection criteria were as follows: (1) site recode International Classification of Disease (ICD)-O-3/WHO 2008: Cervix Uteri; (2) survival time was known and more than 0; (3) age 20 years or older; (4) ICD-O-3 histology codes were defined as any of the following: 8041/3, 8042/3, 8044/3, or 8045/3. In addition, patients who with an unknown surgery status were excluded. A flowchart of patients' selection for the SCCC analysis in the SEER is presented in the Supplemental Figure 2.

After screening, a total of 578 patients meeting the above conditions were included. The clinical pathological variables incorporated comprised age, stage, tumor size, lymph nodes, surgery, radiation therapy, chemotherapy, and survival in months. We defined tumour size (cm) as a categorical variable (≤2.0, >2.0－4.0, >4.0, and unknown) according to the definition in clinical guidelines. Status of lymph node was classified as: lymph node not involved, lymph node involved, or unknown. We restaged the original staging information according to the FIGO 2018 staging rules(https://www.figo.org/).

Supplemental Figure 2 The flowchart of patient selection for SCCC analysis in SEER



Supplemental Table 1 The true value of marginal RMTLd under large samples in three situations in simulation study (Unit: Days)

	$\tau = 180.5$	$\tau = 365$	$\tau = 730$
Scenario A			
RMTLd	-11.68	-28.84	-63.51
Scenario B			
RMTLd	-13.87	-29.20	-48.91
Scenario C			
RMTLd	-0.9125	-9.49	-41.61

Supplemental Table 2 Results of the simulation study comparing the aRMTLd-Conner and aRMTLd-proposed estimators

Performance measures		Bias (days)		RMSE		Coverage		RSE	
Method N	 CR	aRMTLd-Conner	aRMTLd-Proposed	aRMTLd-Conner	aRMTLd-Proposed	aRMTLd-Conner	aRMTLd-Proposed	aRMTLd-Conner	aRMTLd-Proposed
τ = 182.5 days									
Scenario A									
500	15%	-0.0949	-0.1309	0.0165	0.0165	0.9734	0.9438	1.1168	0.9785
	30%	-0.0583	-0.0953	0.0167	0.0167	0.9720	0.9446	1.1160	0.9819
1000	15%	-0.0234	-0.0414	0.0115	0.0115	0.9752	0.9460	1.1248	0.9914
	30%	-0.0472	-0.0658	0.0117	0.0117	0.9718	0.9502	1.1203	0.9917
Scenario B									
500	15%	-0.0451	-0.0838	0.0161	0.0161	0.9716	0.9440	1.1248	0.9903
	30%	-0.0069	-0.0467	0.0163	0.0163	0.9716	0.9424	1.1238	0.9937
1000	15%	0.0072	-0.0121	0.0114	0.0114	0.9708	0.9482	1.1158	0.9885
	30%	-0.0165	-0.0365	0.0116	0.0117	0.9682	0.9468	1.1090	0.9866
Scenario C									
500	15%	0.0170	-0.0076	0.0168	0.0169	0.9764	0.9456	1.1472	0.9855
	30%	0.0423	0.0171	0.0171	0.0171	0.9758	0.9466	1.1460	0.9890
1000	15%	0.0990	0.0867	0.0119	0.0119	0.9762	0.9452	1.1418	0.9859
	30%	0.0850	0.0722	0.0122	0.0122	0.9758	0.9504	1.1355	0.9851
τ = 365 days									
Scenario A									
500	15%	0.1955	0.0556	0.0352	0.0353	0.9784	0.9412	1.1655	0.9748
	30%	0.2521	0.1040	0.0362	0.0363	0.9768	0.9452	1.1605	0.9798
1000	15%	0.3507	0.2807	0.0245	0.0245	0.9804	0.9462	1.1816	0.9919
	30%	0.2899	0.2156	0.0252	0.0252	0.9780	0.9454	1.1739	0.9952
Scenario B									
500	15%	-0.0290	-0.1734	0.0345	0.0346	0.9768	0.9450	1.1742	0.9846
	30%	0.0169	-0.1359	0.0354	0.0356	0.9750	0.9448	1.1680	0.9889
1000	15%	0.0737	0.0015	0.0243	0.0243	0.9762	0.9462	1.1732	0.9877
	30%	0.0174	-0.0592	0.0251	0.0251	0.9736	0.9460	1.1628	0.9887
Scenario C									
500	15%	-0.0742	-0.1858	0.0354	0.0355	0.9802	0.9440	1.1953	0.9857
	30%	-0.0660	-0.1843	0.0364	0.0365	0.9794	0.9512	1.1915	0.9927
1000	15%	0.0986	0.0425	0.0249	0.0249	0.9828	0.9480	1.1968	0.9899
	30%	0.0452	-0.0145	0.0258	0.0259	0.9794	0.9480	1.1834	0.9891
τ = 730 days									
Scenario A									
500	15%	0.0369	-0.4112	0.0742	0.0745	0.9814	0.9416	1.1935	0.9759
	30%	0.2969	-0.2058	0.0786	0.0789	0.9786	0.9408	1.1761	0.9801
1000	15%	0.3679	0.1431	0.0514	0.0515	0.9832	0.9456	1.2151	0.9964
	30%	0.2935	0.0394	0.0545	0.0546	0.9802	0.9486	1.1956	0.9994
Scenario B									
500	15%	0.2308	-0.1765	0.0731	0.0734	0.9790	0.9422	1.1915	0.9779
	30%	0.4992	0.0438	0.0776	0.0780	0.9756	0.9446	1.1718	0.9811
1000	15%	0.3785	0.1734	0.0510	0.0511	0.9774	0.9456	1.2041	0.9910
	30%	0.3085	0.0771	0.0541	0.0543	0.9776	0.9474	1.1846	0.9948
Scenario C									
500	15%	0.0324	-0.4194	0.0733	0.0736	0.9824	0.9458	1.2266	0.9924
	30%	0.1269	-0.3891	0.0782	0.0786	0.9820	0.9494	1.2040	0.9936
1000	15%	0.3648	0.1373	0.0516	0.0517	0.9842	0.9474	1.2277	0.9953
	30%	0.2525	-0.0090	0.0552	0.0554	0.9800	0.9430	1.2021	0.9941

Note:

(1) N: sample sizes; CR: censoring rates;

the reasonable range for empirical coverage is 0.9440~0.9560, calculated using the following formula: $0.95 \pm 1.96\sqrt{0.95(1-0.95)/5000}$)

Supplemental Table 3 Baseline characteristics (Surgical groups)

Characteristics	level	All	No surgery	Surgery	P value
No. of patients		578	345	233	
Age (mean (SD))		49.85(15.61	53.13 (15.43)	45.01 (14.60)	<0.001
Tumor size (%)	≤2.0	48 (8.3)	14 (2.4)	34 (5.9)	<0.001
	>2.0-4.0	70 (12.1)	16 (2.8)	54 (9.3)	
	>4.0/Unknown	460 (79.6)	315 (54.5)	145 (25.1)	
Stage(%)	Localized	114 (19.7)	23 (4.0)	91 (15.7)	<0.001
	Regional	168 (29.1)	92 (15.9)	76 (13.1)	
	Distant	177 (30.6)	114 (24.9)	33 (5.7)	
	Unknown/ unstaged	119 (20.6)	86 (14.9)	33 (5.7)	
Lymph node involvement (%)	No/Unknown	130 (22.5)	62 (10.7)	68 (11.8)	0.002
	Yes	448 (77.5)	283 (49.0)	165 (28.5)	
FIGO 2018 stage(%)	Ⅰ	111 (19.2)	25 (4.3)	33 (14.9)	<0.001
	Ⅱ	40 (6.9)	28 (4.8)	86 (2.1)	
	Ⅲ	143 (24.7)	79 (13.7)	12 (11.1)	
	Ⅳ	186 (32.2)	148 (25.6)	64 (6.6)	
	Unknown	98 (17.0)	65 (11.2)	38 (5.7)	
Chemotherapy(%)	No/Unknown	106 (18.3)	73 (12.6)	33 (5.7)	0.043
	Yes	472 (81.7)	272 (47.1)	200 (34.6)	
Radiation(%)	No/Unknown	219 (37.9)	119 (20.6)	100 (17.3)	0.050
	Yes	359 (62.1)	226 (39.1)	133 (23.0)	

Supplemental Table 4 Baseline characteristics (Radiotherapy groups)

Characteristics	level	All	No radiation	Radiation	P value
No. of patients		578	219	359	
Age (mean (SD))		49.85(15.61	51.63 (15.46)	48.77 (15.62)	0.032
Tumor size (%)	≤2.0	48 (8.3)	21 (3.6)	27 (4.7)	0.270
	>2.0-4.0	70 (12.1)	21 (3.6)	49 (8.5)	
	>4.0/Unknown	460 (79.6)	177 (30.6)	188 (49.0)	
Stage(%)	Localized	114 (19.7)	40 (6.9)	74 (12.8)	<0.001
	Regional	168 (29.1)	37 (6.4)	37 (22.7)	
	Distant	177 (30.6)	89 (15.4)	89 (15.2)	
	Unknown/ unstaged	119 (20.6)	53 (9.2)	53 (11.4)	
Lymph node involvement (%)	No/Unknown	130 (22.5)	46 (8.0)	84 (14.5)	0.571
	Yes	448 (77.5)	173 (30.0)	275 (47.6)	
FIGO 2018 stage(%)	Ⅰ	111 (19.2)	39 (6.7)	72 (12.5)	<0.001
	Ⅱ	40 (6.9)	5 (0.9)	35(6.1)	
	Ⅲ	143 (24.7)	40 (6.9)	103 (17.8)	
	Ⅳ	186 (32.2)	91 (15.7)	95 (16.4)	
	Unknown	98 (17.0)	44 (7.6)	54 (9.3)	
Chemotherapy(%)	No/Unknown	106 (18.3)	70 (12.1)	36 (6.2)	<0.001
	Yes	472 (81.7)	149 (25.8)	323 (55.9)	
Surgery(%)	No/Unknown	219 (37.9)	119 (20.6)	226 (39.1)	0.050
	Yes	359 (62.1)	100 (17.3)	133 (23.0)	

Supplemental Table 5 Baseline characteristics (Treatment groups: combined radiotherapy with chemotherapy)

Characteristics	level	All	No	Yes	P value
No. of patients		578	219	359	
Age (mean (SD))		49.85(15.61	52.68 (16.12)	47.63 (14.84)	<0.001
Tumor size (%)	≤2.0	48 (8.3)	22 (3.8)	26 (4.5)	0.129
	>2.0-4.0	70 (12.1)	23 (4.0)	47 (8.1)	
	>4.0/Unknown	460 (79.6)	210 (36.3)	250 (43.3)	
Stage(%)	Localized	114 (19.7)	46 (8.0)	68 (11.8)	<0.001
	Regional	168 (29.1)	49 (8.5)	119 (20.6)	
	Distant	177 (30.6)	105 (18.2)	72 (12.5)	
	Unknown/ unstaged	119 (20.6)	55 (9.5)	64(11.1)	
Lymph node involvement (%)	No/Unknown	130 (22.5)	54 (9.3)	76 (13.1)	0.567
	Yes	448 (77.5)	201 (34.8)	247 (42.7)	
FIGO 2018 stage(%)	Ⅰ	111 (19.2)	46 (8.0)	65 (11.2)	<0.001
	Ⅱ	40 (6.9)	11 (1.9)	29 (5.0)	
	Ⅲ	143 (24.7)	47 (8.1)	96 (16.6)	
	Ⅳ	186 (32.2)	106 (18.3)	20 (13.8)	
	Unknown	98 (17.0)	45 (7.8)	53 (9.2)	
Surgery(%)	No/Unknown	219 (37.9)	147 (25.4)	198 (34.3)	0.422
	Yes	359 (62.1)	108 (18.7)	125 (21.6)	