Univariable survival analysis

S8

Michael Hauptmann
Netherlands Cancer Institute
Amsterdam, The Netherlands
m.hauptmann@nki.nl

NKI-AVL
The Netherlands Cancer Institute
Antoni van Leeuwenhoek Hospital

Survival analysis

- Cohort studies (incl. clinical trials & patient series)
- Survival time outcomes (also called failure time, time-to-event), censoring
- Kaplan-Meier survival estimate
- Logrank test (2-sample, 2-sample stratified, >2-sample, trend)
- Cox regression (next session)
- Proportional hazards assumption (next session)

Tumor volume study

ORIGINAL ARTICLE

TUMOR VOLUME AS PROGNOSTIC FACTOR IN CHEMORADIATION FOR ADVANCED HEAD AND NECK CANCER

Joost L. Knegjens, MD,[1] Michael Hauptmann, PhD,[2] Frank A. Pameijer, MD,[3] Alfons J. Balm, MD, PhD,[4,5] Frank J. Hoebers, MD, PhD,[1] Josien A. de Bois, RTT,[1] Johannes H. Kaanders, MD, PhD,[6] Carla M. van Herpen, MD, PhD,[7] Cornelia G. Verhoef, MD, PhD,[8] Oda B. Wijers, MD, PhD,[9] Ruud G. Wiggenraad, MD,[10] Jan Buter, MD, PhD,[11] Coen R. Rasch, MD, PhD[1]

- 360 patients (pts) w/ advanced head & neck squamous cell carcinoma (oral cavity, oropharynx, hypopharynx)
- Diagnosed 1997–2006 in 6 hospitals
- Chemoradiation for functionally/anatomically unresectable disease
- Primary tumor volume measured by pretreatment MRI or CT
- 72% male, mean age 56 yrs (range, 25–85)
- Followed for local recurrence (median, 20 months)

Data set `tumorvolume.sav`

- Includes several transformations of original variables based on code in `transformations_syntax.txt`

- Codebook `labels.doc` describes variable labels

Quantitative aspects of a cohort study

Exposed/Treated	Subjects	Events	Person-time
Yes	N_1	D_1	PYR_1
No	N_0	D_0	PYR_0
Total	N	D	PYR

- Cumulative incidence: proportion of subjects with event during given interval D/N

- Incidence rate (incidence density): #events per person-time (rate) D/PYR

- Incidence rate ratio: ratio of incidence rates by exposure/treatment (approximates relative risk for rare diseases)

$$\frac{D_1/PYR_1}{D_0/PYR_0}$$

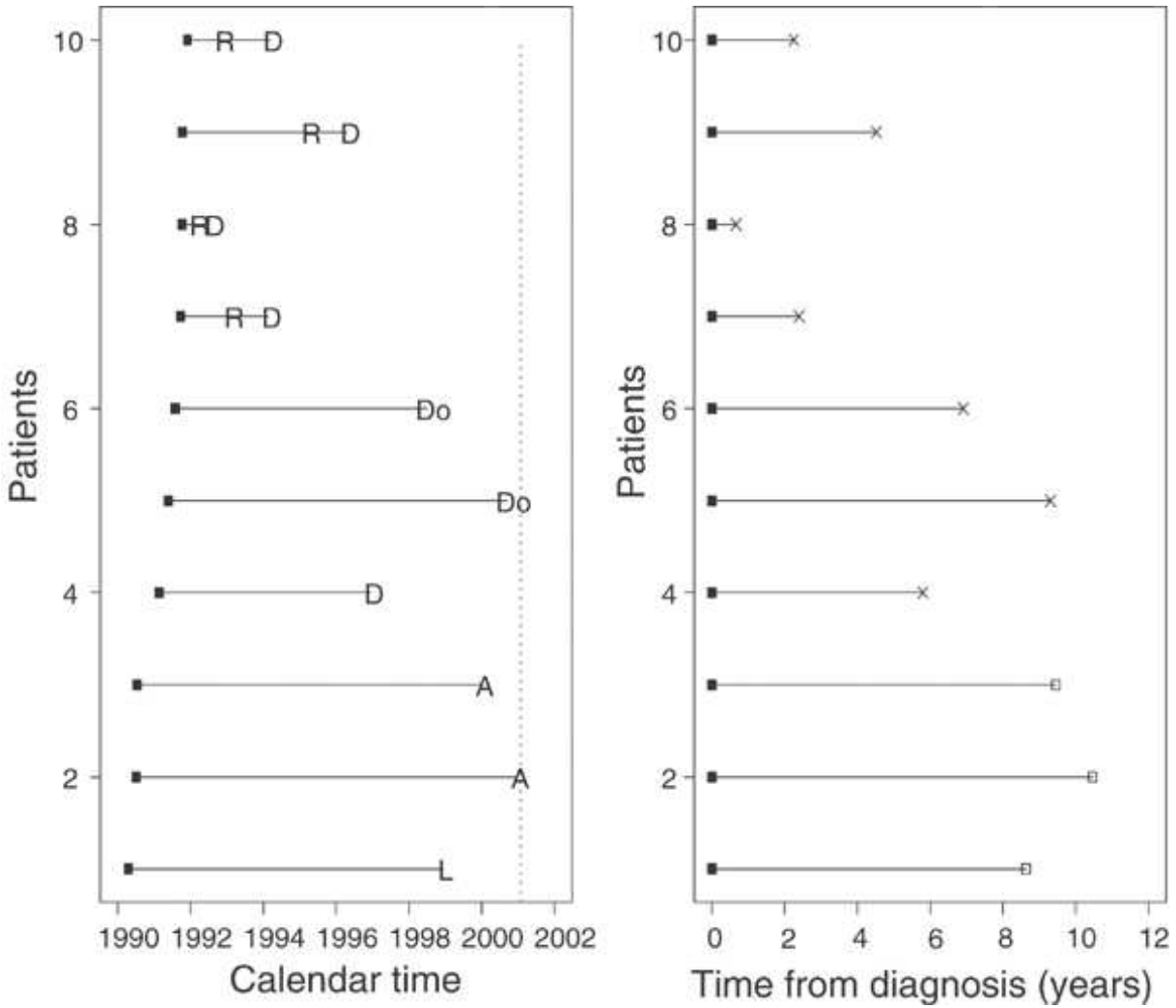
Tumor volume study with first local recurrence as event[1]

	Tumor volume (cm^3)		
	<50	≥ 50	Overall
#Pts	279	84	363
#Person-months observed	7141	1455	8595
#Events	71	34	105
#Pts w/ event	71	34	105
Mean #person-months/pt	25.6	17.3	$\frac{8595}{363}=23.7$
Incidence rate/year	11.9%	28.0%	$\frac{105}{8595/12}=14.7\%$
Incidence rate ratio	1.0 (ref)	$\frac{28.0}{11.9}=2.4$	

Survival outcome: time to an event of interest

- From cancer treatment to relapse/progression of disease (clinical trial)
- From birth or entry into a cohort to death (epidem. mortality study)
- From exposure to death (rodent carcinogenesis bioassay)
- From cancer diagnosis (DX) to death from underlying cancer or all causes (pt series)
- From exposure to palpable tumor (rodent carcinogenesis bioassay)

Schematic of survival data



R=relapse, D=death from disease, Do=death from other cause, A=attended last clinic visit (alive),
L=loss to follow-up, X=death, open square=censored

Why is this special?

- If event occurred in all study subjects, many methods of analysis would be applicable: comparison of two or more continuous distributions

- In cancer trials, survival times rarely normally distributed but skewed with many early & relatively few late events $\longrightarrow$ nonparametric test

- Survival times are censored when event of interest can no longer be observed (e.g., at end of follow-up)
  - Loss to follow-up
  - Death before event of interest
  - Other event which precludes observation of event of interest

- Special statistical methods necessary, logistic regression inappropriate

Length of follow-up

- Needs to be long enough so that sufficiently many events occur
- Needs to be tailored to event of interest (short-term, long-term)
- Interpretation of findings needs to take length of follow-up into account (conclusions beyond length of follow-up are speculation)
- In terms of statistical power, size of survival study is #events, not #pts
- Other indicator of size of study: total #person-yrs contributed by all pts

Censoring

- Examples
  - Pt has not (yet) experienced event by the close of the study
  - Loss to follow-up during study period
  - Different event makes further follow-up impossible (death from other cause, amputation)
  - Non-compliance with treatment
  - Mouse sacrificed due to other disease or death without tumor
- True time to event unknown, only: no event happened until censoring
- Censored survival time underestimates true (unknown) time to event

NKI-AVL
The Netherlands Cancer Institute
Antoni van Leeuwenhoek Hospital

Important

Always state event of interest & start/end of period of observation (including censoring)

"Local control, locoregional control & disease-free survival were calculated from the time of randomization until DX of a local recurrence, regional recurrence, distant metastasis, new tumor, death, loss to follow-up or end of follow-up, whichever came first. Pts with distant metastasis, new tumor, or death as the first event were censored at that time as well as pts lost to follow-up or event-free at the end of follow-up. For calculation of local control, pts with regional recurrence were censored as well.

Overall survival was the time from randomization until death, loss to follow-up, or end of follow-up, whichever came first. Pts alive at end of follow-up or lost to follow-up were censored at that time, respectively.

Disease-specific survival & metastasis-free survival were calculated accordingly." (Rasch et al. Cancer 2010)

Tumor volume study

- Primary outcome: first local recurrence

- Represented by 2 variables
  - Survival time: time (in months) from start of treatment until DX of local recurrence, regional recurrence, distant metastasis, new tumor, death, loss to follow-up or end of follow-up, whichever came first
  - Event indicator: binary variable indicating whether local recurrence was diagnosed at survival time (event of interest occurred, value=1) or not (censored, value=0)

Survival & hazard

- Survival probability (survivor function)

$$S(t) = P(T \geq t)$$

  Probability that survival time T is at least t, i.e., that subject survives from time origin (e.g., cancer DX) to future time t

- Hazard

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}$$

  Probability that subject who is at risk at time t has event at that time (instantaneous event rate for subject who has already survived to t), difficult to interpret but useful for modeling

- Hazard relates to incident current event rate, survival reflects cumulative non-occurrence

Kaplan-Meier (product-limit) survival estimate

- Nonparametric estimate of survival S(t) from observed (censored & uncensored) survival times

- k pts have events at distinct times $t_1 < t_2 < \ldots < t_k$

- Assume censoring is unrelated to event probability & events occur independently from each other

$$S(t_j) = S(t_{j-1})(1 - \frac{d_j}{n_j})$$

$S(t_{j-1})$ probability of being alive at t_{j-1}
n_j #pts alive just before t_j
d_j #events at t_j
$t_0 = 0$
$S(0) = 1$

Survival time calculation[2]

volumeg2=1 if tumor volume<50 cm^3 and 2 otherwise

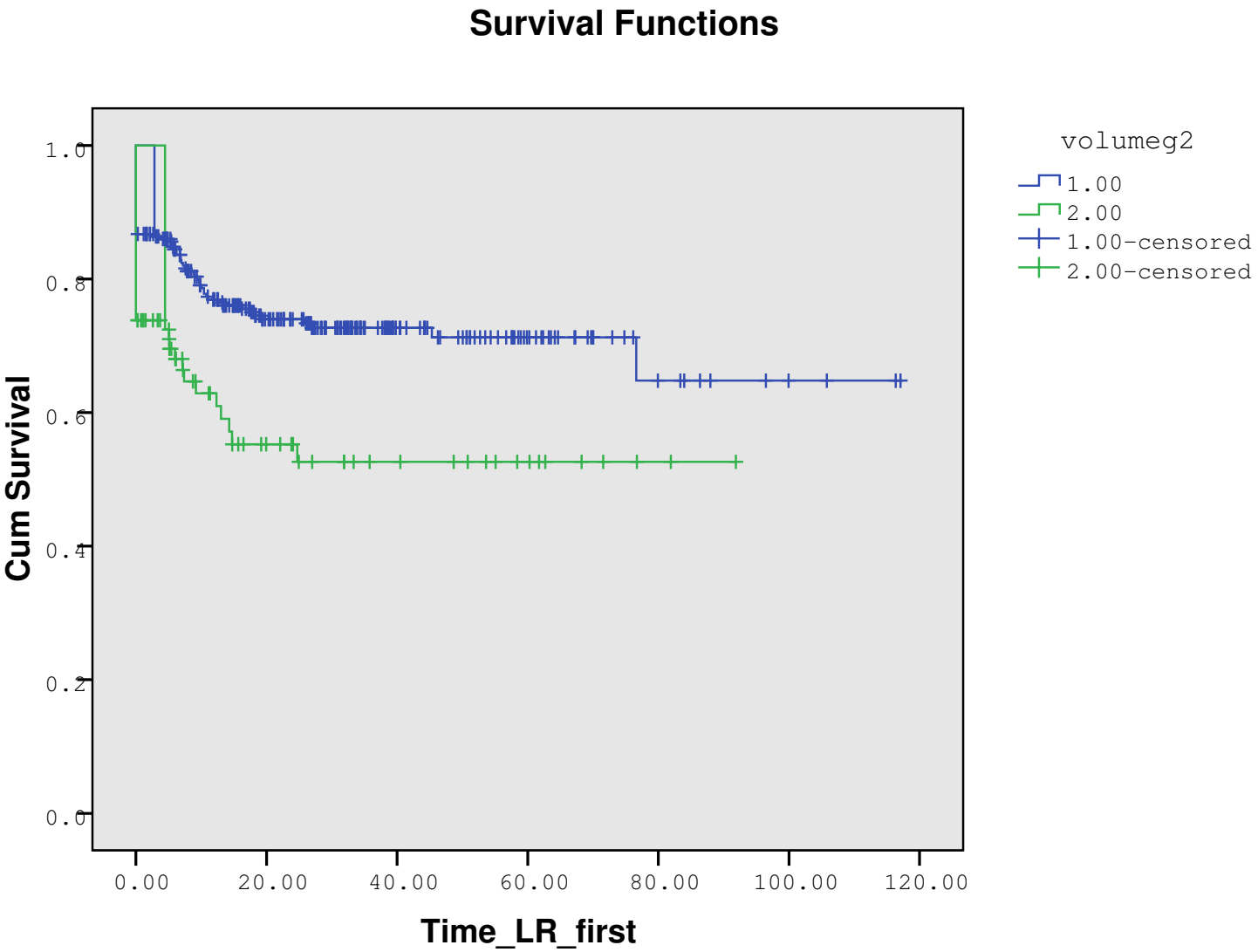
Here volumeg2=1:

T		KM
0	37 events	1*(1-37/279)=.867
.3	1 censored	.867*(1-0/242)=.867
...	...	...
2.63	1 censored	.867*(1-0/234)=.867
2.83	1 event	.867*(1-1/233)=.864
3.06	1 censored	.864*(1-0/232)=.864
...	...	...
3.45	1 censored	.864*(1-0/229)=.864
3.78	1 event	.864*(1-1/228)=.860
...	...	...

Survival Table

volumeg2		Time	Status	Cumulative Proportion Surviving at the Time		N of Cumulative Events	N of Remaining Cases
				Estimate	Std. Error		
1.00	1	.000	1	.	.	1	278
	2	.000	1	.	.	2	277
	3	.000	1	.	.	3	276
	4	.000	1	.	.	4	275
	5	.000	1	.	.	5	274
	6	.000	1	.	.	6	273
	7	.000	1	.	.	7	272
	8	.000	1	.	.	8	271
	9	.000	1	.	.	9	270
	10	.000	1	.	.	10	269
	11	.000	1	.	.	11	268
	12	.000	1	.	.	12	267
	13	.000	1	.	.	13	266
	14	.000	1	.	.	14	265
	15	.000	1	.	.	15	264
	16	.000	1	.	.	16	263
	17	.000	1	.	.	17	262
	18	.000	1	.	.	18	261
	19	.000	1	.	.	19	260
	20	.000	1	.	.	20	259
	21	.000	1	.	.	21	258
	22	.000	1	.	.	22	257
	23	.000	1	.	.	23	256
	24	.000	1	.	.	24	255
	25	.000	1	.	.	25	254
	26	.000	1	.	.	26	253
	27	.000	1	.	.	27	252
	28	.000	1	.	.	28	251
	29	.000	1	.	.	29	250
	30	.000	1	.	.	30	249
	31	.000	1	.	.	31	248
	32	.000	1	.	.	32	247
	33	.000	1	.	.	33	246
	34	.000	1	.	.	34	245
	35	.000	1	.	.	35	244
	36	.000	1	.	.	36	243
	37	.000	1	.867	.020	37	242
	38	.300	0	.	.	37	241
	39	1.180	0	.	.	37	240
	40	1.510	0	.	.	37	239
	41	1.640	0	.	.	37	238
	42	1.680	0	.	.	37	237
	43	2.140	0	.	.	37	236
	44	2.140	0	.	.	37	235
	45	2.630	0	.	.	37	234
	46	2.830	1	.864	.021	38	233
	47	3.060	0	.	.		

Survival curve

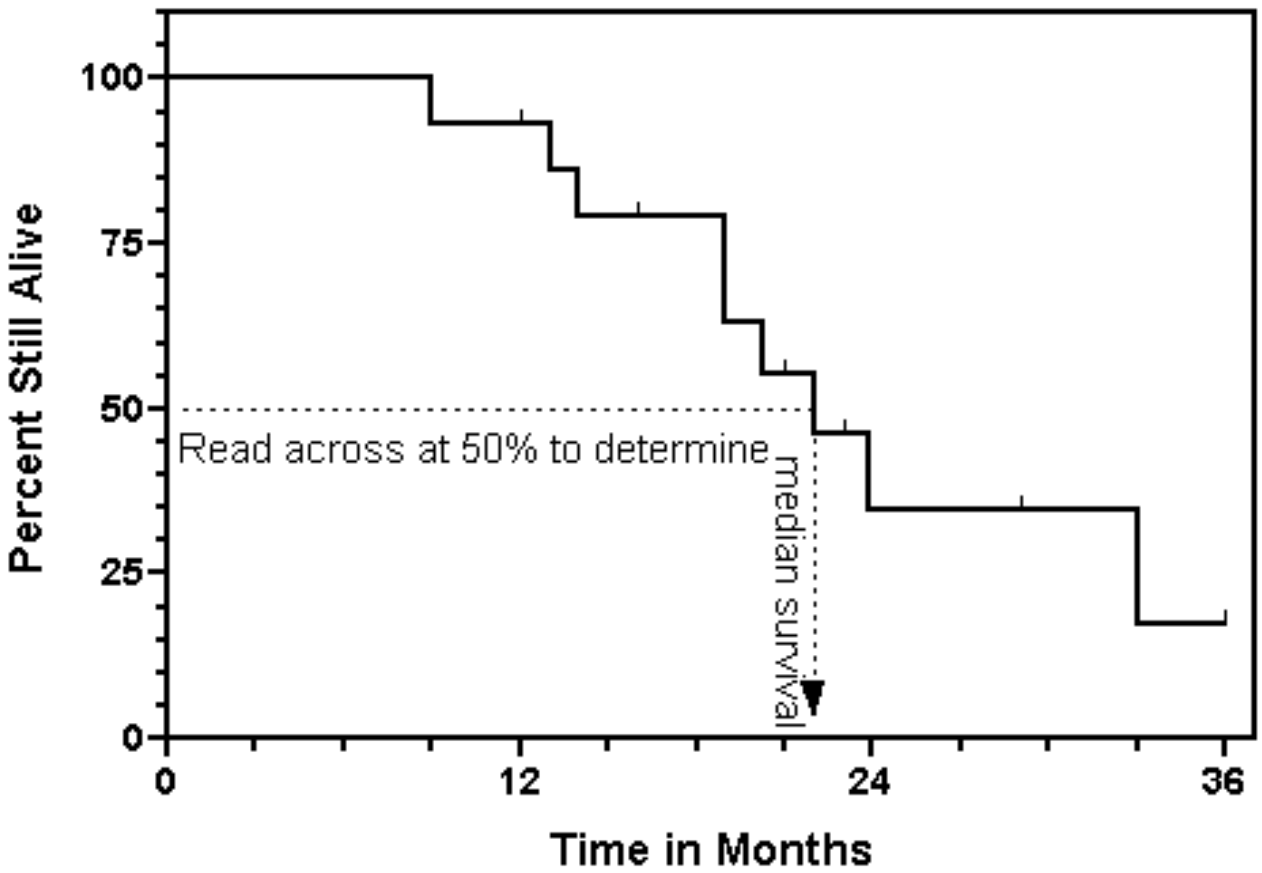


- 5-yr local recurrence-free survival for pts w/ lower volume tumors is about .71 (71%)
- Steep decline in first yr indicates generally poor prognosis
- #Lost to follow-up can be calculated from total #pts & #events, and #pts at risk at end of follow-up

Properties of KM estimator

- Value of $S(t)$ constant between event times
- Step function that changes value only at times of events
- Pts contribute information to calculations for as long as they are known to be event-free
- If all pts experience event: ratio of #pts event-free at t divided by #pts in the study
- KM survival curve is plot of KM survival probability against time
- KM plot can be used to estimate median survival (mean survival not often used due to large skew in distribution of most survival data)

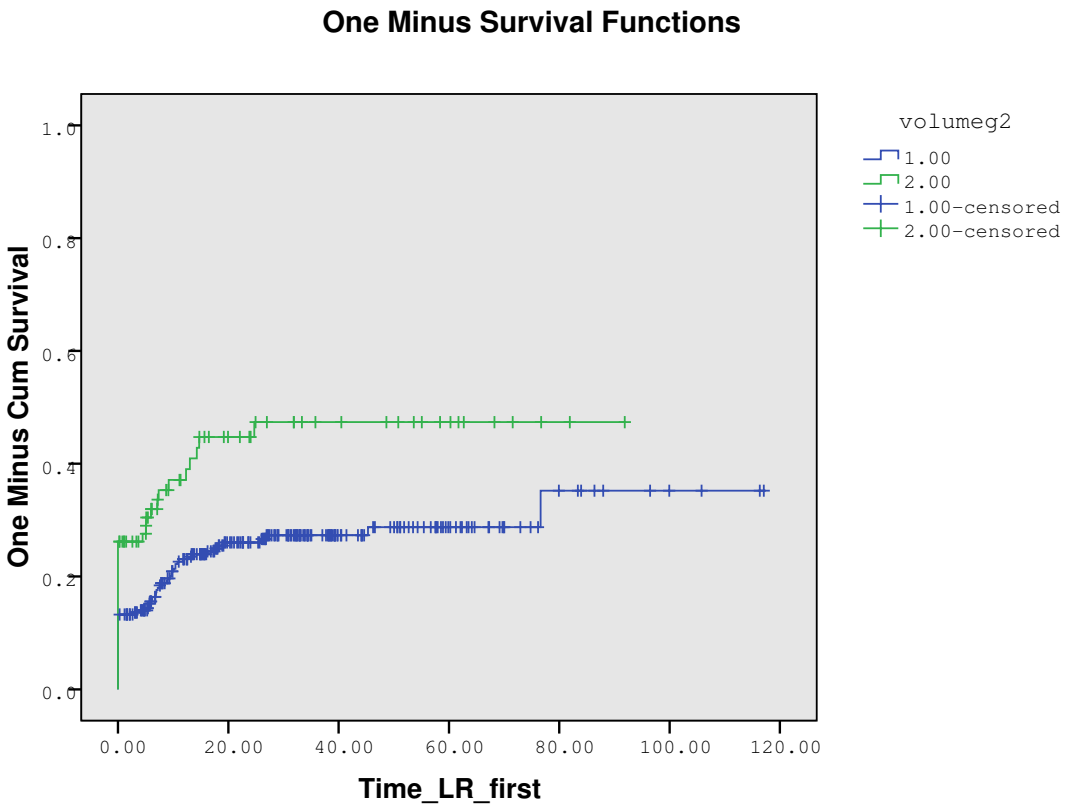
Median survival time



Median local recurrence-free survival unknown for tumor volume study because follow-up is too short (or recurrence rate too low)

Cumulative incidence of local recurrence curve

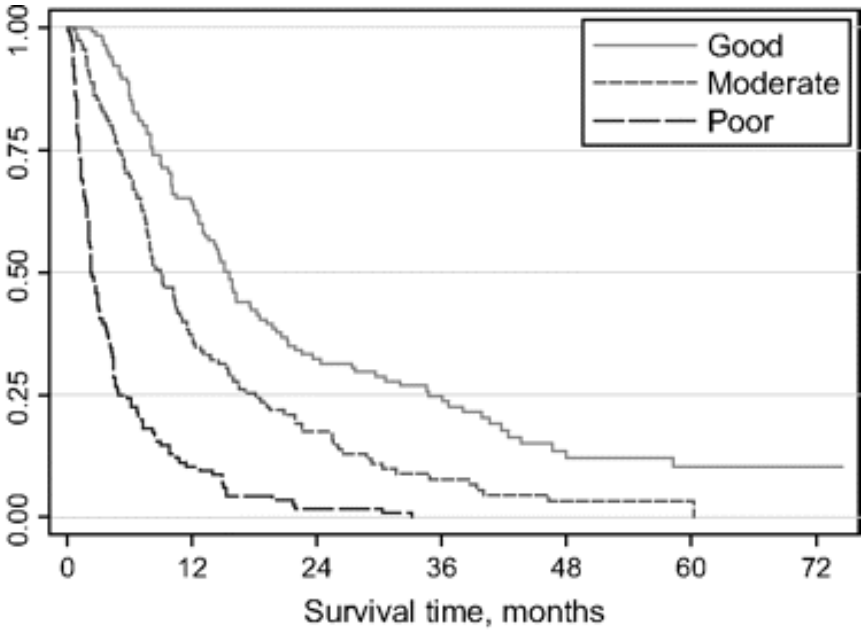
Cumulative incidence = 1–S(t)



Cumulative incidence among pts w/ lower volume tumors is 29% for the 1st 5 yrs (5-yr survival 71%)

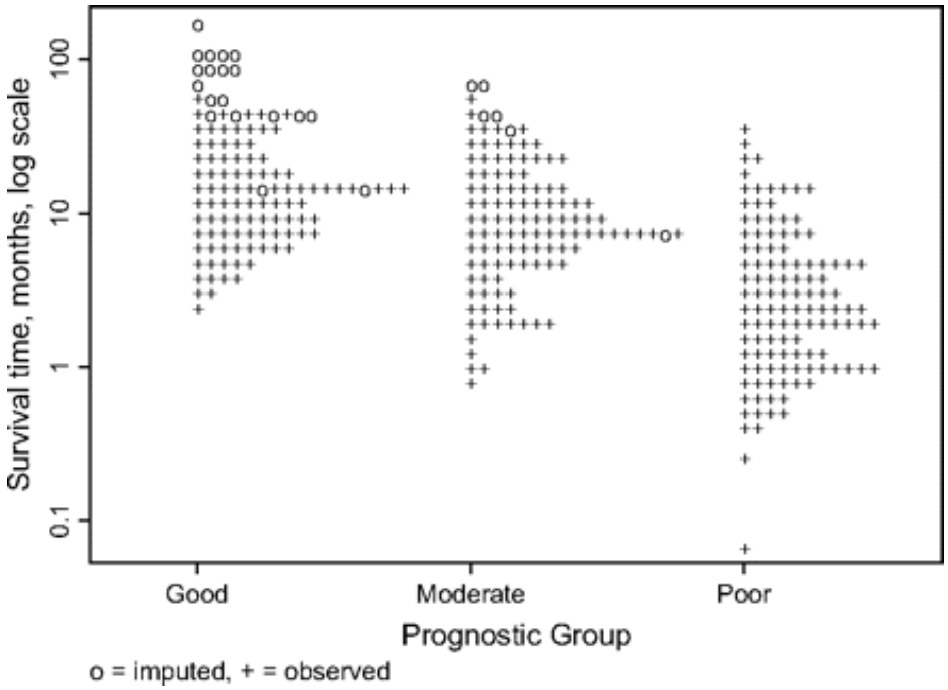
Separation of survival curves often overestimated by KM plots

Survival of pts by prognostic group in the Medical Research Council RE01 trial



Number at risk

Good	115	74	36	8
Moderate	116	42	20	3
Poor	116	12	2	0

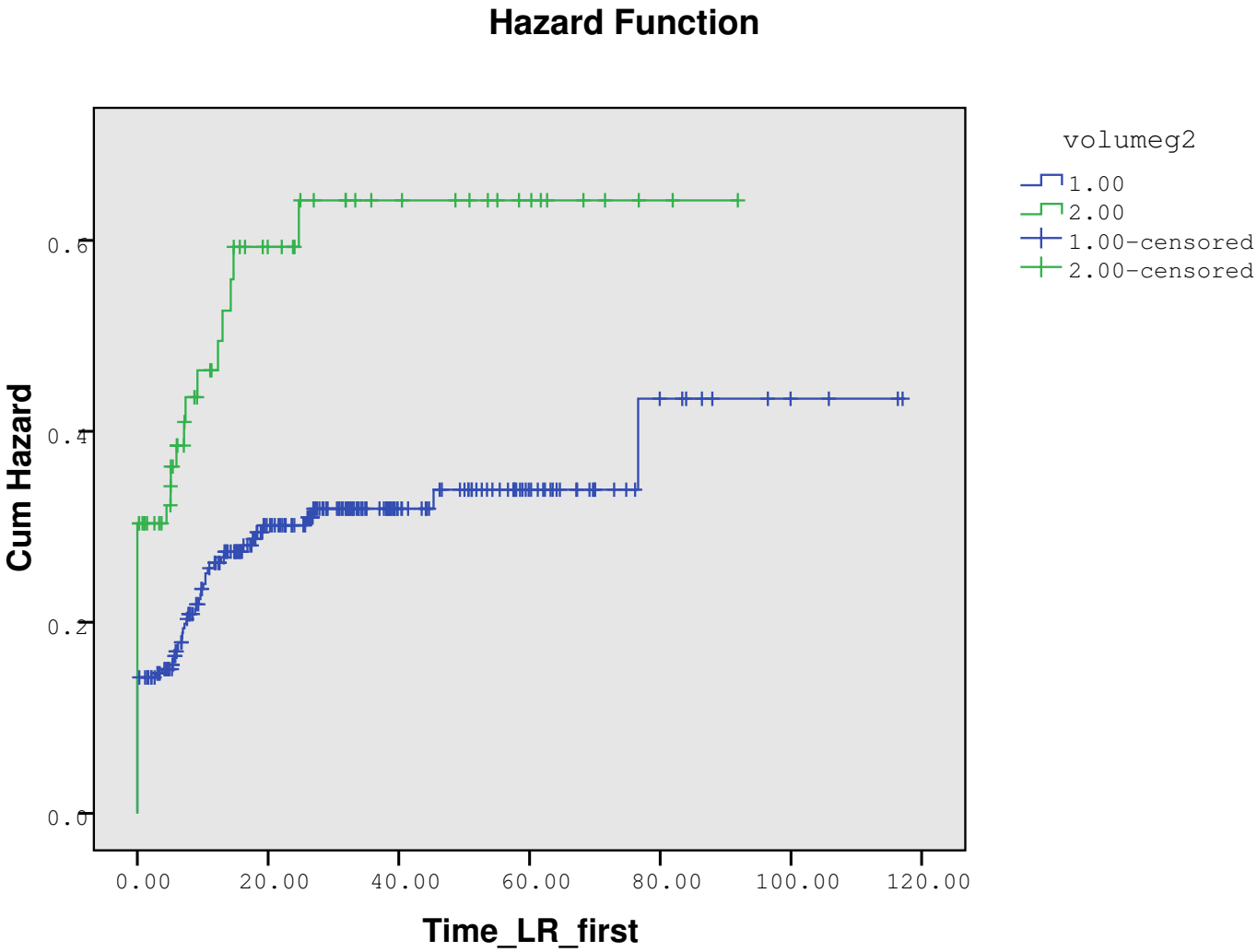


Royston et al., *J Natl Cancer Inst* 2008

Hazard & cumulative hazard

- Relationship between $S(t)$ & $h(t)$: $h(t) = -\frac{d}{dt} \log S(t)$
- $h(t)$ cannot be easily estimated
- Instead, estimate cumulative hazard $H(t) = \int_0^t h(t)dt = -\log S(t)$
  "area under the hazard function between times 0 and t"
- Interpretation difficult: cumulative force of mortality or expected number of events for each subject if event was repeatable process
- $H(t)$ is used as intermediary measure for estimating $h(t)$ & for model diagnostics

Cumulative hazard curves



Two sample logrank (Mantel-Haenszel) test

- Comparing 2 survival curves nonparametrically (e.g., treatment arms)

- H_0: $h_A(t) = h_B(t)$ vs. H_1: $h_A(t) = \theta * h_B(t)$

- 2-by-2 table at each failure time

Treatment	Died at time t_j	Alive at time t_j	At risk just before t_j
A	d_{Aj}	$n_{Aj} - d_{Aj}$	n_{Aj}
B	d_{Bj}	$n_{Bj} - d_{Bj}$	n_{Bj}
Total	d_j	$n_j - d_j$	n_j

- Cell frequencies conditional on marginals under $H_0 \rightarrow d_{Aj}$ determines the rest of the table
  Example: 60% group A, 40% group B, under H_0, at each failure time 60% of all observed events are expected in group A

Test statistic

Treatment	Died at time t_j	Alive at time t_j	At risk just before t_j
A	d_{Aj}	$n_{Aj} - d_{Aj}$	n_{Aj}
B	d_{Bj}	$n_{Bj} - d_{Bj}$	n_{Bj}
Total	d_j	$n_j - d_j$	n_j

$$\frac{\left\{\sum_j \left[d_{Aj} - \mathrm{E}(d_{Aj})\right]\right\}^2}{\mathrm{VAR}} = \frac{\left[\sum_j (\mathrm{O}_i - \mathrm{E}_i)\right]^2}{\mathrm{VAR}} \sim \chi^2(1 \text{ DF})$$

At each event time & given the observed #events which occurred, the #events expected in the 1st group is calculated under the null hypothesis of no group differences, then summed up over all event times for the expected total #events in the first group, test compares observed #events with expected

Does local recurrence differ by tumor volume?[3]

Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	10.689	1	.001

Test of equality of survival distributions for the different levels of volumeg2.

Significantly better survival for pts w/ lower volume tumors (p=.001)

Stratified two sample logrank test

- Similar to Cochran-Mantel-Haenszel test for stratified 2-by-2 tables
- Test statistic

$$\frac{\left[\sum_{\text{strata}} \sum_j \left(\mathsf{O}_{ij} - \mathsf{E}_{ij}\right)\right]^2}{\mathsf{VAR}} \sim \chi^2(1 \text{ DF})$$

- Assumption: homogeneity of treatment effects across strata
- Evaluate with Breslow-Day test of homogeneity for stratified 2-by-2 tables (M-1 DF for M strata)

Does volume measurement modality (CT vs. MRI) confound the results?[4]

Case Processing Summary

MRI_CT	volumeg2	Total N	N of Events	Censored	
				N	Percent
1	1.00	205	49	156	76.1%
	2.00	70	28	42	60.0%
	Overall	275	77	198	72.0%
2	1.00	73	21	52	71.2%
	2.00	14	6	8	57.1%
	Overall	87	27	60	69.0%
Overall	Overall	362	104	258	71.3%

Overall Comparisons[a]

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	12.066	1	.001

Test of equality of survival distributions for the different levels of volumeg2.

a. Adjusted for MRI_CT.

Volume measurement modality does apparently not confound the association between tumor volume & local recurrence

More than 2 samples

- H_0: $h_1(t) = h_2(t) = \ldots = h_G(t)$

- Test statistic

$$\frac{\left[\sum_{\text{samples}} \sum_j \left(\mathsf{O}_{ij} - \mathsf{E}_{ij}\right)\right]^2}{\mathsf{VAR}} \sim \chi^2(G - 1 \text{ DF})$$

O_{ij} observed #failures at event time i in group j

E_{ij} expected #failures at event time i in group j

Local recurrence-free survival by 4 tumor volume categories[5]

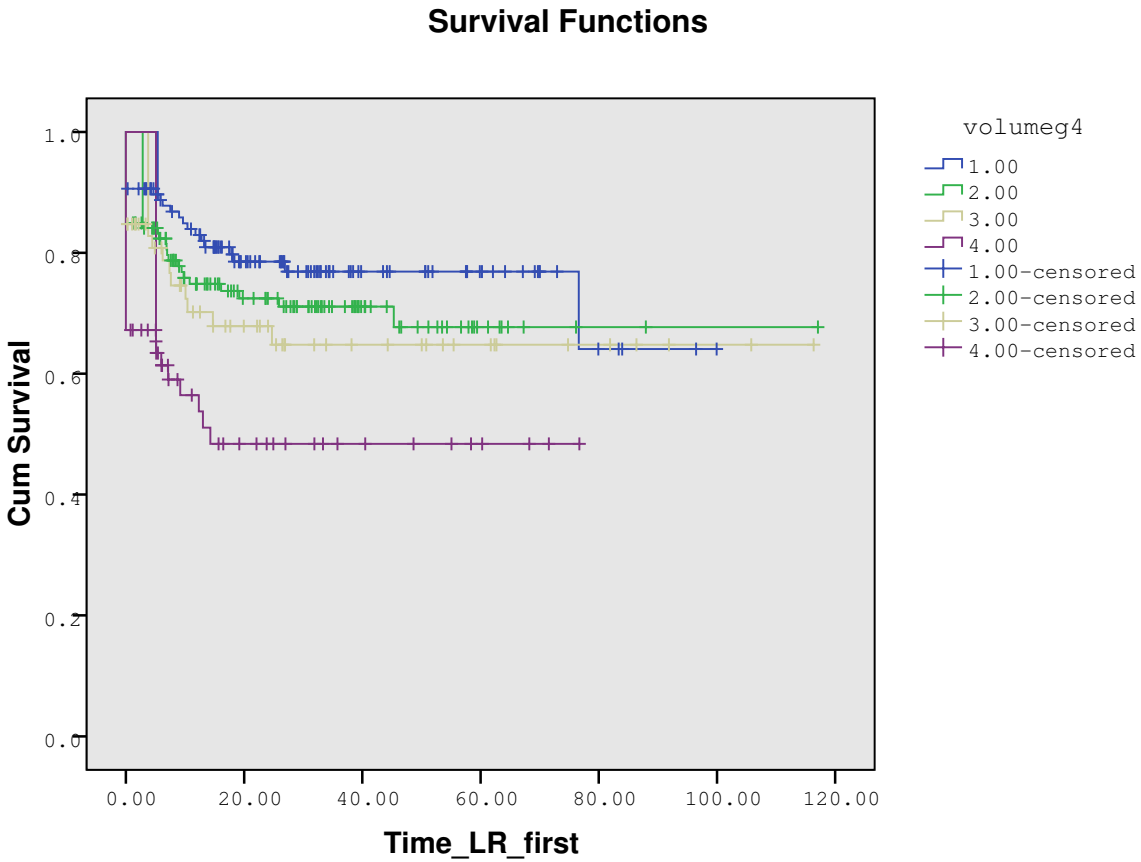
Case Processing Summary

volumeg4	Total N	N of Events	Censored	
			N	Percent
1.00	117	25	92	78.6%
2.00	126	34	92	73.0%
3.00	59	18	41	69.5%
4.00	61	28	33	54.1%
Overall	363	105	258	71.1%

Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	17.966	3	.000

Test of equality of survival distributions for the different levels of volumeg4.



Strong evidence of heterogeneity between 4 tumor volume groups ($p<.001$)

Test of trend

- If groups naturally ordered (age groups, stages of cancer, lo-med-hi, none-moderate-severe, doses of treatment/exposure), consider trend in survival across groups
  $\rightarrow$ null hypothesis the same as above
  $\rightarrow$ ordered alternative, more powerful test, only 1 DF

- Scores $w_1, w_2, \ldots, w_G$ represent groups (1, 2, 3, ... or actual values)

- Test statistic

$$\frac{\left[\sum_{\text{samples}} w_j \sum_j \left(\mathrm{O}_{ij} - \mathrm{E}_{ij}\right)\right]^2}{\mathrm{VAR}} \sim \chi^2(1 \text{ DF})$$

  O_{ij} observed #failures at event time i in group j
  E_{ij} expected #failures at event time i in group j

Survival trend across 4 tumor volume categories?[6]

Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	15.518	1	.000

The vector of trend weights is -3, -1, 1, 3. This is the default.

- Significant trend in survival with tumor volume ($p<.001$)

- Compare with k-sample logrank test

Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	17.966	3	.000

Test of equality of survival distributions for the different levels of volumeg4.

Comments on logrank test

- Assumptions same as for KM survival estimates
  - Censoring unrelated to event probability (uninformative)
  - Survival similar for subjects recruited early & late in study

  Deviations matter most when different across groups

- Purely test of significance, no estimate of magnitude of difference

- Most likely detects difference when risk of event is consistently greater in one group

- Unlikely to detect difference if survival curves cross $\rightarrow$ always plot survival curves

- For 2 groups, logrank test tests H_0 that ratio of the hazard rates in the two groups equals 1

- Hazard ratio (HR) is measure of relative survival experience, can be estimated by

$$\text{HR} = \frac{O_1/E_1}{O_2/E_2}$$

  where O_i/E_i is estimated relative (excess) hazard in group i

- In practice, better estimate HR by Cox regression

Uninformative censoring

- Subjects who are censored at a given point in time should be as likely to have a subsequent event as those individuals who remain in the study

- Examples for informative censoring: withdrawal from clinical trial because of drug toxicity or worsening clinical condition

- If censoring occurs in a small #pts, impact of informative censoring is likely small

SPSS code (syntax and clicking)

1. Tumor volume study with first local recurrence as event
   Click Analyze – Descriptive Statistics – Frequencies and select variable `volumeg2`. Click Analyze – Compare Means – Means and select variables `Time_LR_first` and `volumeg2`. Click Analyze – Descriptive Statistics – Crosstabs and select variables `First_LR` and `volumeg2`.

```
FREQUENCIES VARIABLES=volumeg2
  /ORDER=ANALYSIS.
MEANS TABLES=Time_LR_first BY volumeg2
  /CELLS MEAN SUM.
CROSSTABS
  /TABLES=First_LR BY volumeg2
  /FORMAT=AVALUE TABLES
  /CELLS=COUNT
  /COUNT ROUND CELL.
```

2. Survival time calculation
   Click Analyze – Survival – Kaplan-Meier amd select variable `Time_LR_first` as Time, `First_LR` as Status with 1 indicating an event and `volumeg2` as Factor. Under Options, select Survival and One Minus Survival in Plots.

```
KM Time_LR_first BY volumeg2
   /STATUS=First_LR(1)
   /PRINT TABLE MEAN
   /PLOT SURVIVAL OMS.
```

3. Does local recurrence differ by tumor volume?
   Click Analyze - Survival - Kaplan-Meier and select variables as above. Request logrank test under Compare Factors.

```
KM Time_LR_first BY volumeg2
  /STATUS=First_LR(1)
  /PRINT NONE
  /TEST LOGRANK
  /COMPARE OVERALL POOLED.
```

4. Does volume measurement modality (CT vs. MRI) confound the results?
   Request Kaplan-Meier plot and logrank test above, but select variable `MRI_CT` as stratification variable.

```
KM Time_LR_first BY volumeg2
  /STRATA=MRI_CT
  /STATUS=First_LR(1)
  /PRINT NONE
  /TEST LOGRANK
  /COMPARE OVERALL POOLED.
```

5. Local recurrence-free survival by 4 tumor volume categories
   Click Analyze - Survival - Kaplan-Meier and select variables as above.
   Request logrank test under Compare Factors.

```
KM Time_LR_first BY volumeg4
  /STATUS=First_LR(1)
  /PRINT NONE
  /TEST LOGRANK
  /COMPARE OVERALL POOLED.
```

6. Survival trend across 4 tumor volume categories
   Click Analyze - Survival - Kaplan-Meier and select variables as above.
   Under Compare Factors, request logrank test and linear trend test.

```
KM Time_LR_first BY volumeg4
  /STATUS=First_LR(1)
  /PRINT NONE
  /TEST LOGRANK
  /TREND
  /COMPARE OVERALL POOLED.
```