

# Solutions for Session 10

Mark Lunt

September 20, 2011

```
. do solution.do

. use http://www.stata-press.com/data/r8/leukemia, clear
(Leukemia Remission Study)

. stset weeks, fail(relapse)
      failure event:  relapse != 0 & relapse < .
obs. time interval:  (0, weeks]
exit on or before:  failure
```

---

```
42 total obs.
0 exclusions
```

---

```
42 obs. remaining, representing
30 failures in single record/single failure data
541 total analysis time at risk, at risk from t = 0
      earliest observed entry t = 0
      last observed exit t = 35
```

```
. sts list if treatment1 == 1
      failure _d: relapse
analysis time _t: weeks
```

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
6	21	3	1	0.8571	0.0764	0.6197	0.9516
7	17	1	0	0.8067	0.0869	0.5631	0.9228
9	16	0	1	0.8067	0.0869	0.5631	0.9228
10	15	1	1	0.7529	0.0963	0.5032	0.8894
11	13	0	1	0.7529	0.0963	0.5032	0.8894
13	12	1	0	0.6902	0.1068	0.4316	0.8491
16	11	1	0	0.6275	0.1141	0.3675	0.8049
17	10	0	1	0.6275	0.1141	0.3675	0.8049
19	9	0	1	0.6275	0.1141	0.3675	0.8049
20	8	0	1	0.6275	0.1141	0.3675	0.8049
22	7	1	0	0.5378	0.1282	0.2678	0.7468
23	6	1	0	0.4482	0.1346	0.1881	0.6801
25	5	0	1	0.4482	0.1346	0.1881	0.6801
32	4	0	2	0.4482	0.1346	0.1881	0.6801
34	2	0	1	0.4482	0.1346	0.1881	0.6801
35	1	0	1	0.4482	0.1346	0.1881	0.6801

---

1.2 At 23 weeks, the survivor function drops from 0.54 to 0.45  
 1.3 Total in the Net Lost column is 12

```
. sts list if treatment1 == 0
      failure _d: relapse
      analysis time _t: weeks
```

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
1	21	2	0	0.9048	0.0641	0.6700	0.9753
2	19	2	0	0.8095	0.0857	0.5689	0.9239
3	17	1	0	0.7619	0.0929	0.5194	0.8933
4	16	2	0	0.6667	0.1029	0.4254	0.8250
5	14	2	0	0.5714	0.1080	0.3380	0.7492
8	12	4	0	0.3810	0.1060	0.1831	0.5778
11	8	2	0	0.2857	0.0986	0.1166	0.4818
12	6	2	0	0.1905	0.0857	0.0595	0.3774
15	4	1	0	0.1429	0.0764	0.0357	0.3212
17	3	1	0	0.0952	0.0641	0.0163	0.2612
22	2	1	0	0.0476	0.0465	0.0033	0.1970
23	1	1	0	0.0000	.	.	.

1.4 8 weeks  
 1.5 None  
 1.6 Median survival before relapse is better on Drug A (23 weeks) than standard treatment (8 weeks)

```
. sts graph, by(treatment1)
      failure _d: relapse
      analysis time _t: weeks
```

```
. graph export graph1.eps replace
(file graph1.eps written in EPS format)
```

1.7 Yes, survival appears to be better on Drug A

```
. sts graph, by(treatment1) yline(0.5)
      failure _d: relapse
      analysis time _t: weeks
```

```
. graph export graph2.eps replace
(file graph2.eps written in EPS format)
```

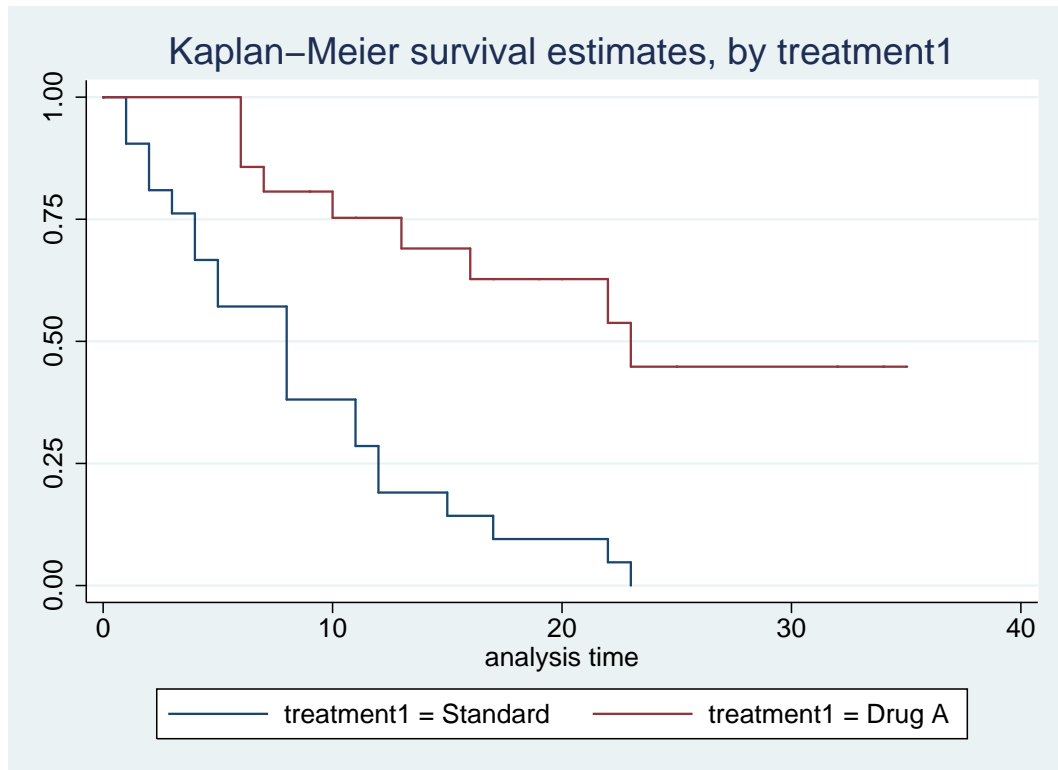


Figure 1: . sts graph, by(treatment1)

```
. sts graph, by(treatment1) yline(0.5) lost
      failure _d: relapse
      analysis time _t: weeks

. graph export graph3.eps replace
(file graph3.eps written in EPS format)

1.9 12 on Drug A, 0 on standard treatment, as before

. sts graph, by(treatment1) yline(0.5) lost gwood
      failure _d: relapse
      analysis time _t: weeks

. graph export graph4.eps replace
(file graph4.eps written in EPS format)
```

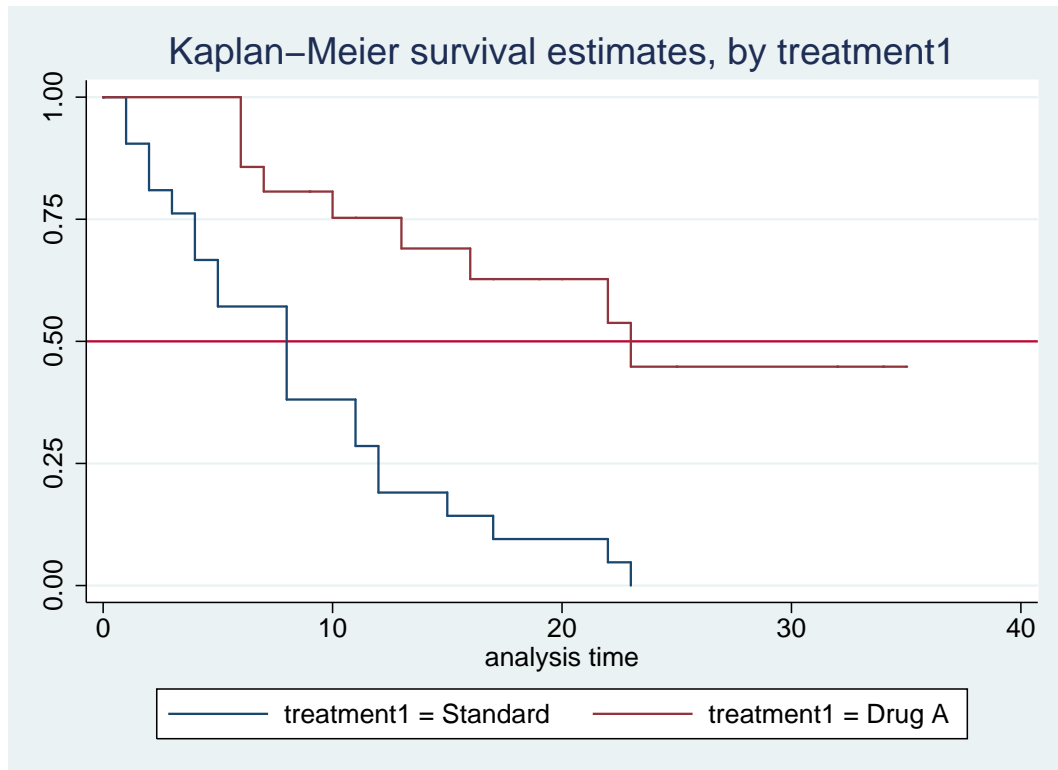


Figure 2: `. sts graph, by(treatment1) yline(0.5)`

1.10 Confidence bands get wider since they are based on smaller numbers

```
. sts test treatment1
      failure _d: relapse
      analysis time _t: weeks

Log-rank test for equality of survivor functions
```

treatment1	Events observed	Events expected
Standard	21	10.75
Drug A	9	19.25
Total	30	30.00

```

      chi2(1) = 16.79
      Pr>chi2 = 0.0000
```

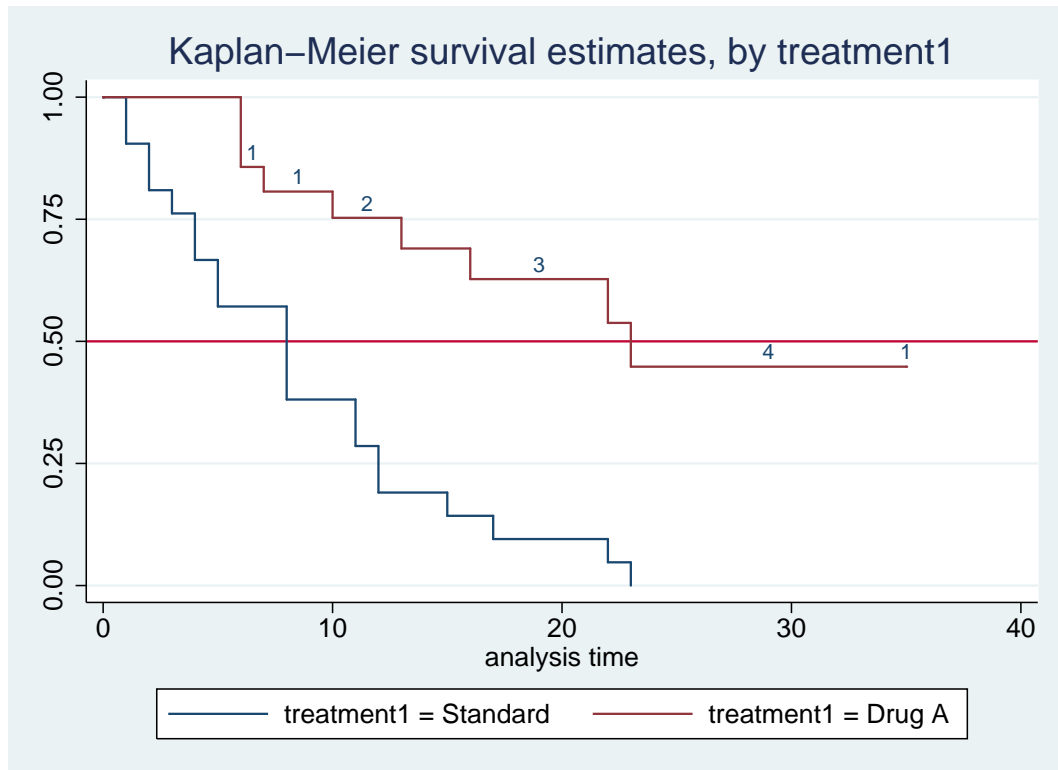


Figure 3: . sts graph, by(treatment1) yline(0.5) lost

1.11 Yes, there are far fewer relapses on Drug A than expected

```
. sts test treatment1, wilcoxon
      failure_d: relapse
      analysis time _t: weeks

Wilcoxon (Breslow) test for equality of survivor functions
```

treatment1	Events observed	Events expected	Sum of ranks
Standard	21	10.75	271
Drug A	9	19.25	-271
Total	30	30.00	0

```

      chi2(1) = 13.46
      Pr>chi2 = 0.0002

```

1.12 Yes, the results are very similar

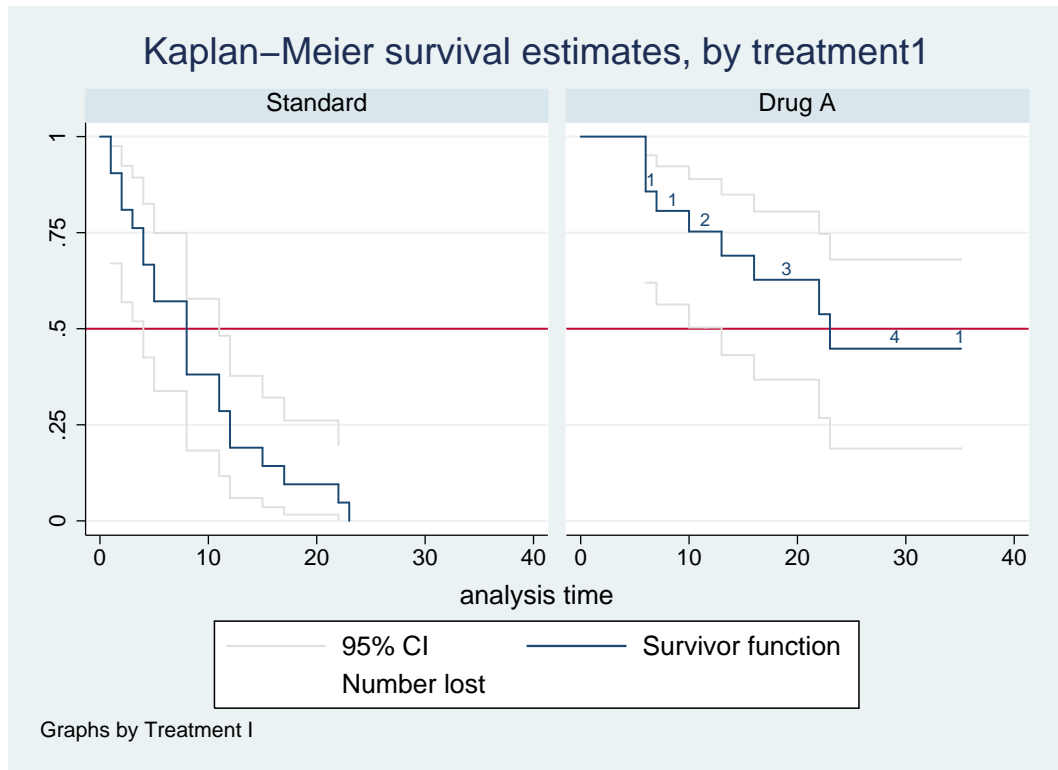


Figure 4: `. sts graph, by(treatment1) yline(0.5) lost gwood`

```

. sts graph, by(wbc3cat)
      failure _d: relapse
      analysis time _t: weeks

. graph export graph5.eps replace
(file graph5.eps written in EPS format)
  
```

*2.1 Yes, survival is best in the Normal group and worst in the High group*

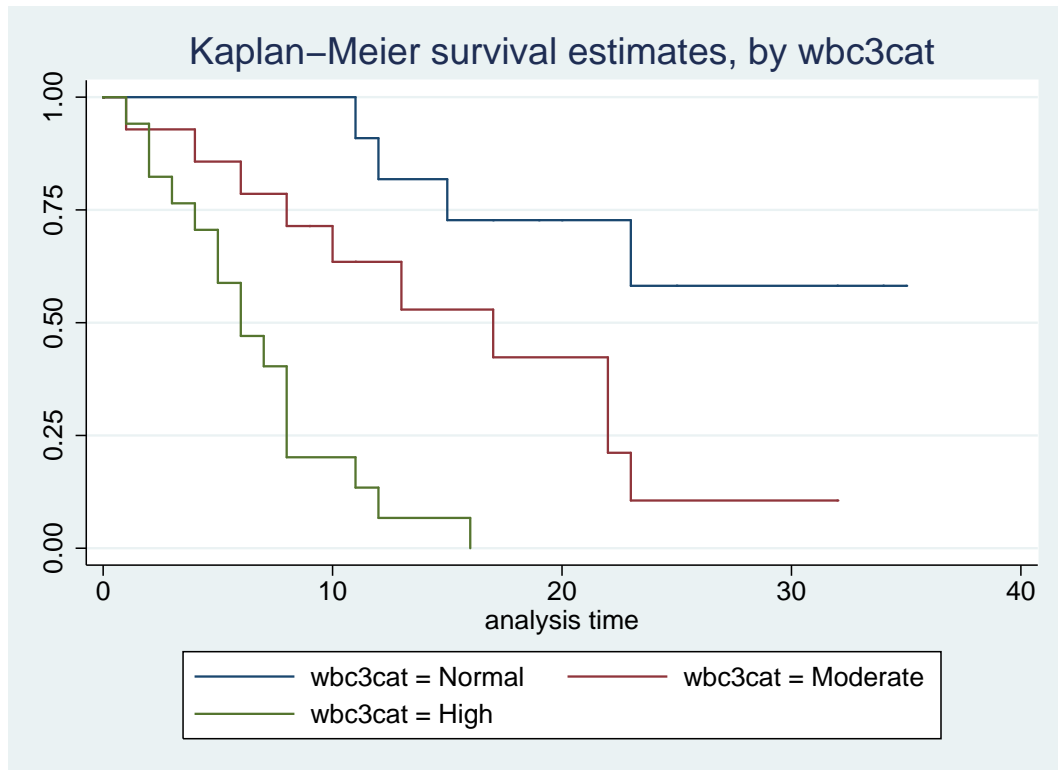


Figure 5: . sts graph, by(wbc3cat)

```
. tab wbc3cat treatment1, co
```

Key			
	<i>frequency</i>	<i>column</i>	<i>percentage</i>
White Blood Cell Count	Treatment I		Total
	Standard	Drug A	
Normal	4 19.05	7 33.33	11 26.19
Moderate	5 23.81	9 42.86	14 33.33
High	12 57.14	5 23.81	17 40.48
Total	21 100.00	21 100.00	42 100.00

2.2 No, there are more in the High category on standard treatment than Drug A

2.3 Survival would be worse in the standard treatment arm

```
. stcox treatment1
      failure _d: relapse
      analysis time _t: weeks
Iteration 0:  log likelihood = -93.98505
Iteration 1:  log likelihood = -86.385606
Iteration 2:  log likelihood = -86.379623
Iteration 3:  log likelihood = -86.379622
Refining estimates:
Iteration 0:  log likelihood = -86.379622
Cox regression -- Breslow method for ties
No. of subjects =          42          Number of obs =          42
No. of failures =          30
Time at risk   =          541
Log likelihood = -86.379622          LR chi2(1) =          15.21
                                          Prob > chi2 =          0.0001
```

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
treatment1	.2210887	.0905501	-3.68	0.000	.0990706	.4933877

2.4 HR = 0.22, 95% CI = 0.10, 0.49

```

. xi: stcox treatment1 i.wbc3cat
i.wbc3cat      _Iwbc3cat_1-3      (naturally coded; _Iwbc3cat_1 omitted)
      failure _d: relapse
      analysis time _t: weeks
Iteration 0:  log likelihood = -93.98505
Iteration 1:  log likelihood = -78.027142
Iteration 2:  log likelihood = -77.480851
Iteration 3:  log likelihood = -77.476906
Refining estimates:
Iteration 0:  log likelihood = -77.476905
Cox regression -- Breslow method for ties
No. of subjects =          42          Number of obs =          42
No. of failures =          30
Time at risk   =          541
Log likelihood = -77.476905          LR chi2(3) =          33.02
                                          Prob > chi2 =          0.0000

```

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
treatment1	.2834551	.1229874	-2.91	0.004	.1211042 .6634517
_Iwbc3cat_2	3.637825	2.201306	2.13	0.033	1.111134 11.91015
_Iwbc3cat_3	10.92214	7.088783	3.68	0.000	3.06093 38.97284

2.5 HR = 0.28, 95% CI = 0.12, 0.66

2.6 The beneficial effect of Drug A was exaggerated by the difference in white blood cell counts between the groups

```

. stcoxkm, by(treatment1)
      failure _d: relapse
      analysis time _t: weeks

. graph export graph6.eps replace
(file graph6.eps written in EPS format)

```

2.7 Yes

```

. stcoxkm, by(wbc3cat)
      failure _d: relapse
      analysis time _t: weeks

. graph export graph7.eps replace
(file graph7.eps written in EPS format)

```

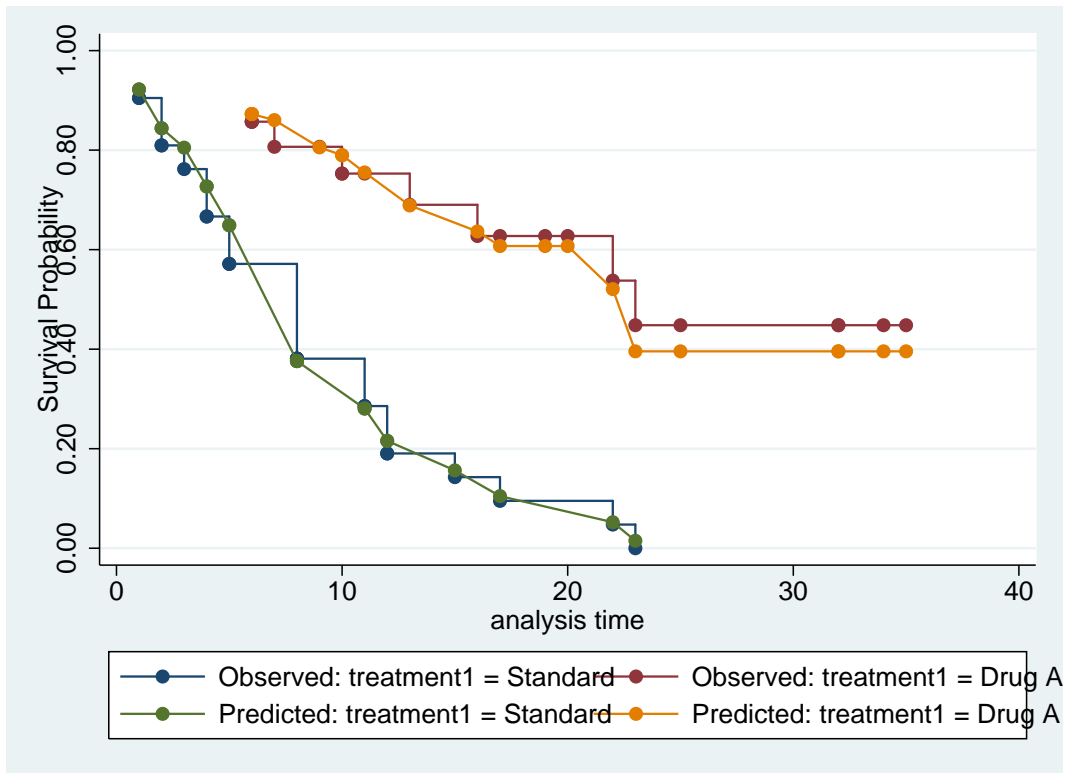


Figure 6: . stcoxkm, by(treatment1)

2.8 Yes

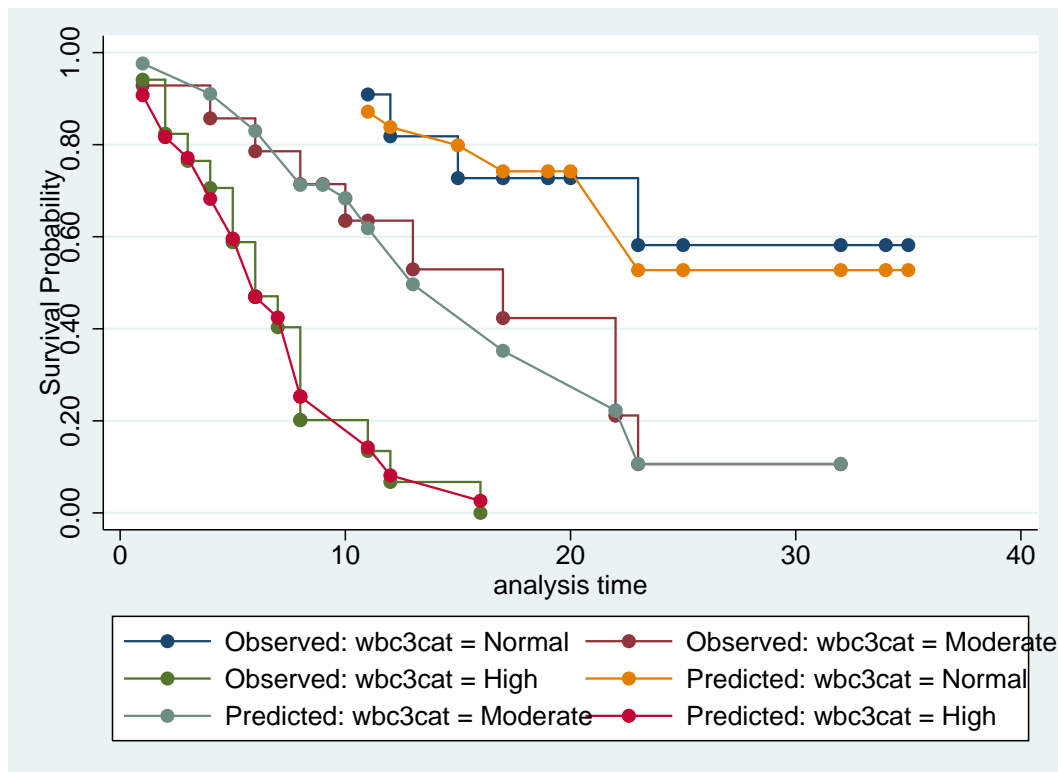


Figure 7: . stcoxkm, by(wbc3cat)

```
. xi: stcox treatment1 i.wbc3cat, sca(sca*) sch(sch*)
i.wbc3cat      _Iwbc3cat_1-3      (naturally coded; _Iwbc3cat_1 omitted)
      failure_d: relapse
      analysis time _t: weeks
Iteration 0:   log likelihood = -93.98505
Iteration 1:   log likelihood = -78.027142
Iteration 2:   log likelihood = -77.480851
Iteration 3:   log likelihood = -77.476906
Refining estimates:
Iteration 0:   log likelihood = -77.476905
Cox regression -- Breslow method for ties
No. of subjects =      42                Number of obs   =      42
No. of failures =      30
Time at risk    =     541
Log likelihood   = -77.476905            LR chi2(3)      =     33.02
                                                Prob > chi2     =     0.0000
```

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
treatment1	.2834551	.1229874	-2.91	0.004	.1211042 .6634517
_Iwbc3cat_2	3.637825	2.201306	2.13	0.033	1.111134 11.91015
_Iwbc3cat_3	10.92214	7.088783	3.68	0.000	3.06093 38.97284

```
. stphtest
```

```
Test of proportional hazards assumption
```

```
Time: Time
```

	chi2	df	Prob>chi2
global test	0.33	3	0.9551

*2.9 Yes*

```
. stphtest, detail
```

```
Test of proportional hazards assumption
```

```
Time: Time
```

	rho	chi2	df	Prob>chi2
treatment1	-0.07019	0.15	1	0.6948
_Iwbc3cat_2	-0.03223	0.03	1	0.8650
_Iwbc3cat_3	0.01682	0.01	1	0.9237
global test		0.33	3	0.9551

*2.10 No, the PH assumption holds for all three variables*

```
. sts graph, by(treatment2)
```

```
failure _d: relapse
```

```
analysis time _t: weeks
```

```
. graph export graph8.eps replace
```

```
(file graph8.eps written in EPS format)
```

*3.1 Survival on drug B is worse than standard treatment for the first 10 weeks*

*3.2 After 10 weeks, survival is better on Drug B than standard treatment*

```
. stcoxkm, by(treatment2)
```

```
failure _d: relapse
```

```
analysis time _t: weeks
```

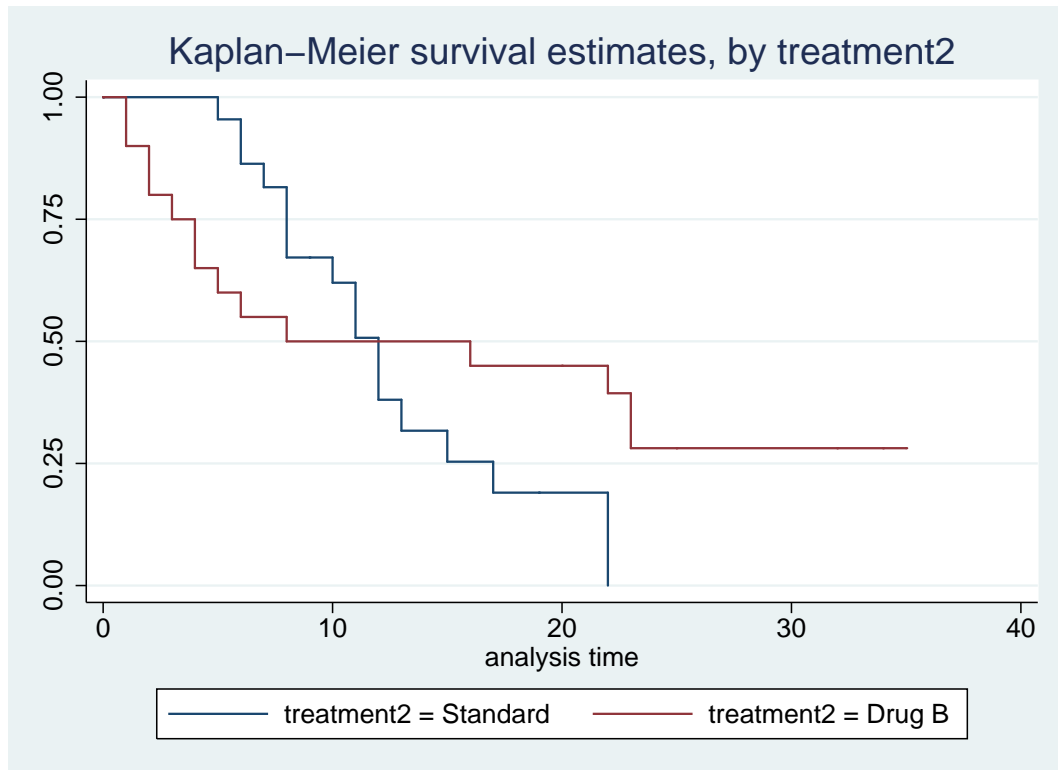


Figure 8: . sts graph, by(treatment2)

```
. graph export graph9.eps replace
(file graph9.eps written in EPS format)
```

3.3 The observed curves cross, the predicted curves do not

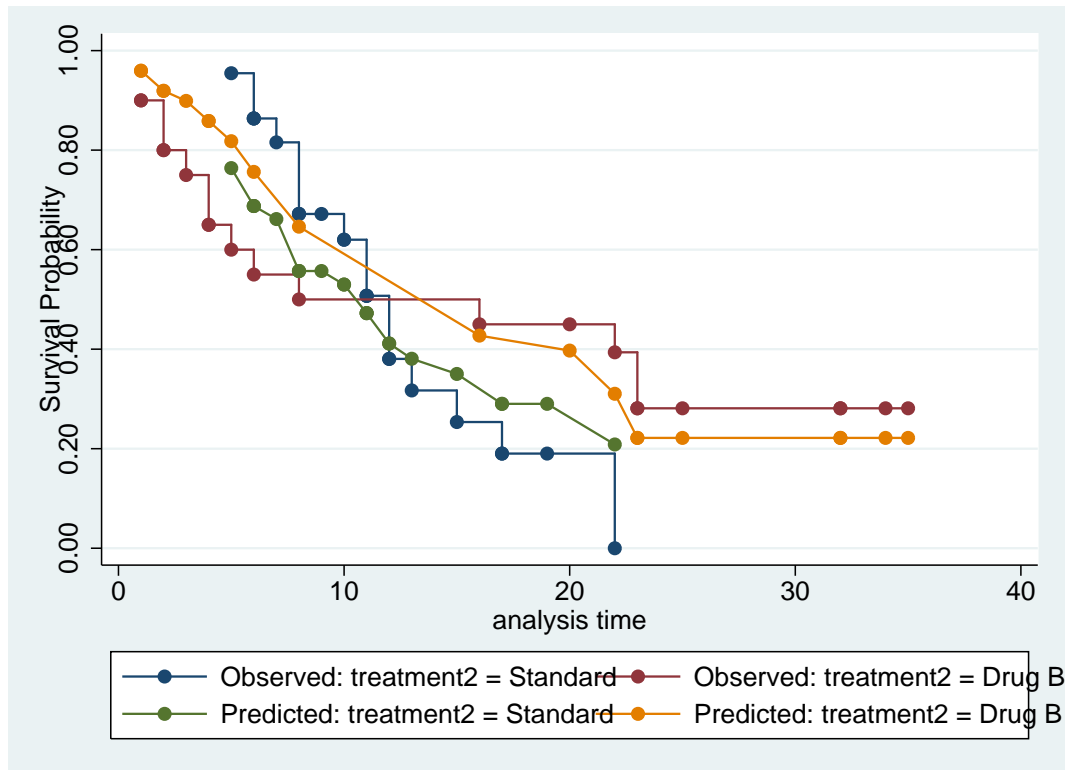


Figure 9: . stcoxkm, by(treatment2)

```
. xi: stcox treatment2 i.wbc3cat
i.wbc3cat      _Iwbc3cat_1-3      (naturally coded; _Iwbc3cat_1 omitted)
      failure_d: relapse
      analysis time _t: weeks
Iteration 0:   log likelihood = -93.98505
Iteration 1:   log likelihood = -82.714241
Iteration 2:   log likelihood = -82.027782
Iteration 3:   log likelihood = -82.019054
Iteration 4:   log likelihood = -82.019053
Refining estimates:
Iteration 0:   log likelihood = -82.019053
Cox regression -- Breslow method for ties
No. of subjects =      42                Number of obs   =      42
No. of failures =      30                Time at risk    =
Time at risk    =      541
Log likelihood   = -82.019053            LR chi2(3)      =      23.93
                                                Prob > chi2     =      0.0000
```

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
treatment2	.8483777	.3469054	-0.40	0.688	.3806529 1.890816
_Iwbc3cat_2	3.409628	2.050784	2.04	0.041	1.048905 11.08353
_Iwbc3cat_3	14.0562	8.873693	4.19	0.000	4.078529 48.44314

3.4 No, the HR is 0.85, but this reduction is not statistically significant

```
. drop sca* sch*

. xi: stcox treatment2 i.wbc3cat, sca(sca*) sch(sch*)
i.wbc3cat      _Iwbc3cat_1-3      (naturally coded; _Iwbc3cat_1 omitted)
      failure _d: relapse
      analysis time _t: weeks
Iteration 0:   log likelihood = -93.98505
Iteration 1:   log likelihood = -82.714241
Iteration 2:   log likelihood = -82.027782
Iteration 3:   log likelihood = -82.019054
Iteration 4:   log likelihood = -82.019053
Refining estimates:
Iteration 0:   log likelihood = -82.019053
Cox regression -- Breslow method for ties
No. of subjects =      42                Number of obs   =      42
No. of failures =      30
Time at risk    =      541
Log likelihood  = -82.019053            LR chi2(3)      =      23.93
                                                Prob > chi2    =      0.0000
```

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
treatment2	.8483777	.3469054	-0.40	0.688	.3806529 1.890816
_Iwbc3cat_2	3.409628	2.050784	2.04	0.041	1.048905 11.08353
_Iwbc3cat_3	14.0562	8.873693	4.19	0.000	4.078529 48.44314

```
. stphtest
      Test of proportional hazards assumption
      Time: Time
```

	chi2	df	Prob>chi2
global test	10.24	3	0.0166

3.5 No: the test shows a significant departure from proportional hazards

```
. stphtest, detail
```

```
Test of proportional hazards assumption
```

```
Time: Time
```

	rho	chi2	df	Prob>chi2
treatment2	-0.51672	10.19	1	0.0014
_Iwbc3cat_2	-0.09860	0.29	1	0.5903
_Iwbc3cat_3	-0.03559	0.04	1	0.8448
global test		10.24	3	0.0166

*3.6 Only the treatment variable does not satisfy the PH assumption*

```
. sts list
```

```
failure _d: relapse
```

```
analysis time _t: weeks
```

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
1	42	2	0	0.9524	0.0329	0.8227	0.9879
2	40	2	0	0.9048	0.0453	0.7658	0.9631
3	38	1	0	0.8810	0.0500	0.7373	0.9486
4	37	2	0	0.8333	0.0575	0.6819	0.9168
5	35	2	0	0.7857	0.0633	0.6286	0.8822
6	33	3	1	0.7143	0.0697	0.5521	0.8265
7	29	1	0	0.6897	0.0715	0.5262	0.8065
8	28	4	0	0.5911	0.0764	0.4269	0.7228
9	24	0	1	0.5911	0.0764	0.4269	0.7228
10	23	1	1	0.5654	0.0773	0.4017	0.7002
11	21	2	1	0.5116	0.0788	0.3495	0.6523
12	18	2	0	0.4547	0.0796	0.2958	0.6006
13	16	1	0	0.4263	0.0795	0.2700	0.5739
15	15	1	0	0.3979	0.0791	0.2449	0.5468
16	14	1	0	0.3695	0.0784	0.2204	0.5191
17	13	1	1	0.3411	0.0774	0.1966	0.4909
19	11	0	1	0.3411	0.0774	0.1966	0.4909
20	10	0	1	0.3411	0.0774	0.1966	0.4909
22	9	2	0	0.2653	0.0765	0.1311	0.4204
23	7	2	0	0.1895	0.0710	0.0753	0.3431
25	5	0	1	0.1895	0.0710	0.0753	0.3431
32	4	0	2	0.1895	0.0710	0.0753	0.3431
34	2	0	1	0.1895	0.0710	0.0753	0.3431
35	1	0	1	0.1895	0.0710	0.0753	0.3431

```
. gen id = _n
```

```
. stset weeks, fail(relapse) id(id)
      id: id
      failure event: relapse != 0 & relapse < .
obs. time interval: (weeks[_n-1], weeks]
exit on or before: failure
```

---

```
42 total obs.
0 exclusions
```

---

```
42 obs. remaining, representing
42 subjects
30 failures in single failure-per-subject data
541 total analysis time at risk, at risk from t = 0
      earliest observed entry t = 0
      last observed exit t = 35
```

```
. stsplot split_time, at(10)
(21 observations (episodes) created)
```

```
. sts list
```

```
      failure _d: relapse
analysis time _t: weeks
      id: id
```

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
1	42	2	0	0.9524	0.0329	0.8227	0.9879
2	40	2	0	0.9048	0.0453	0.7658	0.9631
3	38	1	0	0.8810	0.0500	0.7373	0.9486
4	37	2	0	0.8333	0.0575	0.6819	0.9168
5	35	2	0	0.7857	0.0633	0.6286	0.8822
6	33	3	1	0.7143	0.0697	0.5521	0.8265
7	29	1	0	0.6897	0.0715	0.5262	0.8065
8	28	4	0	0.5911	0.0764	0.4269	0.7228
9	24	0	1	0.5911	0.0764	0.4269	0.7228
10	23	1	1	0.5654	0.0773	0.4017	0.7002
11	21	2	1	0.5116	0.0788	0.3495	0.6523
12	18	2	0	0.4547	0.0796	0.2958	0.6006
13	16	1	0	0.4263	0.0795	0.2700	0.5739
15	15	1	0	0.3979	0.0791	0.2449	0.5468
16	14	1	0	0.3695	0.0784	0.2204	0.5191
17	13	1	1	0.3411	0.0774	0.1966	0.4909
19	11	0	1	0.3411	0.0774	0.1966	0.4909
20	10	0	1	0.3411	0.0774	0.1966	0.4909
22	9	2	0	0.2653	0.0765	0.1311	0.4204
23	7	2	0	0.1895	0.0710	0.0753	0.3431
25	5	0	1	0.1895	0.0710	0.0753	0.3431
32	4	0	2	0.1895	0.0710	0.0753	0.3431
34	2	0	1	0.1895	0.0710	0.0753	0.3431
35	1	0	1	0.1895	0.0710	0.0753	0.3431

3.10 Yes

```
. list id weeks relapse split_time _t0 _t in 1/45
```

	id	weeks	relapse	split_~e	_t0	_t
1.	1	1	1:yes	0	0	1
2.	2	1	1:yes	0	0	1
3.	3	2	1:yes	0	0	2
4.	4	2	1:yes	0	0	2
5.	5	3	1:yes	0	0	3
6.	6	4	1:yes	0	0	4
7.	7	4	1:yes	0	0	4
8.	8	5	1:yes	0	0	5
9.	9	5	1:yes	0	0	5
10.	10	6	1:yes	0	0	6
11.	11	6	0:no	0	0	6
12.	12	6	1:yes	0	0	6
13.	13	6	1:yes	0	0	6
14.	14	7	1:yes	0	0	7
15.	15	8	1:yes	0	0	8
16.	16	8	1:yes	0	0	8
17.	17	8	1:yes	0	0	8
18.	18	8	1:yes	0	0	8
19.	19	9	0:no	0	0	9
20.	20	10	1:yes	0	0	10
21.	21	10	0:no	0	0	10
22.	22	10	.	0	0	10
23.	22	11	1:yes	10	10	11
24.	23	10	.	0	0	10
25.	23	11	0:no	10	10	11
26.	24	10	.	0	0	10
27.	24	11	1:yes	10	10	11
28.	25	10	.	0	0	10
29.	25	12	1:yes	10	10	12
30.	26	10	.	0	0	10
31.	26	12	1:yes	10	10	12
32.	27	10	.	0	0	10
33.	27	13	1:yes	10	10	13
34.	28	10	.	0	0	10
35.	28	15	1:yes	10	10	15
36.	29	10	.	0	0	10
37.	29	16	1:yes	10	10	16
38.	30	10	.	0	0	10
39.	30	17	1:yes	10	10	17
40.	31	10	.	0	0	10
41.	31	17	0:no	10	10	17
42.	32	10	.	0	0	10
43.	32	19	0:no	10	10	19
44.	33	10	.	0	0	10
45.	33	20	0:no	10	10	20

```
. gen t1 = treatment2 * (split_time == 0)
```

```

. gen t2 = treatment2 * (split_time == 10)

. xi: stcox t1 t2 i.wbc3cat
i.wbc3cat      _Iwbc3cat_1-3      (naturally coded; _Iwbc3cat_1 omitted)
      failure _d: relapse
      analysis time _t: weeks
      id: id
Iteration 0:  log likelihood = -93.98505
Iteration 1:  log likelihood = -76.779578
Iteration 2:  log likelihood = -76.392334
Iteration 3:  log likelihood = -76.389699
Iteration 4:  log likelihood = -76.389699
Refining estimates:
Iteration 0:  log likelihood = -76.389699
Cox regression -- Breslow method for ties
No. of subjects =          42      Number of obs =          63
No. of failures =           30
Time at risk   =          541
Log likelihood = -76.389699      LR chi2(4) =          35.19
                                Prob > chi2 =          0.0000

```

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
t1	2.134668	1.02538	1.58	0.114	.832643	5.472702
t2	.1139286	.0943624	-2.62	0.009	.0224707	.577628
_Iwbc3cat_2	3.053347	1.8602	1.83	0.067	.9251154	10.07758
_Iwbc3cat_3	14.31174	9.109685	4.18	0.000	4.110428	49.8308

*3.13 HR for the first 10 weeks == 2.13 (95% CI = 0.83, 5.47)*  
*3.14 HR for after 10 weeks = 0.11 (95% CI = 0.02, 0.58)*  
*3.15 Yes: risk was elevated at first, then reduced*

```

. drop sca* sch*

```

```
. xi: stcox t1 t2 i.wbc3cat, sca(sca*) sch(sch*)
i.wbc3cat      _Iwbc3cat_1-3      (naturally coded; _Iwbc3cat_1 omitted)
      failure _d: relapse
      analysis time _t: weeks
      id: id
```

```
Iteration 0: log likelihood = -93.98505
Iteration 1: log likelihood = -76.779578
Iteration 2: log likelihood = -76.392334
Iteration 3: log likelihood = -76.389699
Iteration 4: log likelihood = -76.389699
```

Refining estimates:

```
Iteration 0: log likelihood = -76.389699
```

Cox regression -- Breslow method for ties

```
No. of subjects =      42      Number of obs   =      63
No. of failures =      30
Time at risk    =      541
LR chi2(4)      =      35.19
Log likelihood  = -76.389699  Prob > chi2   =      0.0000
```

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
t1	2.134668	1.02538	1.58	0.114	.832643 5.472702
t2	.1139286	.0943624	-2.62	0.009	.0224707 .577628
_Iwbc3cat_2	3.053347	1.8602	1.83	0.067	.9251154 10.07758
_Iwbc3cat_3	14.31174	9.109685	4.18	0.000	4.110428 49.8308

```
. stphtest
```

Test of proportional hazards assumption

Time: Time

	chi2	df	Prob>chi2
global test	1.84	4	0.7651

*3.16 Yes, the model is now appropriate*

```
. stphtest, det
```

Test of proportional hazards assumption

Time: Time

	rho	chi2	df	Prob>chi2
t1	-0.20497	1.09	1	0.2971
t2	0.07906	0.16	1	0.6907
_Iwbc3cat_2	-0.11753	0.40	1	0.5261
_Iwbc3cat_3	-0.02192	0.01	1	0.9044
global test		1.84	4	0.7651

*3.17 None of the variables depart from the PH assumption  
end of do-file*