

# Notes on Survival Analysis

Jae-Woo Kim

SPSS ver.11.5 DOES NOT provide commands to run parametric models, frailty models or models using a counting process data layout for recurrent models!

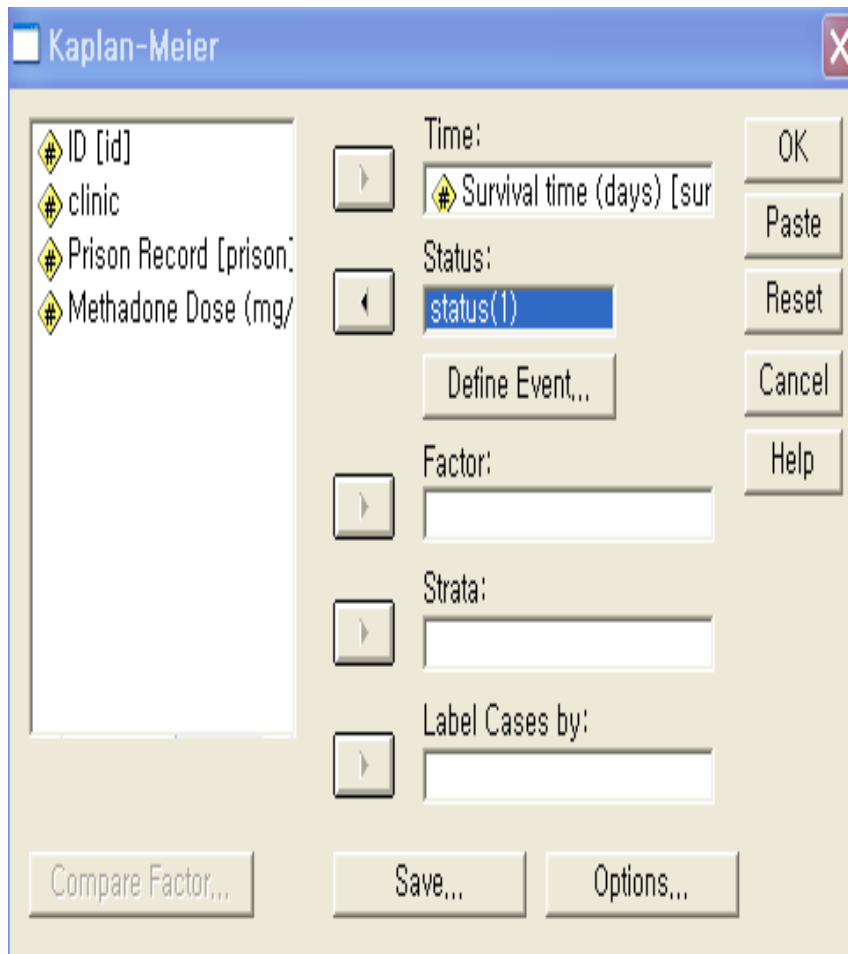
# Using addicts.sav

	id	clinic	status	survt	prison	dose
1	1.00	1.00	1.00	428.00	.00	50.00
2	2.00	1.00	1.00	275.00	1.00	55.00
3	3.00	1.00	1.00	262.00	.00	55.00
4	4.00	1.00	1.00	183.00	.00	30.00
5	5.00	1.00	1.00	259.00	1.00	65.00
6	6.00	1.00	1.00	714.00	.00	55.00
7	7.00	1.00	1.00	438.00	1.00	65.00
8	8.00	1.00	.00	796.00	1.00	60.00
9	9.00	1.00	1.00	892.00	.00	50.00
10	10.00	1.00	1.00	393.00	1.00	65.00
11	11.00	1.00	.00	161.00	1.00	80.00
12	12.00	1.00	1.00	836.00	1.00	60.00
13	13.00	1.00	1.00	523.00	.00	55.00
14	14.00	1.00	1.00	612.00	.00	70.00
15	15.00	1.00	1.00	212.00	1.00	60.00
16	16.00	1.00	1.00	399.00	1.00	60.00
17	17.00	1.00	1.00	771.00	1.00	75.00
18	18.00	1.00	1.00	514.00	1.00	80.00
19	19.00	1.00	1.00	512.00	.00	80.00
20	21.00	1.00	1.00	624.00	1.00	80.00
21	22.00	1.00	1.00	209.00	1.00	60.00
22	23.00	1.00	1.00	341.00	1.00	60.00

# What can I do in SPSS?

- Estimating survival functions (unadjusted) and comparing them across strata
- Assessing ph assumption using KM log-log survival curves
- Running Cox ph model
- Running a stratified Cox model and obtaining Cox adjusted log-log curves
- Assessing ph assumption with a statistical test
- Running an extended Cox model

1. Estimating survival functions (unadjusted) and comparing them across strata
- \* Life table approach in SPSS presents a Display Time Intervals box. This allows the user to define the time intervals used for the life table analysis.

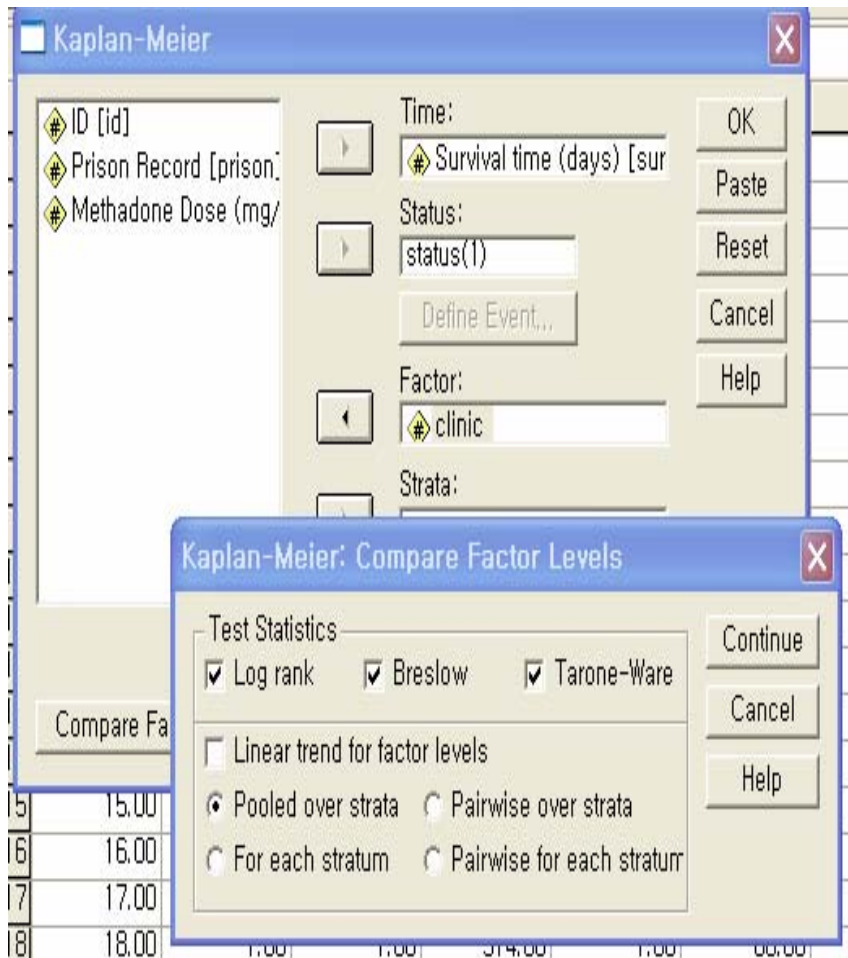


- 1 inserted in the box because STATUS is coded 1 for events

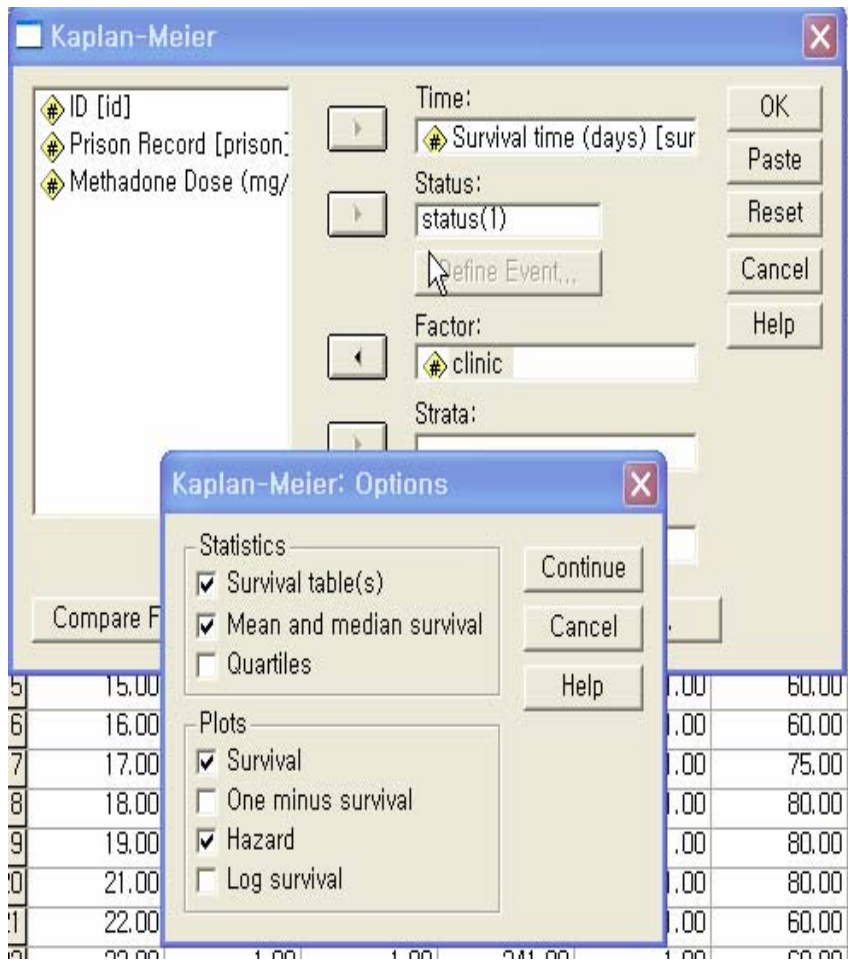
## → Kaplan-Meier

Survival Analysis for SURVT Survival time (days)

Time	Status	Cumulative Survival	Standard Error	Cumulative Events	Number Remaining
2.00	censored			0	237
2.00	censored			0	236
7.00	endpoint	.9958	.0042	1	235
13.00	endpoint	.9915	.0060	2	234
17.00	endpoint	.9873	.0073	3	233
19.00	endpoint	.9831	.0084	4	232
26.00	endpoint	.9788	.0094	5	231
28.00	censored			5	230
28.00	censored			5	229
29.00	endpoint	.9745	.0103	6	228
30.00	endpoint	.9703	.0111	7	227
33.00	endpoint	.9660	.0118	8	226
35.00	endpoint			9	225
35.00	endpoint	.9574	.0132	10	224
37.00	endpoint	.9532	.0138	11	223
41.00	endpoint			12	222
41.00	endpoint	.9446	.0149	13	221
47.00	endpoint	.9403	.0155	14	220
49.00	endpoint	.9361	.0160	15	219
50.00	endpoint	.9318	.0165	16	218
53.00	censored			16	217
53.00	censored			16	216
59.00	endpoint	.9275	.0169	17	215
62.00	endpoint	.9232	.0174	18	214
62.00	censored			18	213
67.00	endpoint	.9188	.0179	19	212

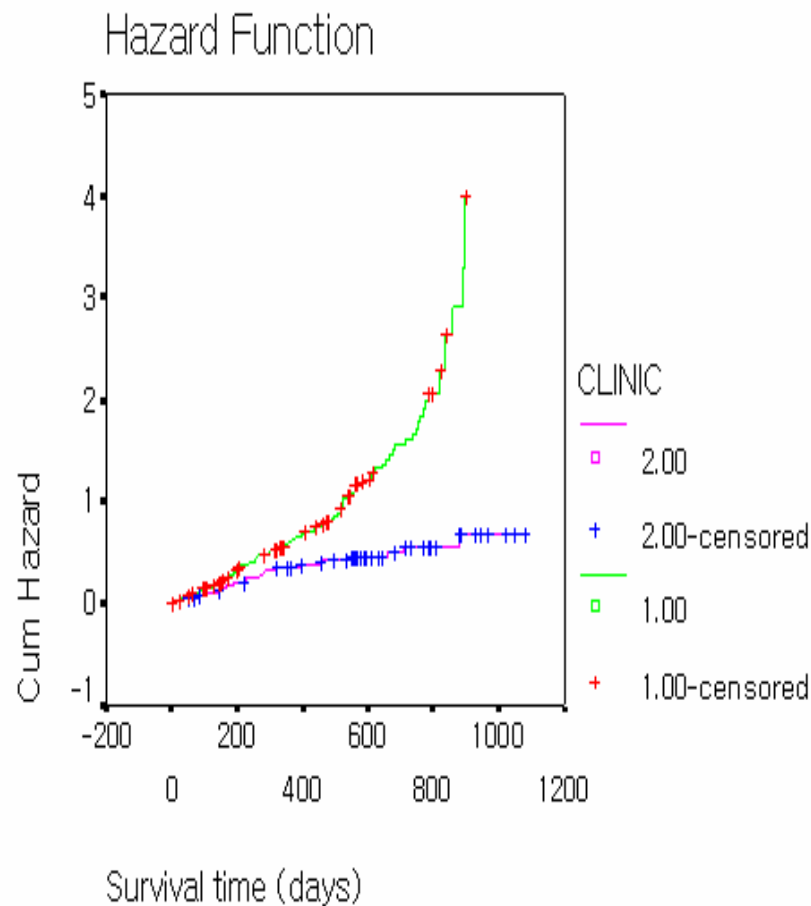
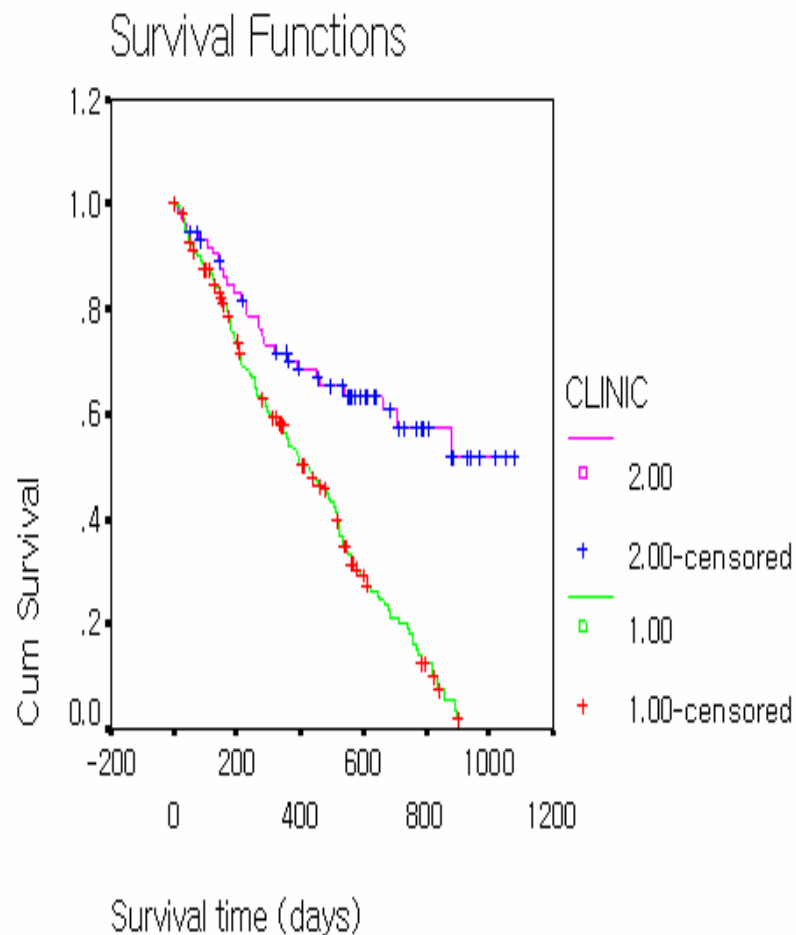


- To obtain KM survival estimates and plots by CLINIC as well as log rank (and other) test statistics, `Survt by clinic/ status=status(1)`
- Breslow in SPSS is the same with Wilcoxon test in Stata (or SAS)



- There are four choices here, but no log-log survival plots in SPSS.

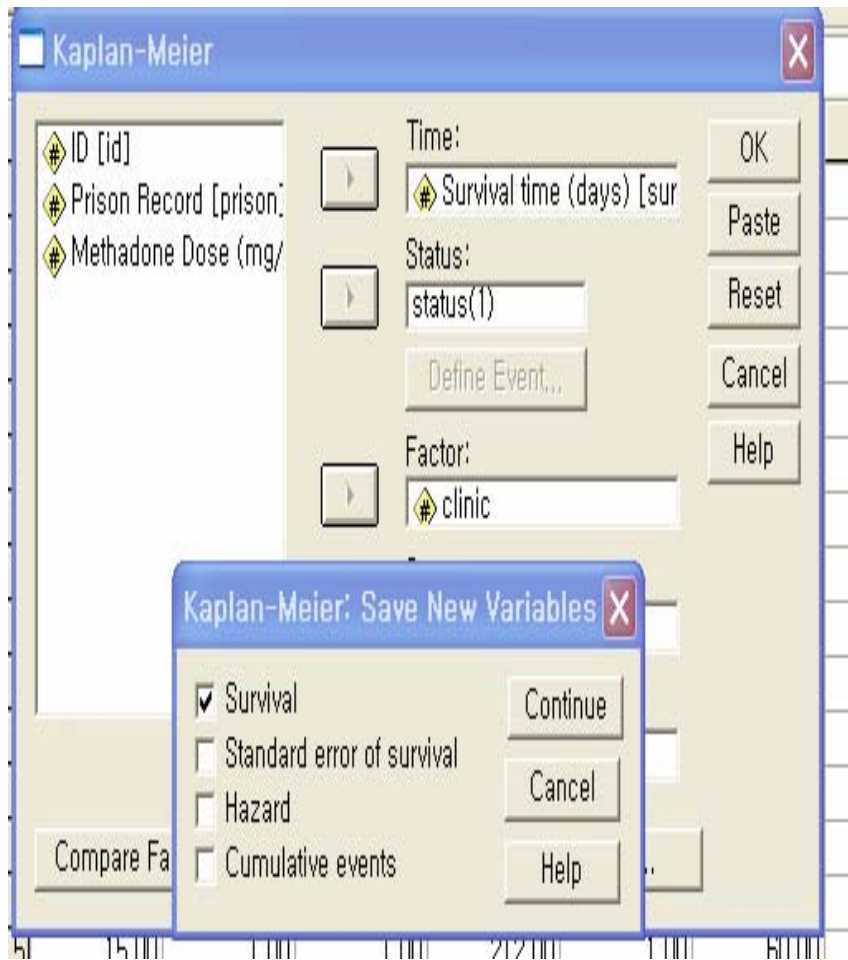
# I don't know why, but no result of test statistics for equality of survival distribution for CLINIC



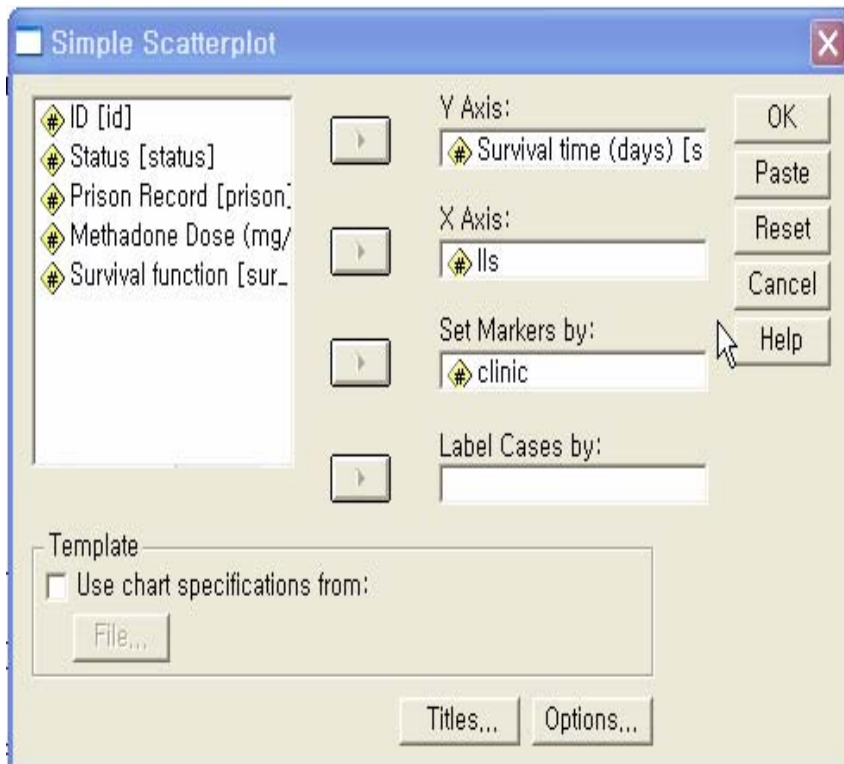
## 2. Assessing PH assumption using KM log-log survival curves

- \* SPSS does not provide unadjusted KM log-log curves, instead provide adjusted log-log curves from running a stratified Cox model. A log-log curve equivalent to the unadjusted KM log-log curve can be obtained in SPSS by running a stratified Cox without including any covariates in the model.

# Save new variables (sur\_1)

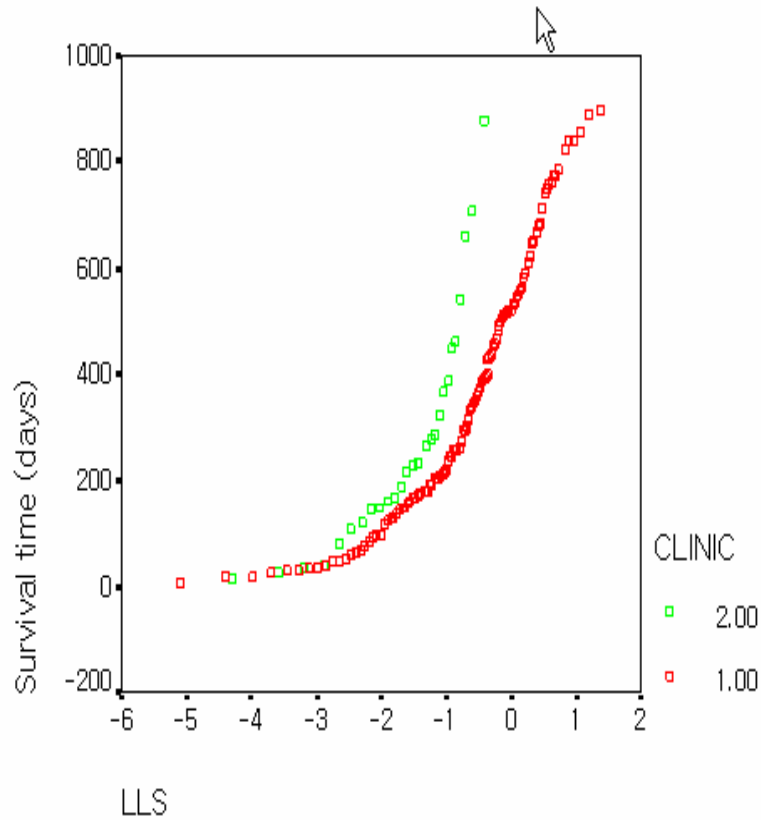


- To create a new variable called lls containing the  $\log(-\log)$  of sur\_1:
- Compute lls= $\ln(-\ln(\text{sur}_1))$ . Execute.
- Or, use Transform->Compute



- To plot lls against survival time

→ Graph



- Approximately parallel curves support PH assumption for CLINIC.

### 3. Running a Cox PH model

➔ Cox Regression

**Case Processing Summary**

		N	Percent
Cases available in analysis	Event <sup>a</sup>	150	63,0%
	Censored	86	36,1%
	Total	236	99,2%
Cases dropped	Cases with missing values	0	,0%
	Cases with negative time	0	,0%
	Censored cases before the earliest event in a stratum	2	,8%
	Total	2	,8%
	Total	238	100,0%

a. Dependent Variable: Survival time (days)

Block 1: Method = Enter

**Omnibus Tests of Model Coefficients<sup>a,b</sup>**

-2 Log Likelihood	Overall (score)			Change From Previous Step			Change From Previous Block		
	Chi-square	df	Sig.	Chi-square	df	Sig.	Chi-square	df	Sig.
1346,805	56,273	3	,000	64,519	3	,000	64,519	3	,000

a. Beginning Block Number 0, initial Log Likelihood function: -2 Log likelihood: 1411,324

b. Beginning Block Number 1, Method = Enter

**Variables in the Equation**

	B	SE	Wald	df	Sig.	Exp(B)
PRISON	,327	,167	3,813	1	,051	1,386
DOSE	-,035	,006	30,785	1	,000	,965
CLINIC	-1,009	,215	22,045	1	,000	,365

**Block 0: Beginning Block**

**Omnibus Tests of Model Coefficients**

-2 Log Likelihood	1411,324
-------------------	----------

**Covariate Means**

	Mean
PRISON	,462
DOSE	60,487
CLINIC	1,314

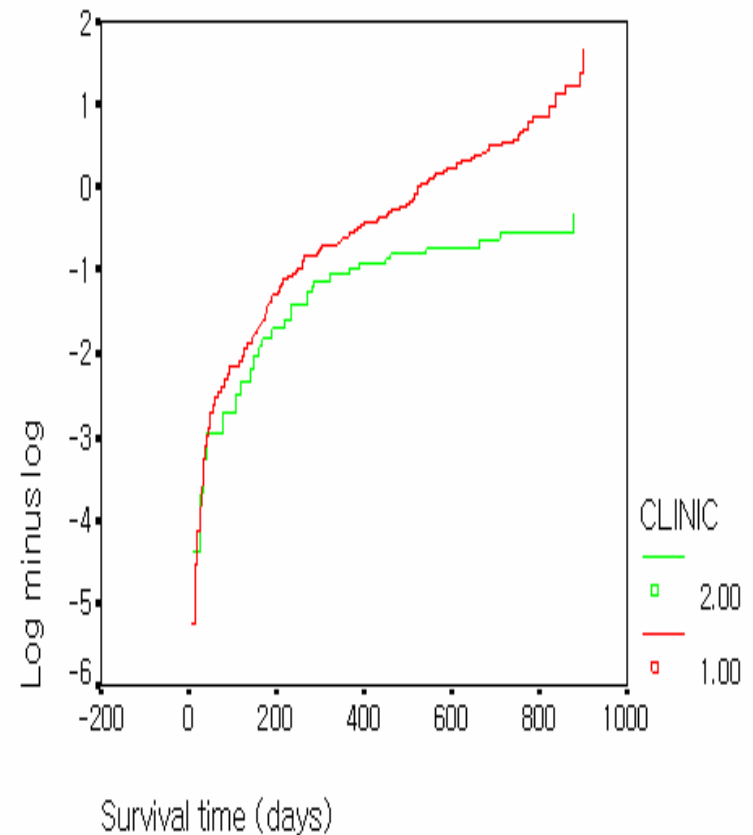
- #### 4. Running a stratified Cox model and Obtaining Cox adjusted log-log curves
- PRISON and DOSE -> Covariate
  - CLINIC -> Strata
  - Click the plots button and check log minus log as the plot type
  - Cox adjusted log-log plots are fitted using the mean values of PRISON and DOSE and used to evaluate PH assumption for CLINIC.

### Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)
PRISON	.389	.169	5.298	1	.021	1.475
DOSE	-.035	.006	29.552	1	.000	.965



LML Function at mean of covariates

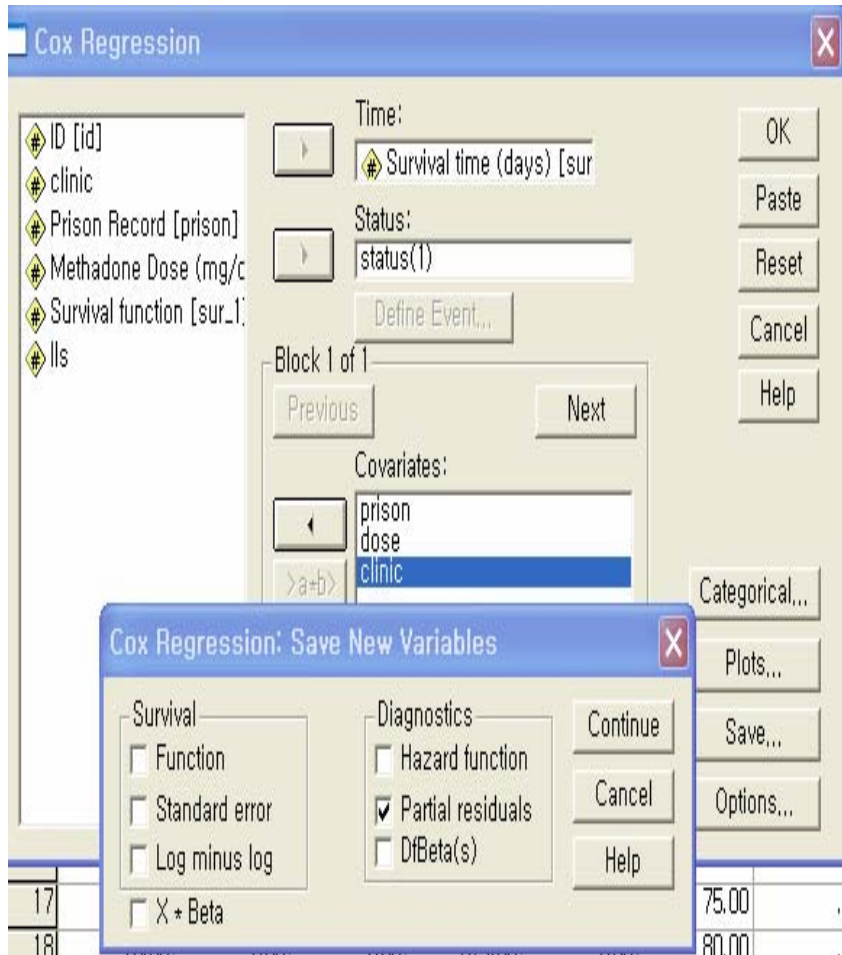


## 5. Assessing PH assumption with a statistical test

- \* SPSS does not easily accommodate a statistical test on PH assumption using the Schoenfeld residuals. However, this can be programmed as follows.

- Run a Cox model to obtain Schoenfeld residuals for all covariates AND save them
- Delete observations censored
- Create a variable that contains the ranked order of survival time.
- Run correlations on the survival rankings with the Schoenfeld residuals
- The p-value for testing whether the correlation is zero between the two is the p-value for the statistical test that PH assumption is violated. The null is that PH assumption is not violated. (Null rejected only for CLINIC)

# pr1\_1, pr2\_1, pr3\_1 for three covariates (partial residuals)



**Cox Regression**

Time: Survival time (days) [sur]

Status: status(1)

Block 1 of 1

Covariates: prison, dose, clinic

**Cox Regression: Save New Variables**

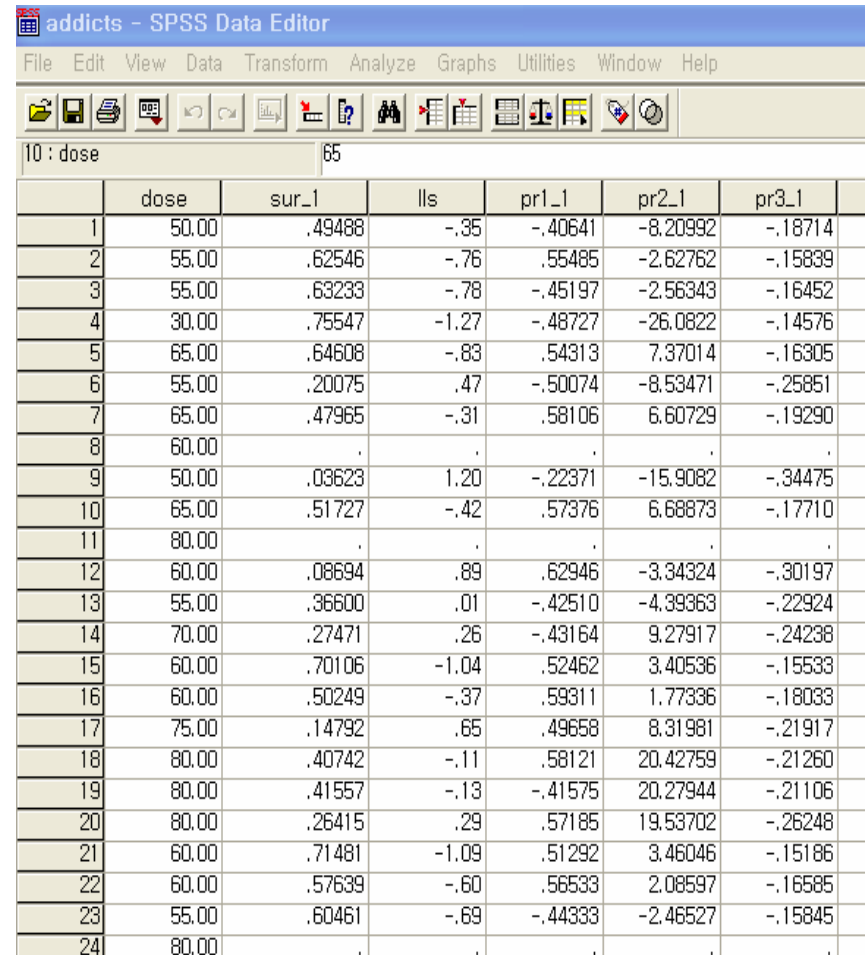
Survival

- Function
- Standard error
- Log minus log
- X \* Beta

Diagnostics

- Hazard function
- Partial residuals
- DiBeta(s)

Buttons: OK, Paste, Reset, Cancel, Help, Categorical..., Plots..., Save..., Options...



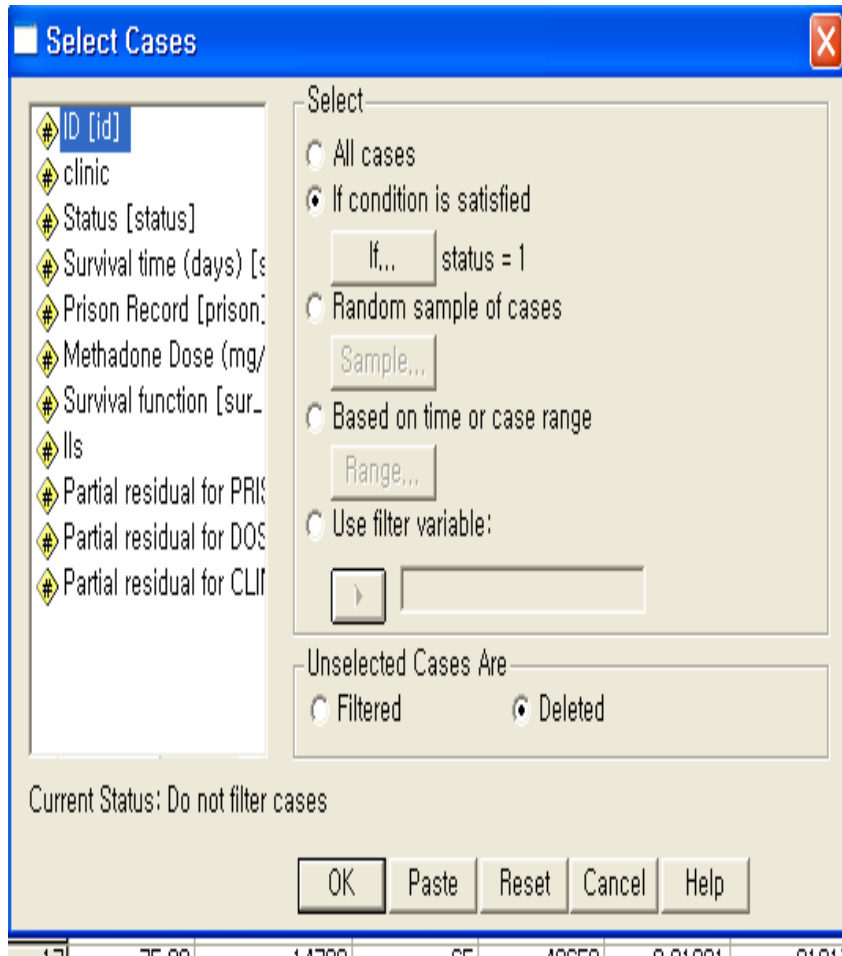
addicts - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

10 : dose 65

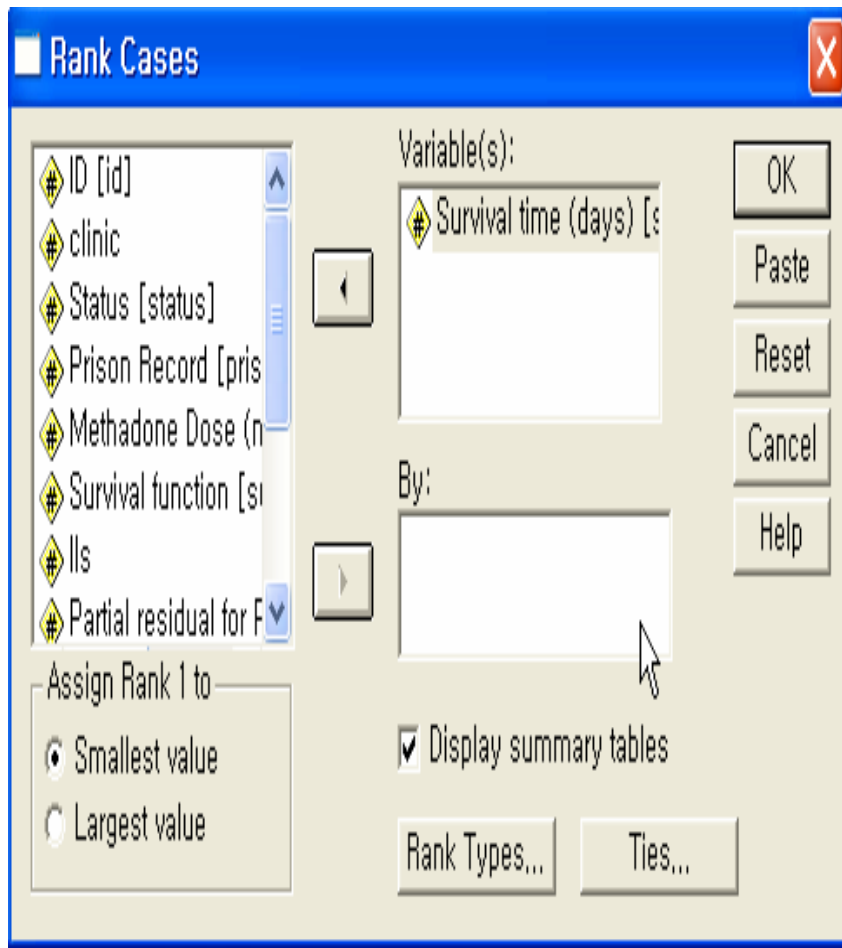
	dose	sur_1	lls	pr1_1	pr2_1	pr3_1
1	50.00	.49488	-.35	-.40641	-8.20992	-.18714
2	55.00	.62546	-.76	.55485	-2.62762	-.15839
3	55.00	.63233	-.78	-.45197	-2.56343	-.16452
4	30.00	.75547	-1.27	-.48727	-26.0822	-.14576
5	65.00	.64608	-.83	.54313	7.37014	-.16305
6	55.00	.20075	.47	-.50074	-8.53471	-.25851
7	65.00	.47965	-.31	.58106	6.60729	-.19290
8	60.00	.	.	.	.	.
9	50.00	.03623	1.20	-.22371	-15.9082	-.34475
10	65.00	.51727	-.42	.57376	6.68873	-.17710
11	80.00	.	.	.	.	.
12	60.00	.08694	.89	.62946	-3.34324	-.30197
13	55.00	.36600	.01	-.42510	-4.39363	-.22924
14	70.00	.27471	.26	-.43164	9.27917	-.24238
15	60.00	.70106	-1.04	.52462	3.40536	-.15533
16	60.00	.50249	-.37	.59311	1.77336	-.18033
17	75.00	.14792	.65	.49658	8.31981	-.21917
18	80.00	.40742	-.11	.58121	20.42759	-.21260
19	80.00	.41557	-.13	-.41575	20.27944	-.21106
20	80.00	.26415	.29	.57185	19.53702	-.26248
21	60.00	.71481	-1.09	.51292	3.46046	-.15186
22	60.00	.57639	-.60	.56533	2.08597	-.16585
23	55.00	.60461	-.69	-.44333	-2.46527	-.15845
24	80.00	.	.	.	.	.

# Delete censored cases



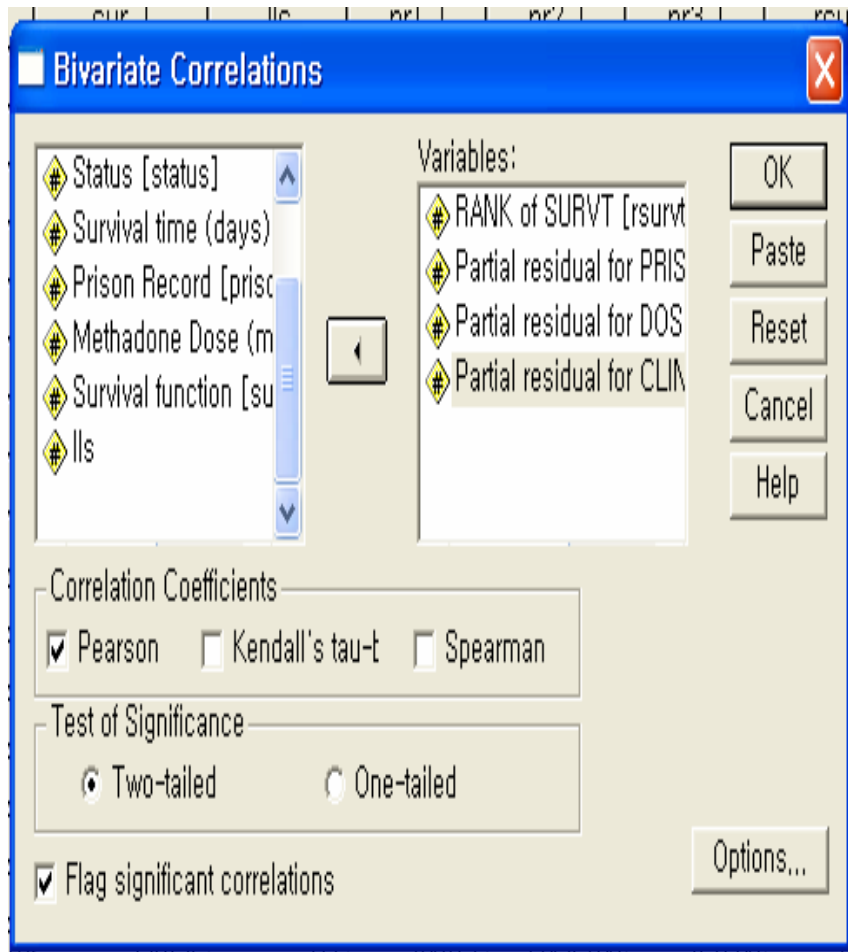
- Checked Deleted in the box called Unselected Cases Are.

# Create a new var including survival ranking (Transform->Ranked cases)



	sur_1	lls	pr1_1	pr2_1	pr3_1	rsurv
1	.49488	-.35	-.40641	-8.20992	-.18714	97.000
2	.62546	-.76	.55485	-2.62762	-.15839	74.000
3	.63233	-.78	-.45197	-2.56343	-.16452	70.500
4	.75547	-1.27	-.48727	-26.0822	-.14576	49.000
5	.64608	-.83	.54313	7.37014	-.16305	69.000
6	.20075	.47	-.50074	-8.53471	-.25851	135.000
7	.47965	-.31	.58106	6.60729	-.19290	99.000
8	.03623	1.20	-.22371	-15.9082	-.34475	149.000
9	.51727	-.42	.57376	6.68873	-.17710	94.000
10	.08694	.89	.62946	-3.34324	-.30197	145.000
11	.36600	.01	-.42510	-4.39363	-.22924	114.500
12	.27471	.26	-.43164	9.27917	-.24238	125.500
13	.70106	-1.04	.52462	3.40536	-.15533	57.500
14	.50249	-.37	.59311	1.77336	-.18033	96.000
15	.14792	.65	.49658	8.31981	-.21917	140.000
16	.40742	-.11	.58121	20.42759	-.21260	110.000
17	.41557	-.13	-.41575	20.27944	-.21106	109.000
18	.26415	.29	.57185	19.53702	-.26248	127.000
19	.71481	-1.09	.51292	3.46046	-.15186	56.000
20	.57639	-.60	.56533	2.08597	-.16585	84.000
21	.60461	-.69	-.44333	-2.46527	-.15845	79.000
22	.63233	-.78	.54803	7.43657	-.16452	70.500
23	.53944	-.48	.56072	-3.20274	-.17300	90.000
24	.59766	-.66	.55223	-7.48997	-.16004	80.000

# Correlation (ranked survival time and three partial residuals)



→ Correlations

**Correlations**

		RANK of SURVT	Partial residual for PRISON	Partial residual for DOSE	Partial residual for CLINIC
RANK of SURVT	Pearson Correlation	1	-.080	.077	-.262**
	Sig. (2-tailed)	.	.332	.347	.001
	N	150	150	150	150
Partial residual for PRISON	Pearson Correlation	-.080	1	.171*	.010
	Sig. (2-tailed)	.332	.	.037	.904
	N	150	150	150	150
Partial residual for DOSE	Pearson Correlation	.077	.171*	1	.023
	Sig. (2-tailed)	.347	.037	.	.776
	N	150	150	150	150
Partial residual for CLINIC	Pearson Correlation	-.262**	.010	.023	1
	Sig. (2-tailed)	.001	.904	.776	.
	N	150	150	150	150

\*\* . Correlation is significant at the 0,01 level (2-tailed),

\* . Correlation is significant at the 0,05 level (2-tailed),

## 6. Running an extended Cox model

- \* An extended Cox model with exactly one tvc can be run using the point and click approach. Defines a tvc (T\_COV\_). If your tvc is the product of tvc DOSE and the log of survival time, then  $\ln(T\_)*DOSE$

### Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)
PRISON	.357	.171	4.337	1	.037	1.428
CLINIC	.255	.214	1.424	1	.233	1.291
DOSE	-.113	.039	8.584	1	.003	.893
T_COV_	.017	.007	5.406	1	.020	1.017

**Compute Time-Dependent Covariate**

Expression for T\_COV\_:

LN(T\_) \* dose

Functions:

- ABS(numexpr)
- ANY(test,value,value,...)
- ARSIN(numexpr)
- ARTAN(numexpr)
- CDFNORM(zvalue)
- CDF,BERNOULLI(q,p)

Model... Reset Cancel Help

**Time-Dependent Cox Regression**

Time: survt

Status: status(1)

Block 1 of 1

Covariates: prison, clinic, dose, T\_COV\_

Method: Enter

Strata:

OK Paste Reset Cancel Help Previous Next Categorical... Save... Options...

**Variables in the Equation**

	B	SE	Wald	df	Sig.	Exp(B)
PRISON	.396	.173	5.226	1	.022	1.486
CLINIC	.392	.256	2.349	1	.125	1.480
DOSE	-.026	.007	11.815	1	.001	.975
T_COV_	-.478	.465	1.055	1	.304	.620

- Heavyside function does not contribute to estimated hazard ratio until day 365. The estimated ratio for CLINIC at 100 days is  $\exp(0.392)$  and that ratio AT 400 days is  $\exp(0.392 - 0.478)$
- A Heavyside function for CLINIC can be created similarly. E.g. Define T\_COV to be  $(T_{\geq 365}) * CLINIC$ . As of now, SPSS takes the value 1 if survival time is more than 365 and 0 otherwise.

```

TIME PROGRAM .
COMPUTE hv2 = (T_ >= 365)*clinic.
COMPUTE hv1 = (T_ < 365)*clinic.
COXREG
  survt /STATUS=status(1)
  /METHOD=ENTER prison dose hv2 hv1
  /CRITERIA=PTN(.05) POUT(.10) ITERATE(20) .

```

### Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)
PRISON	.396	.173	5.226	1	.022	1.486
DOSE	-.026	.007	11.815	1	.001	.975
HV2	-.086	.389	.049	1	.826	.918
HV1	.392	.256	2.349	1	.125	1.480

- It may be interesting to define two heavyside functions and not include CLINIC in the model.
- But, SPSS allows just one tvf using the point and click approach.
- But, I can use syntax.
- Estimated hazard ratio for CLINIC at 100 days  $e(0.392)$  while that ratio at 400 days  $e(-0.086)$ . The same with the result at the previous page.