Survival Analysis in SPSS
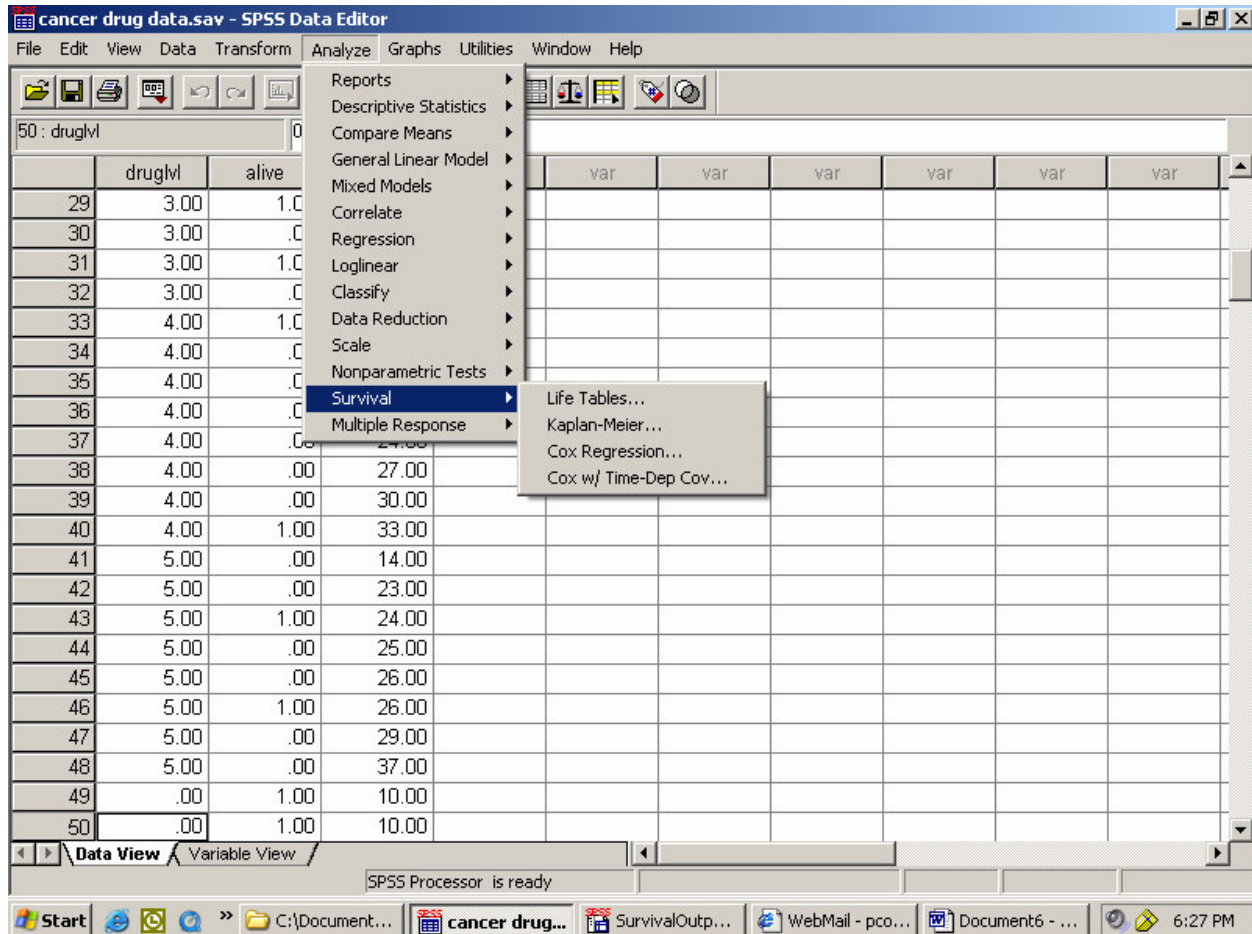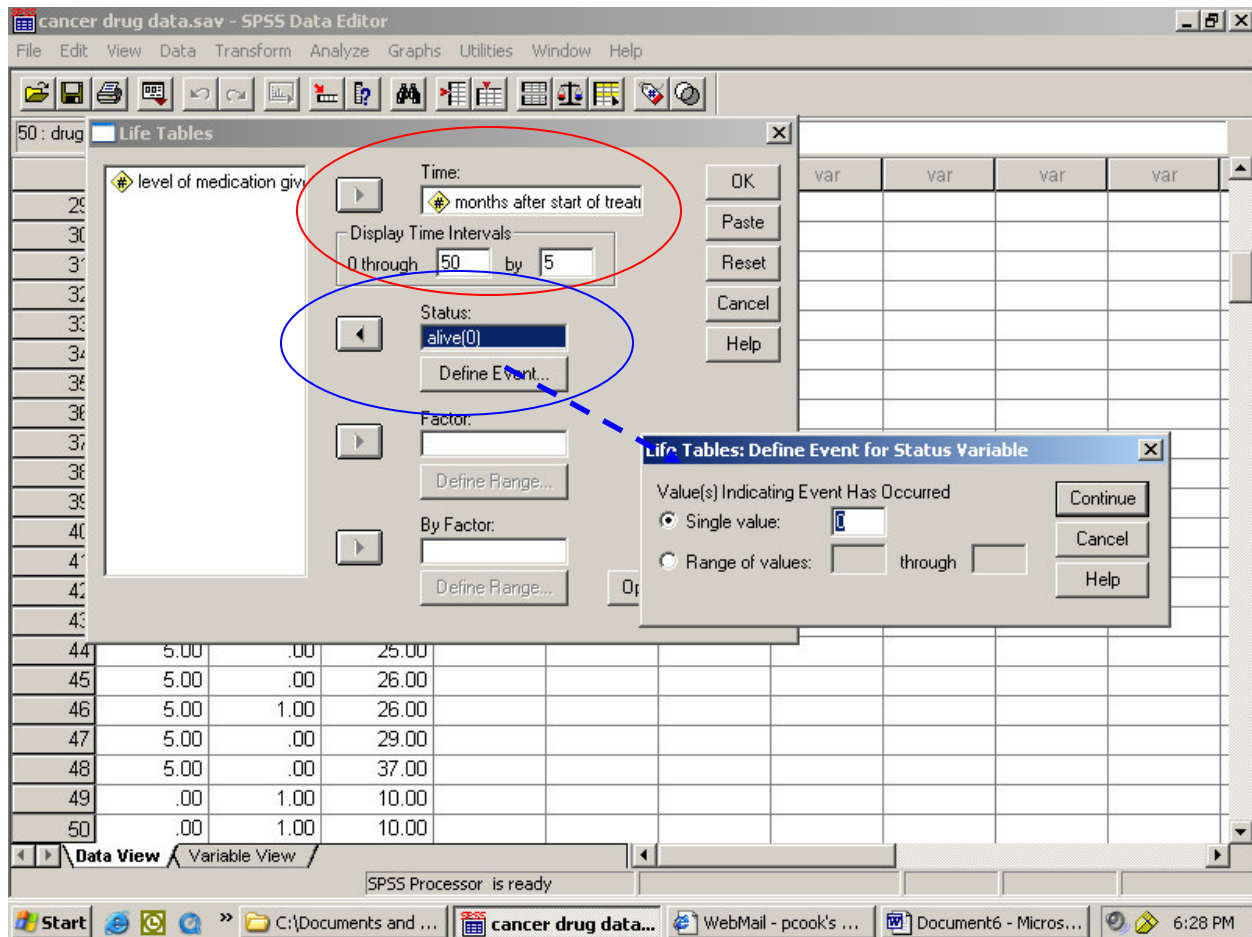Survival analysis is found under its own sub-menu in the "Analyze" menu of SPSS. This example shows survival rates for cancer treatment. Look under "Analyze," then "Survival." You will see four choices in a sub-menu:

We're only going to use two of these four. Start with the "Life Tables" command.

In this dialog box, select two variables:

--The *predictor variable* in survival analysis is always time, and goes in the "Time" box. You also have to use this area under the variable name to select "Intervals," used to break up the time into different amounts. Look at the data to see how high the numbers go: put that highest number in the first box. Then divide the first number by 10 and put the result in the second box (rounding is fine—we're just trying to get *about 10 different intervals* for the data; not too many and not too few).

--The *criterion variable* in survival analysis is the "Status" variable. In this case, the two possible outcomes are alive (1) vs. dead (0). Death is the "event" that we're interested in. Once we have "Alive" entered into the correct box as the status variable, we also need to tell SPSS what code to look for to indicate that *the negative event has occurred*. To do this, click the "Define Event" button, which will show you the sub-dialog pictured above. In this case, the event is indicated by a single value (0, which equals "died"). Type this into the appropriate box, and click "Continue."

Then you can click "OK" in the main dialog box to see the results.

Here are the results:

Survival

This subfile contains: 214 observations

 Life Table
    Survival Variable MONTHS   months after start of treatment

Intrvl Start Time	Number Entrng this Intrvl	Number Wdrawn During Intrvl	Number Exposd to Risk	Number of Termnl Events	Propn Termi- nating	Propn Sur- viving	Cumul Propn Surv at End	Proba- bility Densty	Hazard Rate
.0	214.0	.0	214.0	.0	.0000	1.0000	1.0000	.0000	.0000
5.0	214.0	2.0	213.0	.0	.0000	1.0000	1.0000	.0000	.0000
10.0	212.0	76.0	174.0	24.0	.1379	.8621	.8621	.0276	.0296
15.0	112.0	10.0	107.0	4.0	.0374	.9626	.8298	.0064	.0076
20.0	98.0	12.0	92.0	16.0	.1739	.8261	.6855	.0289	.0381
25.0	70.0	19.0	60.5	22.0	.3636	.6364	.4362	.0499	.0889
30.0	29.0	7.0	25.5	11.0	.4314	.5686	.2481	.0376	.1100
35.0	11.0	1.0	10.5	6.0	.5714	.4286	.1063	.0283	.1600
40.0	4.0	1.0	3.5	.0	.0000	1.0000	.1063	.0000	.0000
45.0	3.0	.0	3.0	.0	.0000	1.0000	.1063	.0000	.0000
50.0+	3.0	.0	3.0	3.0	1.0000	.0000	.0000	**	**

 **   These calculations for the last interval are meaningless.

 The median survival time for these data is 28.72

This table is the main output for this procedure. The left-hand column ("interval start time") shows you the beginning of each step, as the procedure counts forward through time. For instance, the first row shows the results for months 0-5, the next row shows the results for months 5-10, and so forth. Here are the other important columns:

--"Number exposed to risk" is the number of people counted as starting that time interval for purposes of the survival analysis. As you can see, it's a different number from the "number entering this interval"—it has been adjusted for censored data (the cases where we have data, but we don't yet know what the outcome will be).

--"Cumulative Proportion Surviving at End" is the percentage of people (out of 100% at the beginning) who have survived up to the end of the time interval. So, for instance in row 3 (the one marked 10), 86% of those who originally started have made it as far as month 15 (the *end* of that time interval).

--"Hazard Rate" is the % chance of having a terminal event, *for the group of people who were still alive at the start of that particular time interval.* Again in row 3, there's a 2.96% chance of having a terminal event, for people who already made it as far as month 10. If we look for the **largest hazard rate**, we can see that the **time of greatest risk** is between months 35 and 40 (when the hazard rate goes up to 16%).

Finally, at the bottom of this table, we can see the **median survival time**. Recall from the lecture notes that this is *the time at which 50% of those who originally started out have had the terminal event happen*. In this case, the median survival time is 28;72 months, but those people who make it past month 35 are actually *more at risk* later on.

```
          SE of   SE of
 Intrvl   Cumul   Proba-  SE of
 Start    Sur-    bility  Hazard
 Time     viving  Densty   Rate
 -------  ------  ------  ------
      .0  .0000   .0000   .0000
     5.0  .0000   .0000   .0000
    10.0  .0261   .0052   .0060
    15.0  .0297   .0032   .0038
    20.0  .0410   .0066   .0095
    25.0  .0498   .0090   .0185
    30.0  .0513   .0096   .0319
    35.0  .0438   .0096   .0599
    40.0  .0438   .0000   .0000
    45.0  .0438   .0000   .0000
    50.0+ .0000     **      **
```

These are just a few more columns that didn't fit on the page—nothing that we need to see here.

Let's go back to the "life table" command. There's one more thing to try.
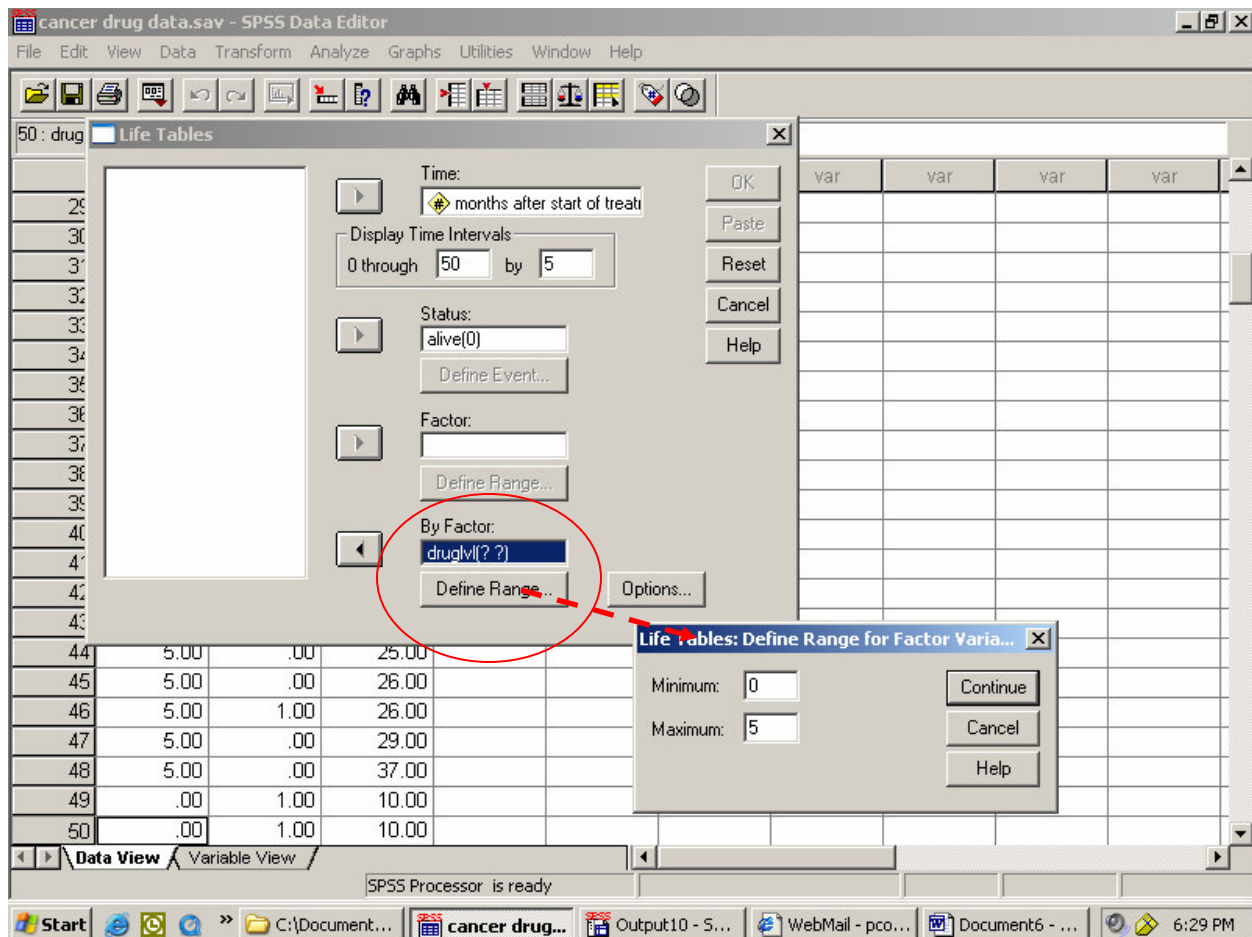
Leave the setup exactly the same as before, except this time, add a *grouping variable* to the procedure. You can use this method to get *different survival times for different subgroups of people*. In this instance, we're interested in whether people who get different drug treatments have different average survival times.

Put "druglvl" into the "by factor" box, and then click "define range" to get the other window shown. We need to tell SPSS what possible values of "druglvl" we want to compare—tell it everthing from druglvl = 0 (the minimum) to druglvl = 5 (the maximum). Then hit "continue" in the small box, and "OK" in the large one to go on.

Here's the revised output. As you can see, you get a different life table for each different value of the factor you entered. Each table has its own pattern of hazard rates and median survival time. If we compare the median survival times (circled below), we can see an effect of the different levels of the drug, in terms of the average number of months that participants in each drug condition survived. When druglvl = 3, we see the longest survival times.

This subfile contains: 214 observations

 Life Table
    Survival Variable MONTHS   months after start of treatment
                  for   DRUGLVL   level of medication given
                                                    =   0

Intrvl Start Time	Number Entrng this Intrvl	Number Wdrawn During Intrvl	Number Exposd to Risk	Number of Termnl Events	Propn Termi- nating	Propn Sur- viving	Cumul Propn Surv at End	Proba- bility Densty	Hazard Rate
.0	48.0	.0	48.0	.0	.0000	1.0000	1.0000	.0000	.0000
5.0	48.0	1.0	47.5	.0	.0000	1.0000	1.0000	.0000	.0000
10.0	47.0	41.0	26.5	4.0	.1509	.8491	.8491	.0302	.0327
15.0	2.0	.0	2.0	1.0	.5000	.5000	.4245	.0849	.1333
20.0	1.0	.0	1.0	.0	.0000	1.0000	.4245	.0000	.0000
25.0	1.0	.0	1.0	.0	.0000	1.0000	.4245	.0000	.0000
30.0	1.0	.0	1.0	.0	.0000	1.0000	.4245	.0000	.0000
35.0	1.0	.0	1.0	.0	.0000	1.0000	.4245	.0000	.0000
40.0	1.0	.0	1.0	.0	.0000	1.0000	.4245	.0000	.0000
45.0	1.0	.0	1.0	.0	.0000	1.0000	.4245	.0000	.0000
50.0+	1.0	.0	1.0	1.0	1.0000	.0000	.0000	**	**

  **   These calculations for the last interval are meaningless.

 The median survival time for these data is   19.11

Life Table
    Survival Variable MONTHS   months after start of treatment
                  for   DRUGLVL   level of medication given
                                                    =   1

Intrvl Start Time	Number Entrng this Intrvl	Number Wdrawn During Intrvl	Number Exposd to Risk	Number of Termnl Events	Propn Termi- nating	Propn Sur- viving	Cumul Propn Surv at End	Proba- bility Densty	Hazard Rate
.0	39.0	.0	39.0	.0	.0000	1.0000	1.0000	.0000	.0000
5.0	39.0	.0	39.0	.0	.0000	1.0000	1.0000	.0000	.0000
10.0	39.0	11.0	33.5	5.0	.1493	.8507	.8507	.0299	.0323
15.0	23.0	4.0	21.0	1.0	.0476	.9524	.8102	.0081	.0098
20.0	18.0	5.0	15.5	4.0	.2581	.7419	.6011	.0418	.0593
25.0	9.0	4.0	7.0	4.0	.5714	.4286	.2576	.0687	.1600
30.0	1.0	.0	1.0	1.0	1.0000	.0000	.0000	.0515	.4000

 The median survival time for these data is   26.47

Life Table
    Survival Variable   MONTHS   months after start of treatment
                  for   DRUGLVL   level of medication given
                                                    =   2

Intrvl Start Time	Number Entrng this Intrvl	Number Wdrawn During Intrvl	Number Exposd to Risk	Number of Termnl Events	Propn Termi- nating	Propn Sur- viving	Cumul Propn Surv at End	Proba- bility Densty	Hazard Rate
.0	40.0	.0	40.0	.0	.0000	1.0000	1.0000	.0000	.0000
5.0	40.0	.0	40.0	.0	.0000	1.0000	1.0000	.0000	.0000
10.0	40.0	9.0	35.5	5.0	.1408	.8592	.8592	.0282	.0303
15.0	26.0	2.0	25.0	1.0	.0400	.9600	.8248	.0069	.0082
20.0	23.0	3.0	21.5	4.0	.1860	.8140	.6713	.0307	.0410
25.0	16.0	7.0	12.5	5.0	.4000	.6000	.4028	.0537	.1000
30.0	4.0	2.0	3.0	1.0	.3333	.6667	.2685	.0269	.0800
35.0	1.0	.0	1.0	1.0	1.0000	.0000	.0000	.0537	.4000

The median survival time for these data is 28.19

Life Table
    Survival Variable   MONTHS   months after start of treatment
                  for   DRUGLVL   level of medication given
                                                    =   3

Intrvl Start Time	Number Entrng this Intrvl	Number Wdrawn During Intrvl	Number Exposd to Risk	Number of Termnl Events	Propn Termi- nating	Propn Sur- viving	Cumul Propn Surv at End	Proba- bility Densty	Hazard Rate
.0	40.0	.0	40.0	.0	.0000	1.0000	1.0000	.0000	.0000
5.0	40.0	1.0	39.5	.0	.0000	1.0000	1.0000	.0000	.0000
10.0	39.0	10.0	34.0	3.0	.0882	.9118	.9118	.0176	.0185
15.0	26.0	.0	26.0	1.0	.0385	.9615	.8767	.0070	.0078
20.0	25.0	2.0	24.0	4.0	.1667	.8333	.7306	.0292	.0364
25.0	19.0	4.0	17.0	4.0	.2353	.7647	.5587	.0344	.0533
30.0	11.0	2.0	10.0	4.0	.4000	.6000	.3352	.0447	.1000
35.0	5.0	1.0	4.5	3.0	.6667	.3333	.1117	.0447	.2000
40.0	1.0	1.0	.5	.0	.0000	1.0000	.1117	.0000	.0000

The median survival time for these data is 31.31

```
Life Table
   Survival Variable  MONTHS    months after start of treatment
                 for  DRUGLVL   level of medication given
                                                =        4


         Number  Number  Number  Number                 Cumul
 Intrvl  Entrng  Wdrawn  Exposd    of    Propn   Propn   Propn   Proba-
 Start    this   During    to    Termnl  Termi-  Sur-    Surv    bility  Hazard
 Time    Intrvl  Intrvl  Risk    Events  nating  viving  at End  Densty  Rate
 ------  ------  ------  ------  ------  ------  ------  ------  ------  ------
    .0    39.0      .0    39.0      .0   .0000  1.0000  1.0000   .0000   .0000
   5.0    39.0      .0    39.0      .0   .0000  1.0000  1.0000   .0000   .0000
  10.0    39.0     5.0    36.5     6.0   .1644   .8356   .8356   .0329   .0358
  15.0    28.0     4.0    26.0      .0   .0000  1.0000   .8356   .0000   .0000
  20.0    24.0     1.0    23.5     3.0   .1277   .8723   .7289   .0213   .0273
  25.0    20.0     3.0    18.5     6.0   .3243   .6757   .4925   .0473   .0774
  30.0    11.0     3.0     9.5     5.0   .5263   .4737   .2333   .0518   .1429
  35.0     3.0      .0     3.0     1.0   .3333   .6667   .1555   .0156   .0800
  40.0     2.0      .0     2.0      .0   .0000  1.0000   .1555   .0000   .0000
  45.0     2.0      .0     2.0      .0   .0000  1.0000   .1555   .0000   .0000
  50.0+    2.0      .0     2.0     2.0  1.0000   .0000   .0000      **      **

    **      These calculations for the last interval are meaningless.

   The median survival time for these data is   29.84


Life Table
   Survival Variable  MONTHS    months after start of treatment
                 for  DRUGLVL   level of medication given
                                                =        5


         Number  Number  Number  Number                 Cumul
 Intrvl  Entrng  Wdrawn  Exposd    of    Propn   Propn   Propn   Proba-
 Start    this   During    to    Termnl  Termi-  Sur-    Surv    bility  Hazard
 Time    Intrvl  Intrvl  Risk    Events  nating  viving  at End  Densty  Rate
 ------  ------  ------  ------  ------  ------  ------  ------  ------  ------
    .0     8.0      .0     8.0      .0   .0000  1.0000  1.0000   .0000   .0000
   5.0     8.0      .0     8.0      .0   .0000  1.0000  1.0000   .0000   .0000
  10.0     8.0      .0     8.0     1.0   .1250   .8750   .8750   .0250   .0267
  15.0     7.0      .0     7.0      .0   .0000  1.0000   .8750   .0000   .0000
  20.0     7.0     1.0     6.5     1.0   .1538   .8462   .7404   .0269   .0333
  25.0     5.0     1.0     4.5     3.0   .6667   .3333   .2468   .0987   .2000
  30.0     1.0      .0     1.0      .0   .0000  1.0000   .2468   .0000   .0000
  35.0     1.0      .0     1.0     1.0  1.0000   .0000   .0000   .0494   .4000

   The median survival time for these data is   27.44
```
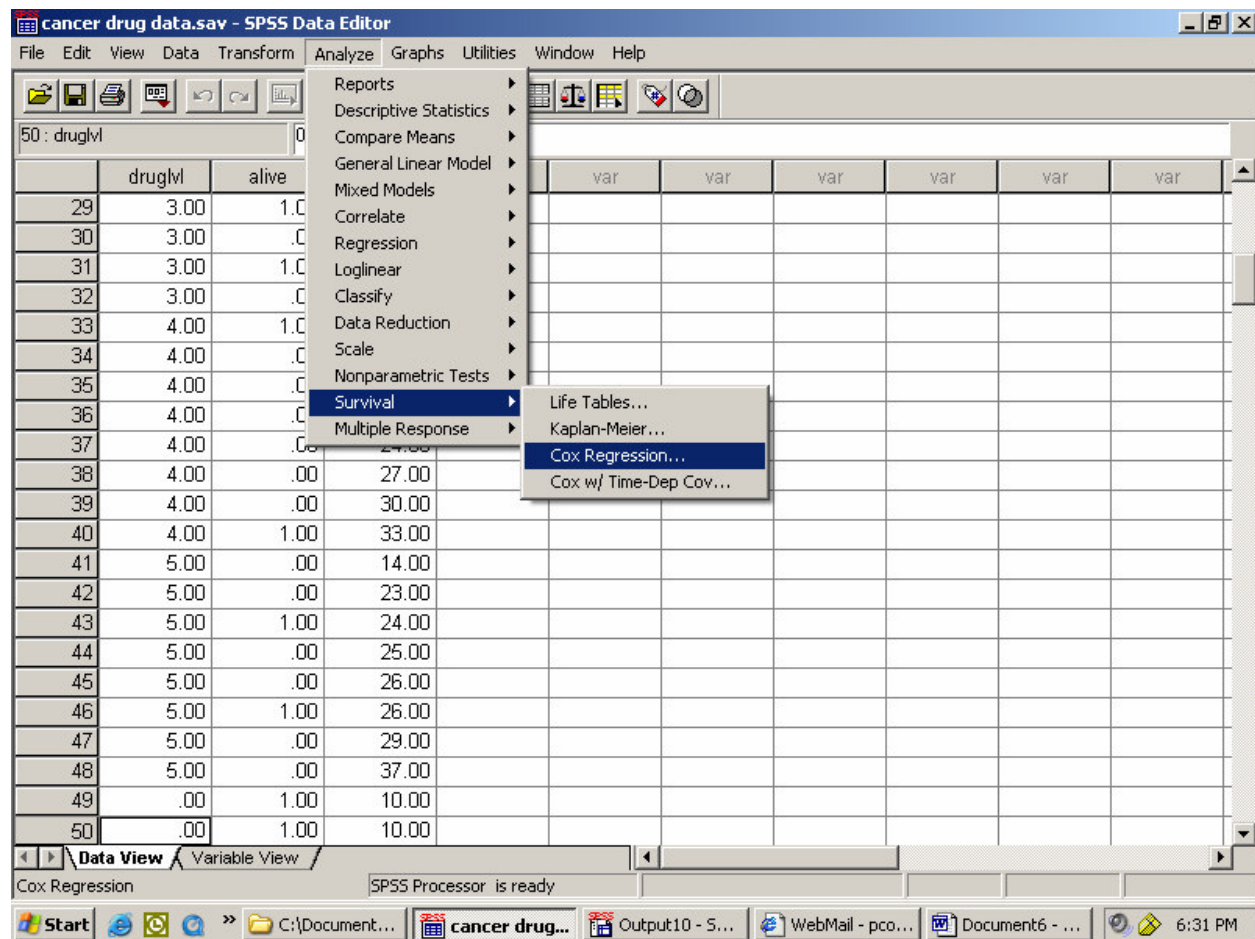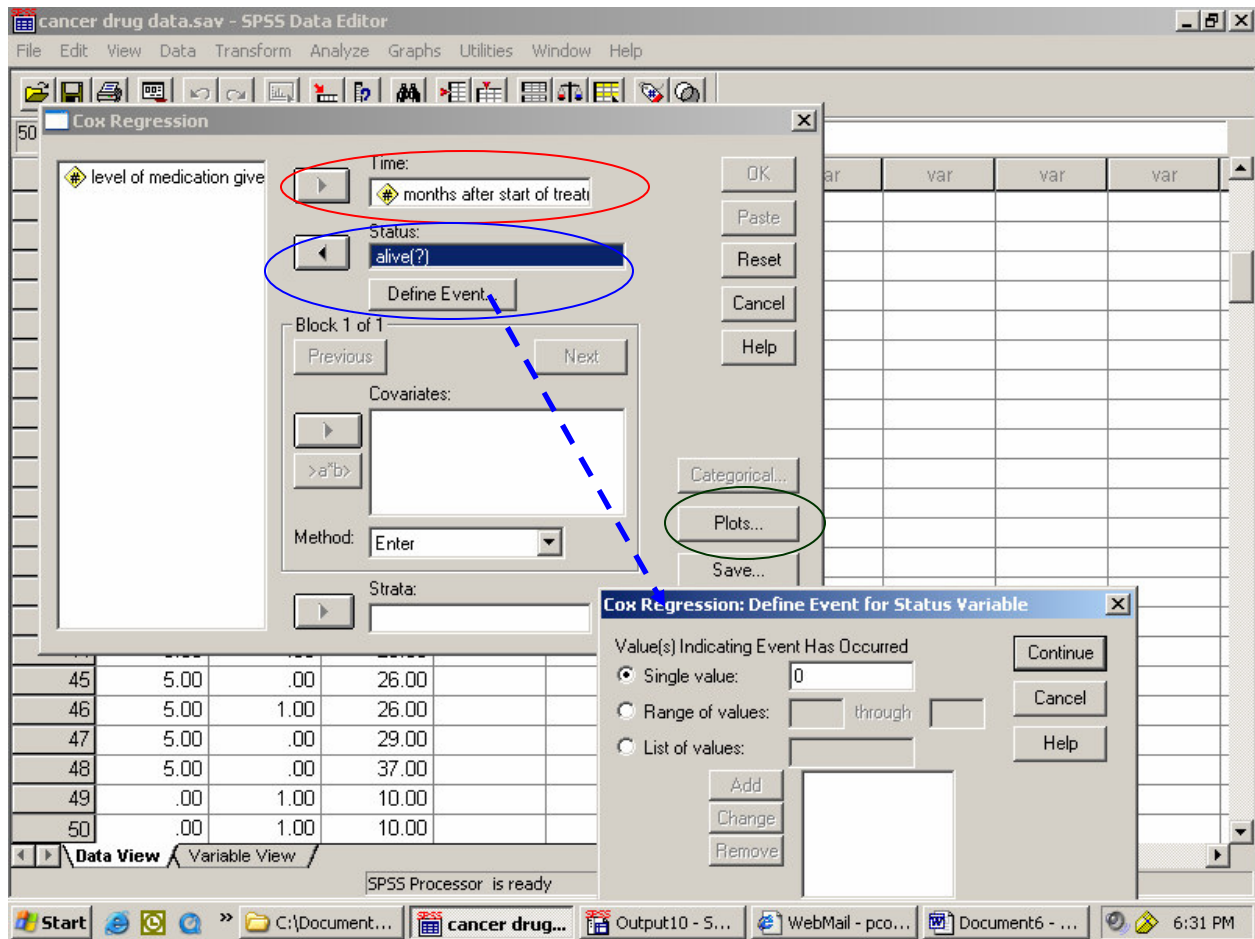
Finally, let's use a different command—Cox Regression—to get a *graph* of the survival curve.

In this dialog box, you *don't have to* put in the time intervals—just put the predictor variable into the "time" box, and SPSS will figure out for itself how to break it out.

You do still have to select the "status" variable, and use the "define event" button to tell SPSS what number indicates that the terminal event has occurred.

Click "Continue" to close the small dialog box. Then click on the "Plots" button to see the next dialog window.
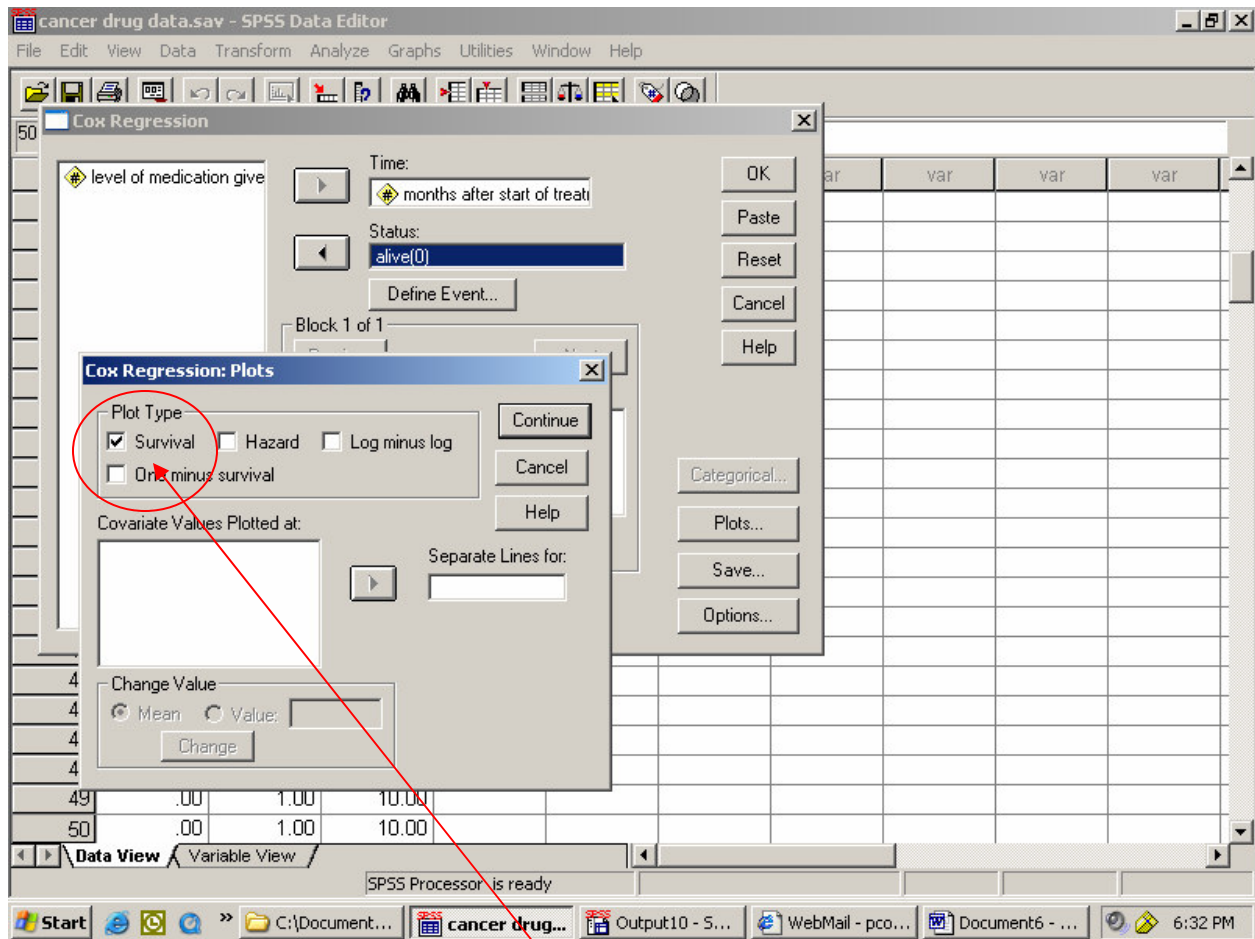
In the "plots" window, just click on the check-box to say that you want a "survival" graph.

Hit "Continue," and then "OK" in the main dialog box to see the graph.

# Cox Regression

Case Processing Summary

		N	Percent
Cases available in analysis	Event(a)	86	40.2%
	Censored	126	58.9%
	Total	212	99.1%
Cases dropped	Cases with missing values	0	.0%
	Cases with negative time	0	.0%
	Censored cases before the earliest event in a stratum	2	.9%
	Total	2	.9%
Total		214	100.0%

a Dependent Variable: months after start of treatment

Omnibus Tests of Model Coefficients

-2 Log Likelihood
698.308

## Survival Function at mean of covariates

months after start of treatment
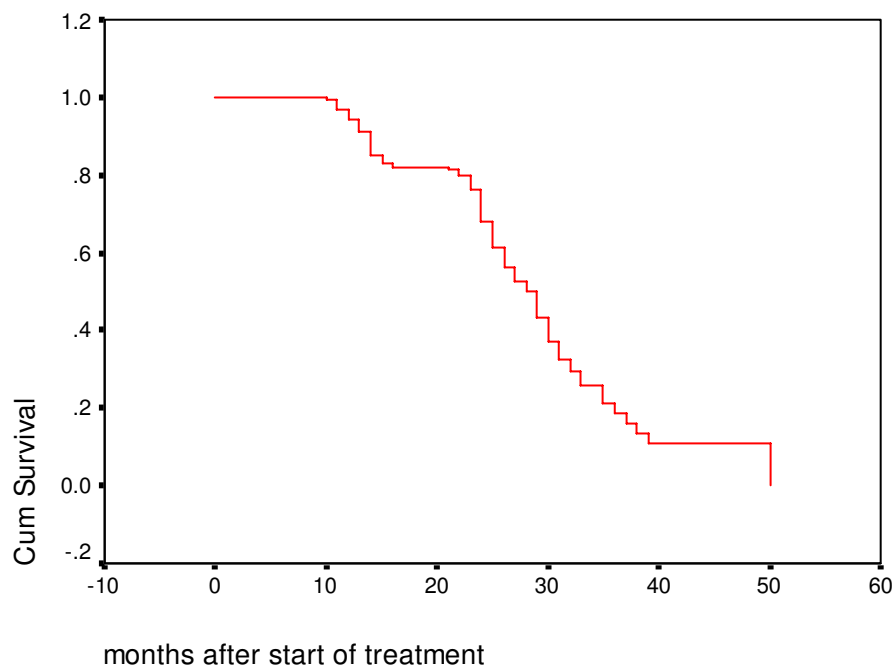
In this output we have a –2 log L statistic, and then the graph.

It's possible to show a different graph for each of several different sub-groups, the same way we got different life tables for different sub-groups. Again, we'll use "druglvl" as the grouping variable.
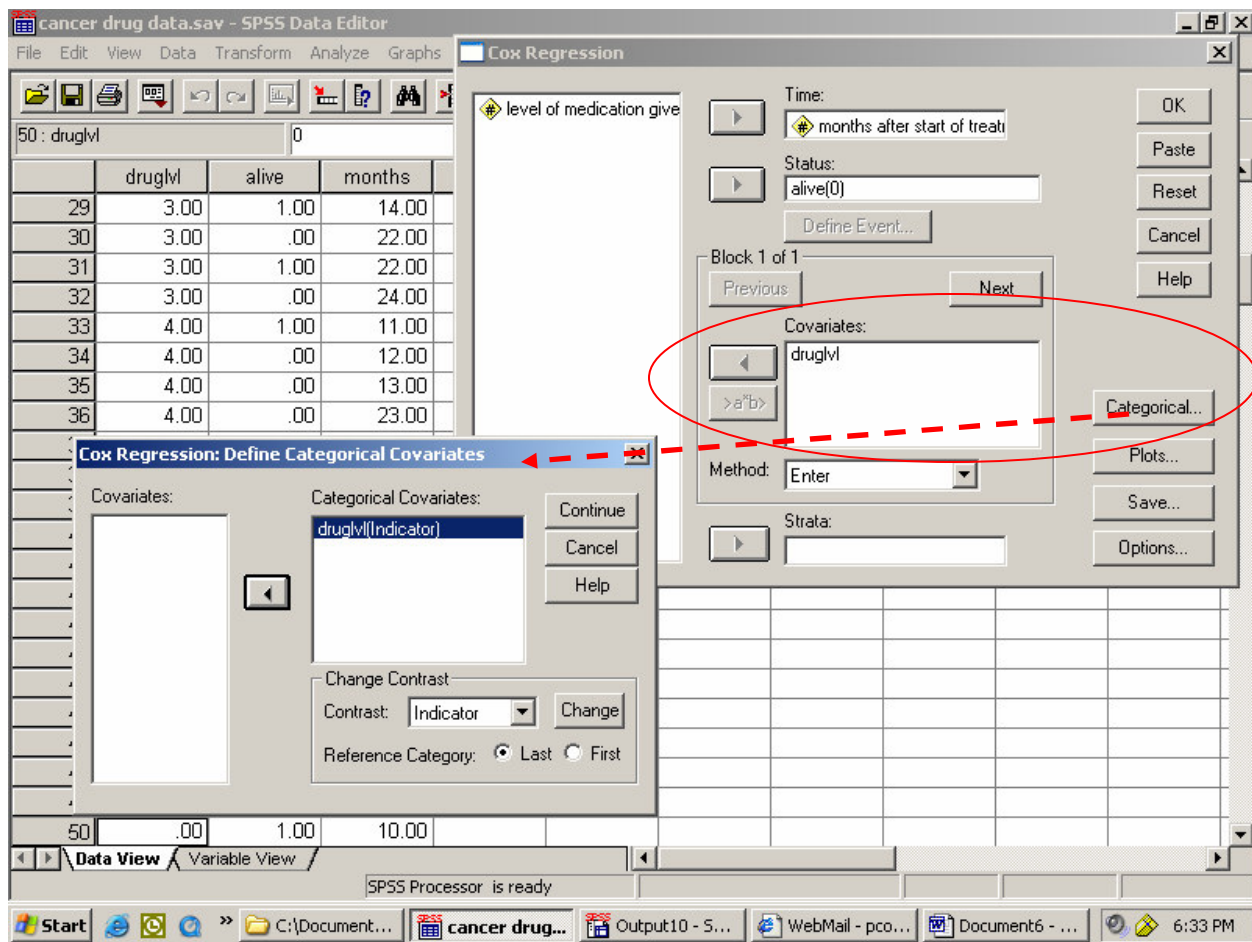

This time, druglvl is called a "covariate" in the dialog box (another one of those inexplicable wording changes in SPSS). It's a grouping variable, but SPSS doesn't immediately recognize that (there would be a "(Cat)" after the variable name if it did). So we need to use the "Categorical" button to tell SPSS that this is a categorical variable.
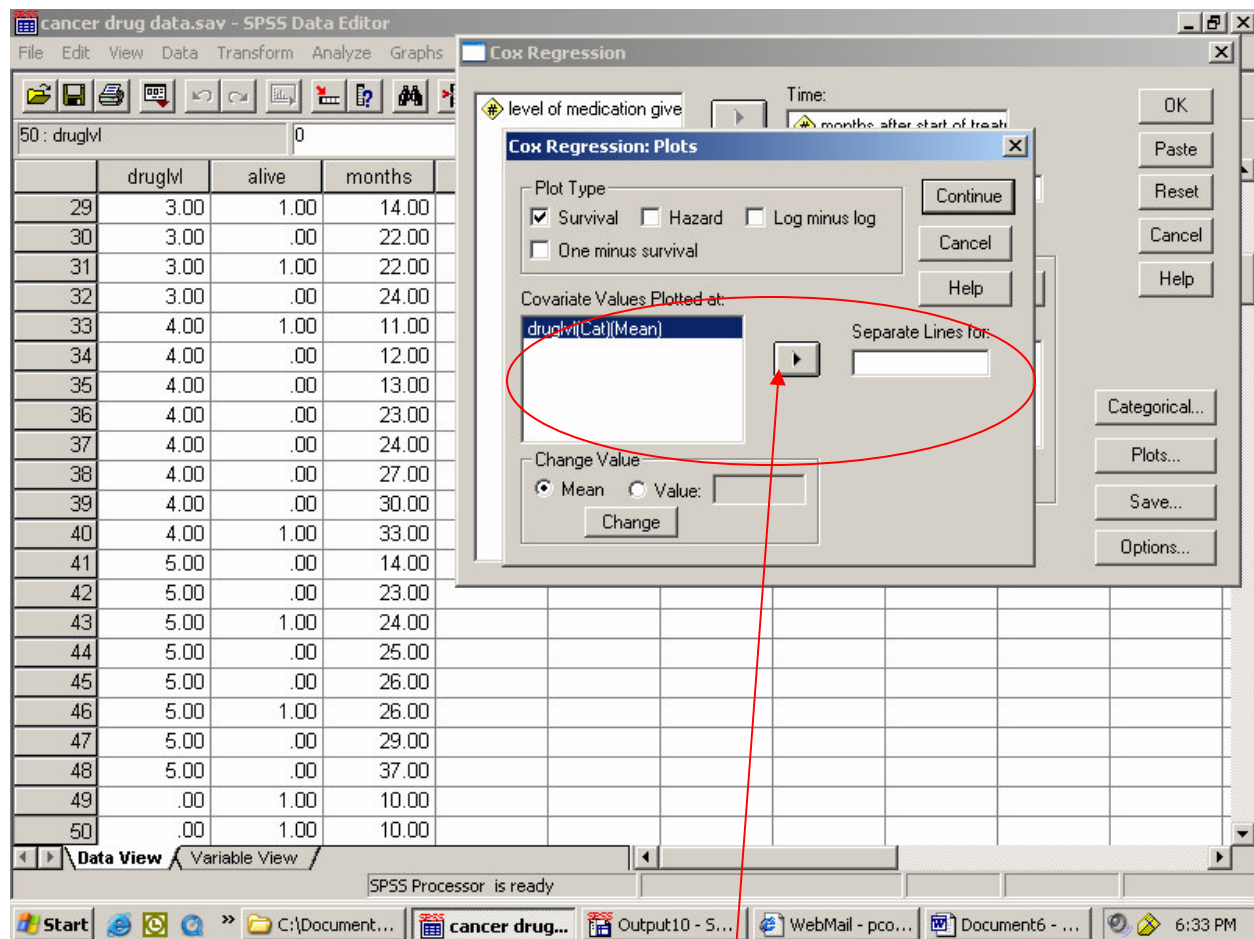
Go back to the "Plots" sub-dialog, and tell SPSS that you want *separate lines* for the different levels of the "druglvl" variable. Do this by using **this arrow** to move "druglvl" from the list on the left to the box on the right.
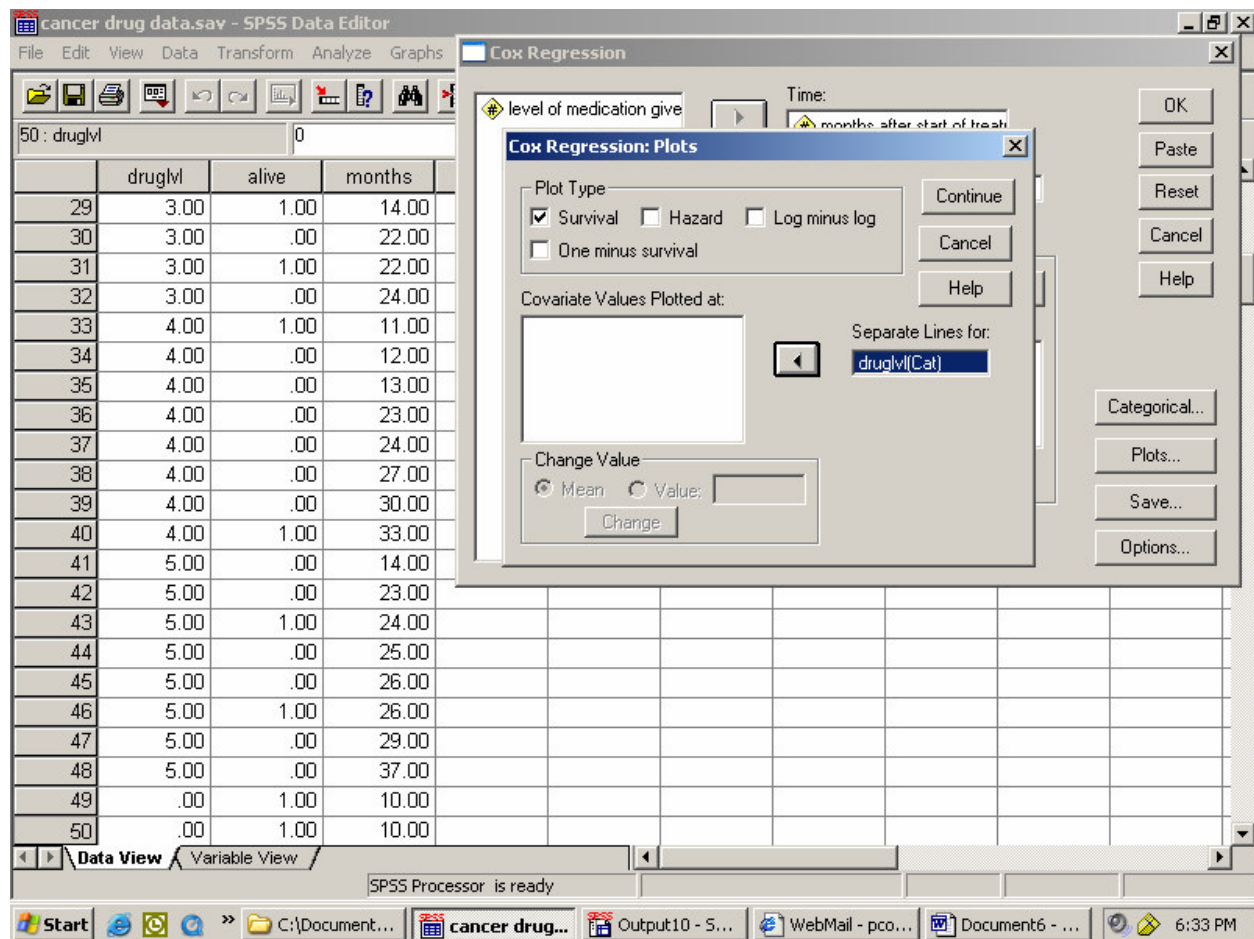
Here it is after "druglvl" has been moved over. Hit "Continue" in the sub-dialog, and then "OK" in the main dialog box, to see the graph.

Here's the final output, showing different-colored lines for the survival pattern seen with each different level of the drug treatment.

Survival Function for patterns 1 - 6

level of medication

- 5.00
- 4.00
- 3.00
- 2.00
- 1.00
- .00

months after start of treatment