"Longitudinal Data – Repeated Measures, Survival (Time to Event) and Cox Proportional Hazards Models"

Melinda K. Higgins, Ph.D.

18 April 2008

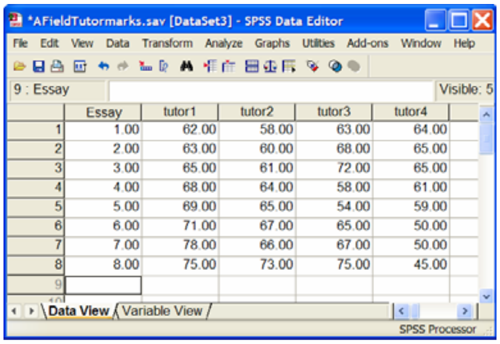
EMORY UNIVERSITY School of Nursing – Research Roundtable *18 April 2008 – M. Higgins*

Outline

- Repeated Measures
  - RM-ANOVA - step-by-step (one-way univariate)
  - "Profile Analysis" approach [Tabachnick & Fidell]
- Time to Event
  - Survival Analysis (Kaplan-Meier)
  - Risk or Hazards Assessment of "covariates" (e.g. Cox Proportional Hazards Regression)

Longitudinal Models: Repeated Measures; Survival Analysis and Cox Regression

Repeated Measures

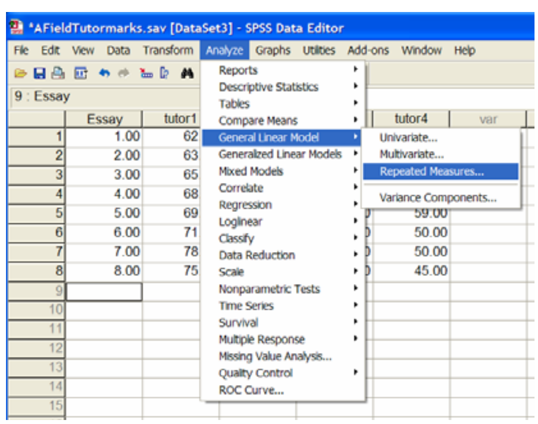


8 Essays ("subjects")
4 Tutors ("repeated for essays")

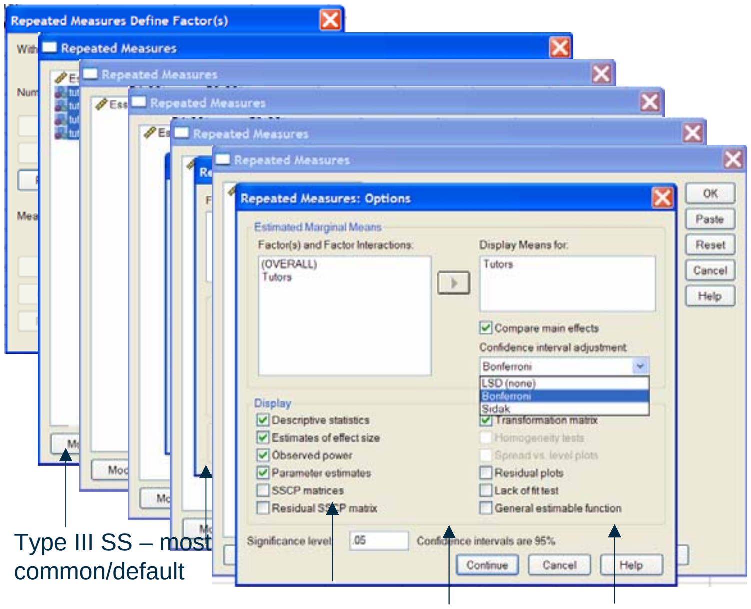
One-way Repeated Measures:
a) evaluate within-subject variance
b) evaluate difference across repeated measure
c) evaluate contrasts between levels of repeated measures
d) Post hoc tests and options (main effects tests)

EMORY UNIVERSITY School of Nursing – Research Roundtable *18 April 2008 – M. Higgins*

RM-ANOVA – set-up in SPSS



Longitudinal Models: Repeated Measures; Survival Analysis and Cox Regression



Type III SS – most common/default

Contrasts – Options

Deviation. Compares the mean of each level (except a reference category) to the mean of all of the levels (grand mean). The levels of the factor can be in any order.

Simple. Compares the mean of each level to the mean of a specified level. This type of contrast is useful when there is a control group. You can choose the first or last category as the reference.

Difference. Compares the mean of each level (except the first) to the mean of previous levels. (Sometimes called reverse Helmert contrasts.)

Order is important

Helmert. Compares the mean of each level of the factor (except the last) to the mean of subsequent levels.

Repeated. Compares the mean of each level (except the last) to the mean of the subsequent level.

Polynomial. Compares the linear effect, quadratic effect, cubic effect, and so on. The first degree of freedom contains the linear effect across all categories; the second degree of freedom, the quadratic effect; and so on. These contrasts are often used to estimate polynomial trends.

First example, we'll use Repeated since there is no good "reference" level or category.

Evaluate Sphericity

Sphericity is the equality of variances of the differences b/t repeated measures

Mauchly's Test of Sphericity[b]

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[a] Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Tutors	.131	11.628	5	.043	.558	.712	.333

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b. Design: Intercept
Within Subjects Design: Tutors

Use G-G, maybe …

If p-value < 0.05, be WARY of F-tests. But a few corrections exist - review Greenhouse-Geisser and Huynh-Feldt

Greenhouse-Geisser (G-G) – closer to 1, more homogeneous the differences (closer to sphericity) – compare to “lower-bound” [want closer to 1 than lower-bound] – “tends to be too conservative”

Huynh-Feldt (H-F) – if G-G>0.75 use H-F correction; if G-G<0.75 use G-G correction

(or take average of p-val between G-G and H-F)

Field, Andy; Discovering Statistics Using SPSS, 2nd ed.; SAGE Publications: London (2005).

Main Repeated Measures Effects (Tests of Within-Subjects)

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Tutors	Sphericity Assumed	554.125	3	184.708	3.700	.028	.346	11.100	.722
	Greenhouse-Geisser	554.125	1.673	331.245	3.700	.063	.346	6.189	.522
	Huynh-Feldt	554.125	2.137	259.329	3.700	.047	.346	7.906	.603
	Lower-bound	554.125	1.000	554.125	3.700	.096	.346	3.700	.383
Error(Tutors)	Sphericity Assumed	1048.375	21	49.923					
	Greenhouse-Geisser	1048.375	11.710	89.528					
	Huynh-Feldt	1048.375	14.957	70.091					
	Lower-bound	1048.375	7.000	149.768					

a. Computed using alpha = .05

Sphericity was violated, and the G-G correction was indicated, although it can “tend to be too conservative.

G-G indicates we should not reject Ho: no difference across within-subjects (between the tutors).

But the H-F is significant and indicates we should reject Ho.

A. Field suggest taking the average to get a p-val = (.063+.047)/2 = .055, do not reject Ho.

Also check Multivariate Tests

Multivariate Tests[c]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Tutors	Pillai's Trace	.741	4.760[b]	3.000	5.000	.063	.741	14.280	.564
	Wilks' Lambda	.259	4.760[b]	3.000	5.000	.063	.741	14.280	.564
	Hotelling's Trace	2.856	4.760[b]	3.000	5.000	.063	.741	14.280	.564
	Roy's Largest Root	2.856	4.760[b]	3.000	5.000	.063	.741	14.280	.564

a. Computed using alpha = .05

b. Exact statistic

c.
Design: Intercept
Within Subjects Design: Tutors

Now the p-value for all multivariate tests is 0.063, which definitely indicates that we should not reject Ho = "all tutors score the same."

Ideally, the analysis stops here since we have concluded there is no significant difference across all of the tutors.

However, for completeness we can look at the contrasts results

Contrasts

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	Tutors	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Tutors	Level 1 vs. Level 2	171.125	1	171.125	18.184	.004	.722	18.184	.953
	Level 2 vs. Level 3	8.000	1	8.000	.152	.708	.021	.152	.063
	Level 3 vs. Level 4	496.125	1	496.125	3.436	.106	.329	3.436	.360
Error(Tutors)	Level 1 vs. Level 2	65.875	7	9.411					
	Level 2 vs. Level 3	368.000	7	52.571					
	Level 3 vs. Level 4	1010.875	7	144.411					

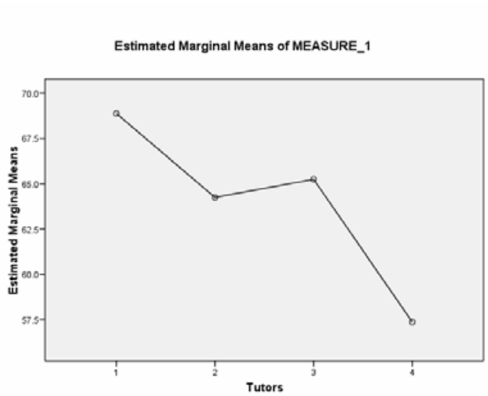
a. Computed using alpha = .05

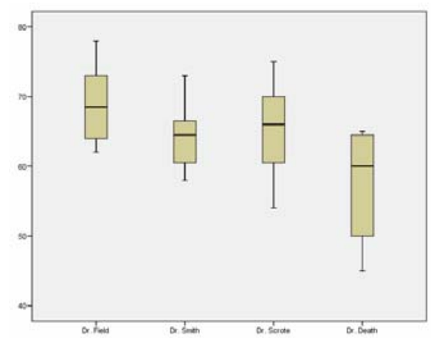
Only the difference between "Level 1 [Dr. Field]" and "Level 2 [Dr. Smith]" was significant. [Although this result would be ignored since the main effect test for tutor was non-significant and sphericity was violated.]

Look at the "Means Plot" as well as boxplots ("summaries of separate variables" for all of the tutors)

EMORY UNIVERSITY **School of Nursing – Research Roundtable** *18 April 2008 – M. Higgins*

Means Plot and Boxplots





Tutors 1=Dr. Field; 2=Dr. Smith; 3=Dr. Scrote 4=Dr. Death

Longitudinal Models: Repeated Measures; Survival Analysis and Cox Regression

EMORY UNIVERSITY **School of Nursing – Research Roundtable** *18 April 2008 – M. Higgins*

Mixed Design ANOVA (Between-groups and Repeated Measures)

Wechsler Intelligence Scale for Children (WISC) – 11 subtests

Primary question: Do profiles of learning-disabled children on the WISC subtests differ if the children are grouped by the age of their playmates?

Between

177 subjects
3 Age-groups (age of playmate) [younger, older, same age]
11 Subtests ("repeated" measured for every subject)
[Information; Comprehension; Arithmetic; Similarities; Vocabulary; Digit Span; Picture Completion; Picture Arrangement; Block Design; Object Assembly; Coding]

Within – multiple DVs

Mixed Design Repeated Measures:
- **a) evaluate between-group effects**
- **b) evaluate within-subject variance**
- **c) evaluate difference across repeated measure ("profiles")**
- **d) evaluate interactions between and within factors**
- **e) optional comparisons**

*Tabachnick, B. G.; Fidell, L. S. **Using Multivariate Statistics (5th Ed.).** Pearson Edu. Inc.: Boston, MA (2007).*

Longitudinal Models: Repeated Measures; Survival Analysis and Cox Regression

"Profile" Analysis vs MANOVA

- In "Profile" Analysis, the "repeated measures" (DVs) are all evaluated together.
  - In MANOVA – the DVs are treated/evaluated directly.
- In "Profile" Analysis, interactions between "repeated measures" and "between-group" factor(s) can be evaluated
  - In MANOVA, since DVs are evaluated directly, no interactions can be evaluated.

"Profile" Analysis Goals/Aims

Major Question: Whether or not groups have different profiles on (over) a set of measures [note the scales of the "repeated measures" assumed to be the same]

1. Parallelism – do different groups have parallel profiles [i.e. is there interaction between the "Between-Group/Subjects" factor and "Within/Repeated Measures."]
2. Levels Test – regardless of parallelism, does one group on average differ significantly from the others [Between-Subjects Main Effects]
3. Flatness – independent of groups, do all DVs/repeated measures elicit the same average response? [Within-Subjects Effects]

"Profile Analysis" of the WISC Data

Within-Subjects Factors

Measure: MEASURE_1

wisc_subtests	Dependent Variable
1	info
2	comp
3	arith
4	simil
5	vocab
6	digit
7	pictcomp
8	parang
9	block
10	object
11	coding

Between-Subjects Factors

		Value Label	N
agemate Preferred age of playmates	1.00	children preference for playmates younger than them	45
	2.00	children preference for playmates older than them	54
	3.00	children preference for playmates same age as them	65

Mauchly's Test of Sphericity[b]

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[a] Greenhouse-Geisser	Huynh-Feldt	Lower-bound
wisc_subtests	.281	199.932	54	.000	.771	.824	.100

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b. Design: Intercept+agemate
Within Subjects Design: wisc_subtests

Sphericity violated and G-G>0.75 – use H-F

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	15889.921	1	15889.921	5724.116	.000
agemate	4.510	2	2.255	.812	.446
Error	446.930	161	2.776		

Levels Test – differences between agemate groups is not significant

Tests of Within-Subjects Effects

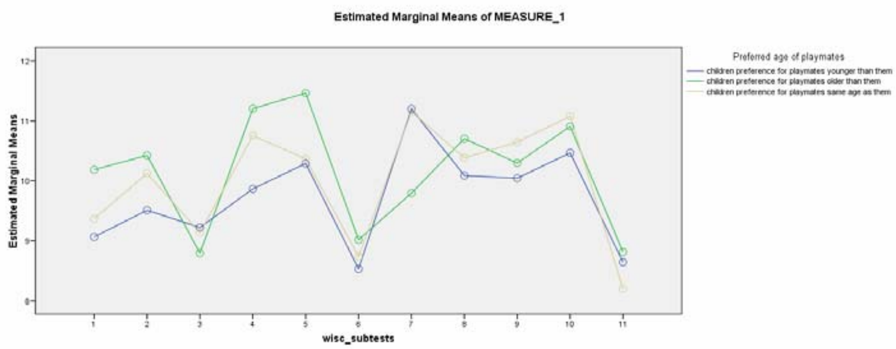
Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
wisc_subtests	Sphericity Assumed	1126.115	10	112.612	19.271	.000
	Greenhouse-Geisser	1126.115	7.710	146.056	19.271	.000
	Huynh-Feldt	1126.115	8.236	136.734	19.271	.000
	Lower-bound	1126.115	1.000	1126.115	19.271	.000
wisc_subtests * agemate	Sphericity Assumed	218.298	20	10.915	1.868	.011
	Greenhouse-Geisser	218.298	15.420	14.157	1.868	.021
	Huynh-Feldt	218.298	16.472	13.253	1.868	.018
	Lower-bound	218.298	2.000	109.149	1.868	.158
Error(wisc_subtests)	Sphericity Assumed	9408.132	1610	5.844		
	Greenhouse-Geisser	9408.132	1241.338	7.579		
	Huynh-Feldt	9408.132	1325.968	7.095		
	Lower-bound	9408.132	161.000	58.436		

Flatness Test – differences within (across) wisc_subtests (repeated measures) is significant

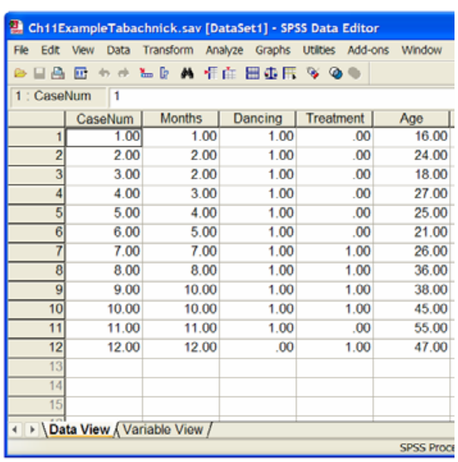
Parallel Test – there are significant interactions between wisc_subtests and agegroups [i.e. profiles are NOT parallel]

Means Plots – Compare Profiles



NOTE: If the Between-Subjects Effects are significant (Levels Test) – you can use these "profiles" as inputs for group classification via Discriminant Analysis

Time to Event – Survival Analysis



	CaseNum	Months	Dancing	Treatment	Age
1	1.00	1.00	1.00	.00	16.00
2	2.00	2.00	1.00	.00	24.00
3	3.00	2.00	1.00	.00	18.00
4	4.00	3.00	1.00	.00	27.00
5	5.00	4.00	1.00	.00	25.00
6	6.00	5.00	1.00	.00	21.00
7	7.00	7.00	1.00	1.00	26.00
8	8.00	8.00	1.00	1.00	36.00
9	9.00	10.00	1.00	1.00	38.00
10	10.00	10.00	1.00	1.00	45.00
11	11.00	11.00	1.00	.00	55.00
12	12.00	12.00	.00	1.00	47.00
13					
14					
15					

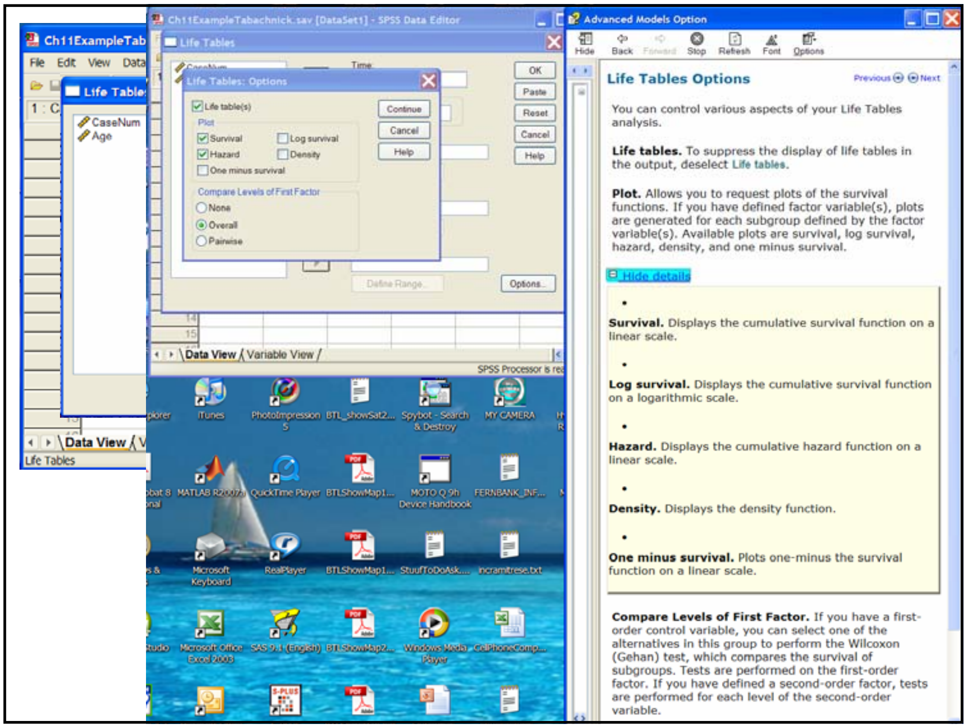
12 Subjects (very small sample!!)

"Event" – stopped dancing (=1)

2 Treatment Levels (0=control; 1=tmt)

Age when subject began dancing is a "covariate"

1. Look at Survival Curves
2. Kaplan-Meier (product-limit method) – group differences in survival times and/or proportions surviving at specific times
3. Cox Regression – prediction survival times from covariates




```
SURVIVAL
  TABLE=Months  BY Treatment(0 1)
  /INTERVAL=THRU 12 BY 1
  /STATUS=Dancing(1)
  /PRINT=TABLE
  /PLOTS ( SURVIVAL  HAZARD )=Months  BY Treatment
  /COMPARE=Months  BY Treatment  .
```

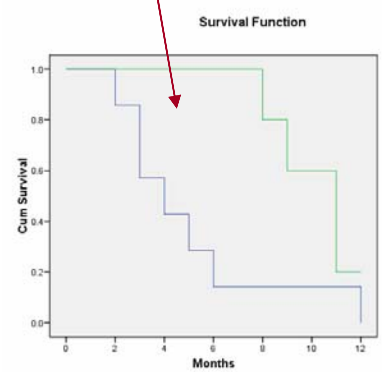
Life Table

First-order Controls		Interval Start Time	Number Entering Interval	Number Withdrawing during Interval	Number Exposed to Risk	Number of Terminal Events	Proportion Terminating	Proportion Surviving	Cumulative Proportion Surviving at End of Interval	Std. Error of Cumulative Proportion Surviving at End of Interval	Probability Density	Std. Error of Probability Density	Hazard Rate	Std. Error of Hazard Rate
Treatment	.00 Control	.000	7	0	7.000	0	.00	1.00	1.00	.00	.000	.000	.00	.00
		1.000	7	0	7.000	1	.14	.86	.86	.13	.143	.132	.15	.15
		2.000	6	0	6.000	2	.33	.67	.57	.19	.286	.171	.40	.28
		3.000	4	0	4.000	1	.25	.75	.43	.19	.143	.132	.29	.28
		4.000	3	0	3.000	1	.33	.67	.29	.17	.143	.132	.40	.39
		5.000	2	0	2.000	1	.50	.50	.14	.13	.143	.132	.67	.63
		6.000	1	0	1.000	0	.00	1.00	.14	.13	.000	.000	.00	.00
		7.000	1	0	1.000	0	.00	1.00	.14	.13	.000	.000	.00	.00
		8.000	1	0	1.000	0	.00	1.00	.14	.13	.000	.000	.00	.00
		9.000	1	0	1.000	0	.00	1.00	.14	.13	.000	.000	.00	.00
		10.000	1	0	1.000	0	.00	1.00	.14	.13	.000	.000	.00	.00
		11.000	1	0	1.000	1	1.00	.00	.00	.00	.143	.132	2.00	.00
	1.00 Treatment	.000	5	0	5.000	0	.00	1.00	1.00	.00	.000	.000	.00	.00
		1.000	5	0	5.000	0	.00	1.00	1.00	.00	.000	.000	.00	.00
		2.000	5	0	5.000	0	.00	1.00	1.00	.00	.000	.000	.00	.00
		3.000	5	0	5.000	0	.00	1.00	1.00	.00	.000	.000	.00	.00
		4.000	5	0	5.000	0	.00	1.00	1.00	.00	.000	.000	.00	.00
		5.000	5	0	5.000	0	.00	1.00	1.00	.00	.000	.000	.00	.00
		6.000	5	0	5.000	0	.00	1.00	1.00	.00	.000	.000	.00	.00
		7.000	5	0	5.000	1	.20	.80	.80	.18	.200	.179	.22	.22
		8.000	4	0	4.000	1	.25	.75	.60	.22	.200	.179	.29	.28
		9.000	3	0	3.000	0	.00	1.00	.60	.22	.000	.000	.00	.00
		10.000	3	0	3.000	2	.67	.33	.20	.18	.400	.219	1.00	.61
		11.000	1	0	1.000	0	.00	1.00	.20	.18	.000	.000	.00	.00

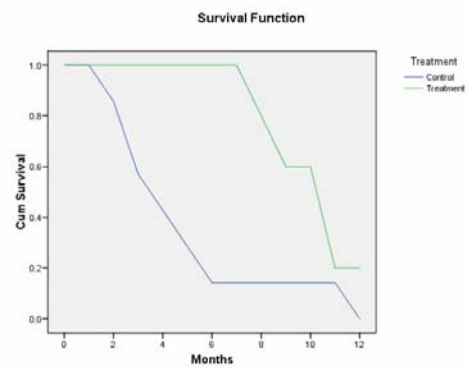
Median Survival Time

First-order Controls		Med Time
Treatment	.00 Control	3.5000
	1.00 Treatment	10.2500

Median Survival Time Differ quite a bit



Line type = step



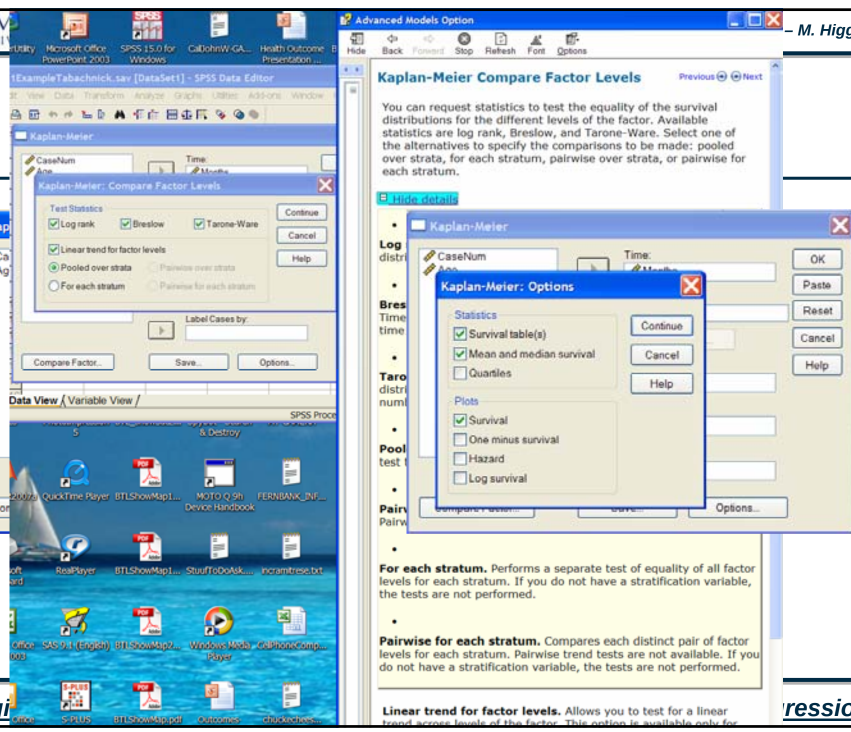
Line type = straight

Overall Comparisons[a]

Wilcoxon (Gehan) Statistic	df	Sig.
4.840	1	.028

a. Comparisons are exact.

Group differences are significant



Case Processing Summary

Treatment	Total N	N of Events	Censored N	Censored Percent
.00 Control	7	7	0	.0%
1.00 Treatment	5	4	1	20.0%
Overall	12	11	1	8.3%

Survival Table

Treatment		Time	Status	Cumulative Proportion Surviving at the Time Estimate	Cumulative Proportion Surviving at the Time Std. Error	N of Cumulative Events	N of Remaining Cases
.00 Control	1	1.000	1.00 Dropped Out	.857	.132	1	6
	2	2.000	1.00 Dropped Out	.	.	2	5
	3	2.000	1.00 Dropped Out	.571	.187	3	4
	4	3.000	1.00 Dropped Out	.429	.187	4	3
	5	4.000	1.00 Dropped Out	.286	.171	5	2
	6	5.000	1.00 Dropped Out	.143	.132	6	1
	7	11.000	1.00 Dropped Out	.000	.000	7	0
1.00 Treatment	1	7.000	1.00 Dropped Out	.800	.179	1	4
	2	8.000	1.00 Dropped Out	.600	.219	2	3
	3	10.000	1.00 Dropped Out	.	.	3	2
	4	10.000	1.00 Dropped Out	.200	.179	4	1
	5	12.000	.00 Still Dancing	.	.	4	0

Means and Medians for Survival Time

	Mean[a]				Median			
			95% Confidence Interval				95% Confidence Interval	
Treatment	Estimate	Std. Error	Lower Bound	Upper Bound	Estimate	Std. Error	Lower Bound	Upper Bound
.00 Control	4.000	1.272	1.506	6.494	3.000	1.309	.434	5.566
1.00 Treatment	9.400	.780	7.872	10.928	10.000	.894	8.247	11.753
Overall	6.250	1.081	4.131	8.369	5.000	2.598	.000	10.092

a. Estimation is limited to the largest survival time if it is censored.

Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	3.747	1	.053
Breslow (Generalized Wilcoxon)	4.926	1	.026
Tarone-Ware	4.522	1	.033

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

Cox Regression

- **Cox Proportional-Hazards Regression Model – models event rates as a log-linear function of predictors (called "covariates").**
- **The regression coefficients give the relative effect (or risk) of each "covariate" on the survivor function.**
- **NOTE: "Covariates" are both the usual covariates, but also include independent predictors and group variables (treatment)**

```
COXREG
  Months  /STATUS=Dancing(1)
  /PATTERN BY Treatment
  /CONTRAST (Treatment)=Indicator
  /METHOD=ENTER Age  /METHOD=ENTER Treatment
  /PLOT SURVIVAL
  /PRINT=CI(95) BASELINE
  /CRITERIA=PIN(.05) POUT(.10) ITERATE(20) .
```

Case Processing Summary

		N	Percent
Cases available in analysis	Event[a]	11	91.7%
	Censored	1	8.3%
	Total	12	100.0%
Cases dropped	Cases with missing values	0	.0%
	Cases with negative time	0	.0%
	Censored cases before the earliest event in a stratum	0	.0%
	Total	0	.0%
Total		12	100.0%

a. Dependent Variable: Months

Block 1: Method = Enter

Omnibus Tests of Model Coefficients[a,b]

-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.
25.395	11.185	1	.001	15.345	1	.000	15.345	1	.000

a. Beginning Block Number 0, initial Log Likelihood function: -2 Log likelihood: 40.740

b. Beginning Block Number 1. Method = Enter

Exp(B) = "odds ratio/hazard ratio"

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
Age	-.199	.072	7.640	1	.006	.819	.711	.944

Age is a significant covariate (via both Wald and -2 Log Likelihood)

Block 2: Method = Enter

Omnibus Tests of Model Coefficients[a,b]

-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.
21.417	14.706	2	.001	3.978	1	.046	3.978	1	.046

a. Beginning Block Number 0, initial Log Likelihood function: -2 Log likelihood: 40.740

b. Beginning Block Number 2. Method = Enter

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
Age	-.230	.089	6.605	1	.010	.795	.667	.947
Treatment	2.542	1.546	2.701	1	.100	12.699	.613	263.064

Age still a significant covariate (Wald)
But treatment is not (according to Wald which is better for smaller sample sizes)

VIII. Statistical Resources and Contact Info

SON S:\Shared\Statistics_MKHiggins

- Shared resource for all of SON – faculty and students
- Will continually update with tip sheets (for SPSS, SAS, and other software), lectures (PPTs and handouts), datasets, other resources and references
- Statistics At Nursing Website: [moving to main website] S:\Shared\Statistics_MKHiggins\website2\index.htm
- And Blackboard Site (in development) for "Organization: Statistics at School of Nursing"

Contact

Dr. Melinda Higgins

Melinda.higgins@emory.edu

Office: 404-727-5180 / Mobile: 404-434-1785