Cox Regression Analysis

- Introducing Cox Regression
- Performing Cox Regression Analysis

Introducing Cox Regression

Cox regression (Cox proportional-hazards model) tests the effects of several factors (predictors) on survival time. Predictors that lower the probability of survival at a given time are called risk factors; predictors that increase the probability of survival at a given time are called protective factors. The Cox proportional-hazards model are similar to a multiple logistic regression that considers time-to-event rather than simply whether an event occurred or not.

In this tutorial, we will use Cox Regression to test the effects of tumor gene expression on survival time while accounting for tumor size.

Performing Cox Regression Analysis

To begin, you should have the Survival Tutorial data set open in Partek Genomics Suite as shown.

- · Select Stat from the main toolbar
- Select Survival Analysis then Cox Regression from the Stat menu (Figure 1)

Partek Genomics Suite - 1 (Survival_Tutorial)								-		×
File Edit Transform View Stat Filter Tools V	Vindow Custom Help									
Descriptive	•					Workflow	ws Choose			~
Analysis X Correlate	•									
Parametric Tests Nonparametric Tests ANOVA	े 🚺 🔍 🍸									
1 (Survival_Tutorial) Remove Batch Effect.	ection 11.833									^
Survival Analysis Fisher Exact	Cox Regression Kaplan-Meier	3. p53 status	4. ER status	5. PgR status	6. age at diagnosis	7. tumor size (mm)	8. Lymph node status	9. 1007_s_at	1 0. 1	
	33 censored	wt	ER+	PgR-	68	9	LN-	11.3908	6.	
Multiple Test Correctio	on 33 censored	mutant	ER-	PgR-	40	12	LN-	11.3651	7.	
Power Analysis	33 censored	mutant	ER+	PgR+	51	26	LN-	11.5778	7.	
4.	3.583 censored	mutant	ER+	PgR+	80	24	LN?	11.2411	6.	
5.	11.75 censored	wt	ER+	PgR+	46	13	LN-	11.3729	6.	
6.	11.333 censored	wt	ER+	PgR+	70	50	LN-	11.8454	6.	
7.	11.667 censored	mutant	ER+	PgR+	74	20	LN-	11.5412	7.	
8.	5.5 censored	wt	ER+	PgR+	38	32	LN-	11.5079	6.	
9.	11.167 censored	wt	ER+	PgR+	67	12	LN+	11.4658	6.	
10.	7.417 censored	wt	ER-	PgR-	69	18	LN-	11.1284	6.	
11.	11.583 censored	wt	ER+	PgR+	34	16	LN-	11.6635	6.	
12.	11.167 censored	wt	ER+	PgR+	79	19	LN-	11.4442	6.	
13.	4.583 censored	wt	ER?	PgR-	71	12	LN?	11.2926	6.	
14.	11.5 censored	wt	ER+	PgR+	64	26	LN-	11.2061	6.	
15.	11.5 censored	wt	ER+	PgR+	54	17	LN-	11.4189	6.	
16.	11.5 censored	mutant	ER-	PgR-	68	6	LN-	10.8191	6.	
17.	11.5 censored	wt	ER+	PgR+	65	23	LN+	11.1728	6.	
18.	10.833 censored	wt	ER+	PgR+	67	25	LN-	10.686	6.	
19.	11.5 censored	wt	ER-	PgR-	67	27	LN-	11.2515	6.	
< > Rows:	236 Columns: 44936 <								>	~

Figure 9. Invoking Cox Regression

The Cox Regression dialog will open. Please note that in this tutorial data set, column 1. Survival (years) indicates the survival time of each patient in years and column 2. Event indicates the event status for each patient, death or censored.

- Set Time Variable to 1. Survival (years) using the drop-down menu
- Set Event Variable to 2. Event using the drop-down menu

Only numeric data are displayed in the Time Variable drop-down list and only categorical data with two categories are displayed in Event Variable.

• Set Event Status to death using the drop-down menu (Figure 2)

Event Status should be set to the primary event outcome. *All Response Variables* will be automatically selected for *Predictor*. This means that Cox Regression will test every probe set for association with the survival (time-to-event).

Cox Regression of Spreadsheet 1	×
Time Variable 1. Survival (years)	~
Event Variable 2. Event	V ? Event Status death
Predictor All Response Variables	~ 🜏
Candidate(s) 1. Survival (years) 2. Event 3. p53 status 4. ER status 5. PgR status 6. age at diagnosis 7. tumor size (mm) 8. Lymph node status	Co-predictor(s) ? Add Factor > <remove add="" factor=""> <add factor=""> <remove <remove="" factor="" factor<="" td=""></remove></add></remove>
	Model Results OK Cancel Apply
0	

Figure 10. Configuring the Cox Regression dialog

Co-predictors are numeric or categorical factors that will be included in the regression model. To evaluate the association between tumor size and gene expression, we can add tumor size to the co-predictors list.

- Select 7. tumor size (mm) from the *Candidate(s)* panel
 Select Add Factor > to add it to the *Co-predictor(s)* panel

Advanced options such as the inclusion of interactions between predictors and co-predictors can be accessed by selecting Model... (Figure 3) and the Resu /ts... button invokes a dialog (Figure 4) with additional output options for the results spreadsheet. We do not need to adjust any of the advanced model or output options for this tutorial.

Model Configure of Spreadsheet 1								
Predictor(s)	Model							
0. Response Variable	Add Factor >	0. Response Variable						
()	Add Interaction >	st daniel sze (mily						
	< Remove Factor							
		OK Cancel						

Figure 11. Configuring advanced options for Cox Regression

Result of Spreadsheet 1 ×
Display selected values in result spreadsheet
Chi-square
Coefficient
DOF
Chi-square of Model
p-value of Model
DOF of Model
Deselect All Select All Close

Figure 12. Configuring output options for Cox Regression

• Select **OK** to run Cox Regression (Figure 5)

-		
Cox Regress	ion of Spreadsheet 1	×
Time Variable	1. Survival (years)	~ ?
Event Variable	2. Event	V Pevent Status death
Predictor	All Response Variables	~ ?
Candidate(s)		Co-predictor(s) ?
1. Survival (years 2. Event 3. p53 status 4. ER status 5. PgR status 6. age at diagnos 8. Lymph node st	s) sis ratus	Add Factor > 7. tumor size (mm) < Remove Factor Stack (Categories)
		Add Factor > < Remove Factor
0		Model Results OK Cancel Apply

Figure 13. Configuring Cox Regression to assess the effect of gene expression and tumor size on survival

The spreadsheet generated by Cox Regression (Figure 6) includes a row for each probe set; the columns provide the following information:

- 1. Column # Column number of probe set in probe intensities spreadsheet
- 2. Probest ID ID of probe set in probe intensities spreadsheet
- 3. HRatio(gene) Hazard ratio for the probe set
- 4. LowCl(gene) lower 95% confidence boundary of the hazard ratio for the probe set
- 5. UpCl(gene) upper 95% confidence boundary of the hazard ratio for the probe set

6. p-value(gene) - P-value of the corresponding Chi-squared test. A low value indicates that the predictor (probe set) poses a large hazard or is associated with shortened survival time

7. to 10. - Effects of the co-predictor on survival time; for each co-predictor, a similar set of columns is added

11. modelfit(0) - P-value of the test assessing the overall model fit, i.e., the relationship between survival time, the predictor, and co-predictors in the model. A modelfit value of > 0.05 indicates a low association between the predictor and/or co-predictors with survival time.

Please note that the Cox Regression results spreadsheet is a temporary file. If you would like to be able to view the spreadsheet again after closing Partek

Genomics Suite, be sure to save it by selecting the Save Active Spreadsheet icon (

Partek Genomics Suite - 1/Cox (ptmp41)	Tools Wir	dow Custo	m Holn									- 0	
Eult Hansionni view Stat Thter	TOOIS WII	luow Custo	in nep							Workflo	ws Choose		
alvsis ×													
			\sim	<u> </u>									-
📨 📩 👧 🔛	🚰 🛞	7 🎫 🤇	2 T	6									
1 (Survival Tutorial)	Curren	t Selection 0.1	34799										
Cox (ptmp41)		1. Column #	2. Probeset ID	3. HRatio(gene)	4. LowCI(gene)	5. UpCI(gene)	6. p-value(gene)	7. HRatio(tumor size (mm))	8. LowCI(tumor size (mm))	9. UpCI(tumor size (mm))	10. p-value(tumor size (mm))	11. modelfit(0)	
	1.	17036	217663_at	0.00202612	3.04539e-05	0.134799	0.00378306	1.05248	1.03171	1.07368	4.93687e-07	6.23673e-07	1
	2.	21592	222224_at	0.00303448	5.12715e-05	0.179595	0.00535731	1.05171	1.03159	1.07222	3.11591e-07	7.84507e-07	
	3.	15366	215986_at	0.0363826	0.00163349	0.810346	0.0363656	1.04793	1.02669	1.06961	7.43324e-06	5.46226e-06	
	4.	9615	210123_s_at	0.0386421	0.00512911	0.291124	0.00159017	1.05556	1.03402	1.07754	2.72499e-07	2.40343e-07	
	5.	6809	207276_at	0.0477627	0.000805976	2.83045	0.144175	1.05188	1.03139	1.07278	4.67564e-07	1.69315e-05	
	6.	19297	219925_at	0.0490878	0.00463365	0.520025	0.0123149	1.04826	1.0276	1.06934	3.47451e-06	2.14722e-06	
	7.	8068	208562_s_at	0.0594123	0.00893444	0.39508	0.00349212	1.04657	1.02583	1.06774	8.32334e-06	5.39573e-07	
	8.	44089	244112_x_at	0.0597695	0.00982303	0.363675	0.00222903	1.05186	1.03112	1.07302	6.51012e-07	1.54865e-07	
	9.	14760	215378_at	0.0602972	0.00984263	0.369388	0.00239011	1.04599	1.02459	1.06784	2.01125e-05	2.20505e-07	
	10.	3646	204111_at	0.0683526	0.00460682	1.01416	0.0512025	1.04885	1.02838	1.06974	2.11107e-06	6.83646e-06	
	11.	14113	214729_at	0.070517	0.00434987	1.14317	0.0620618	1.05026	1.02983	1.07108	9.84872e-07	8.49759e-06	
	12.	16595	217220_at	0.0724613	0.00502776	1.04433	0.0538381	1.04753	1.02685	1.06864	5.02856e-06	7.32092e-06	
	13.	7921	208412_s_at	0.0756429	0.00629298	0.909243	0.0418515	1.05278	1.03199	1.07399	4.32196e-07	5.72169e-06	
	14.	20272	220900_at	0.076119	0.00198217	2.92311	0.166449	1.05046	1.02951	1.07183	1.66453e-06	1.93012e-05	
	15.	36080	236103_at	0.0793831	0.0100955	0.624206	0.0160437	1.05059	1.02973	1.07188	1.41925e-06	2.40582e-06	
	16.	7854	208343_s_at	0.0824477	0.00591038	1.15012	0.063456	1.05033	1.02979	1.07128	1.09784e-06	8.8296e-06	
	17.	32283	232301_at	0.0824969	0.0111905	0.608171	0.0143686	1.04959	1.02981	1.06975	6.16348e-07	2.49777e-06	
	18.	5588	206054_at	0.0836666	0.0086933	0.805229	0.0317522	1.0537	1.0331	1.07472	2.08525e-07	4.97345e-06	
	19.	7402	207875_at	0.083953	0.00469572	1.50096	0.0921881	1.04913	1.0285	1.07017	2.2057e-06	1.20678e-05	
	v - m	20220	220242 -1	0.0000745	0.00010000	0.05740	0.0000304	1.04070	1 00007	1 07070	1 00001- 00	F 00075- 00	

Figure 14. Cox Regression results spreadsheet

The hazard ratio is an effect size measure used to assess the direction and magnitude of the effect of a predictor variable on the relative likelihood of the event occurring at any given point in time, controlling for other predictors in the model.

For continuous predictors, such as gene expression values and tumor size, the hazard ratio is the predicted change in the hazard for a unit increase in the predictor. A hazard ratio greater than 1 indicates that the predictor is associated with shorter time-to-event, hazard ratio less than 1 indicates that the predictor is associated with shorter time-to-event, hazard ratio less than 1 indicates that the predictor is associated with greater time-to-event, and a hazard ratio of 1 indicates that the predictor has no effect on time-to-event. For categorical predictors, the hazard ratio is relative to the reference category.

For any probe set, we can view a detailed HTML report.

- Right-click the row header for row 1
- Select HTML Report from the pop-up menu (Figure 7)

sis X											Workfl	ows Choose	
产 🔒 📘 🔝 🖬 🗄	**		2 T		?								
(Survival_Tutorial)		ent Selection 1.50	096				-		-		-	10	
Cox (ptmp41)		1. Column #	2. Probeset	ID	3. HRatio(gene)	4. LowCI(gene)	5. UpCI(gene)	6. p-value(gene)	7. HRatio(tumor size (mm))	8. LowCI(tumor size (mm))	9. UpCI(tumor size (mm))	p-value(tumor size (mm))	modelfit(0)
	1.	Conv		ət	0.00202612	3.04539e-05	0.134799	0.00378306	1.05248	1.03171	1.07368	4.93687e-07	6.23673e-07
	2.	Paste		at	0.00303448	5.12715e-05	0.179595	0.00535731	1.05171	1.03159	1.07222	3.11591e-07	7.84507e-07
	3.	Filter Include		ət	0.0363826	0.00163349	0.810346	0.0363656	1.04793	1.02669	1.06961	7.43324e-06	5.46226e-06
	4.	Filter Exclude		_at	0.0386421	0.00512911	0.291124	0.00159017	1.05556	1.03402	1.07754	2.72499e-07	2.40343e-07
	5.	Tesert		at	0.0477627	0.000805976	2.83045	0.144175	1.05188	1.03139	1.07278	4.67564e-07	1.69315e-05
	6.	Delete		at	0.0490878	0.00463365	0.520025	0.0123149	1.04826	1.0276	1.06934	3.47451e-06	2.14722e-06
	7.			_at	0.0594123	0.00893444	0.39508	0.00349212	1.04657	1.02583	1.06774	8.32334e-06	5.39573e-07
	8.	Dot Plot (Orig	Data)	_at	0.0597695	0.00982303	0.363675	0.00222903	1.05186	1.03112	1.07302	6.51012e-07	1.54865e-07
	9.	Profile (Orig. I	Data)	at	0.0602972	0.00984263	0.369388	0.00239011	1.04599	1.02459	1.06784	2.01125e-05	2.20505e-07
	10.	Create List		at	0.0683526	0.00460682	1.01416	0.0512025	1.04885	1.02838	1.06974	2.11107e-06	6.83646e-06
	11.	14110	214/29_	at	0.070517	0.00434987	1.14317	0.0620618	1.05026	1.02983	1.07108	9.84872e-07	8.49759e-06
	12.	16595	217220_3	at	0.0724613	0.00502776	1.04433	0.0538381	1.04753	1.02685	1.06864	5.02856e-06	7.32092e-06
	13.	7921	208412_	s_at	0.0756429	0.00629298	0.909243	0.0418515	1.05278	1.03199	1.07399	4.32196e-07	5.72169e-06
	14.	20272	220900_	at	0.076119	0.00198217	2.92311	0.166449	1.05046	1.02951	1.07183	1.66453e-06	1.93012e-05
	15.	36080	236103_	at	0.0793831	0.0100955	0.624206	0.0160437	1.05059	1.02973	1.07188	1.41925e-06	2.40582e-06
	16.	7854	208343_	s_at	0.0824477	0.00591038	1.15012	0.063456	1.05033	1.02979	1.07128	1.09784e-06	8.8296e-06
	17.	32283	232301_	at	0.0824969	0.0111905	0.608171	0.0143686	1.04959	1.02981	1.06975	6.16348e-07	2.49777e-06
	18.	5588	206054_	at	0.0836666	0.0086933	0.805229	0.0317522	1.0537	1.0331	1.07472	2.08525e-07	4.97345e-06
	19.	7402	207875_	at	0.083953	0.00469572	1.50096	0.0921881	1.04913	1.0285	1.07017	2.2057e-06	1.20678e-05
	v [m	20220	220242		0.0003745	0.00010000	0.05743	0.0000304	1.04070	1 00007	1 07070	1 00001- 00	E 0007E- 00

Figure 15. Invoking an HTML report for a probe set

The HTML report (Figure 8) will open in your default web browser.

Cox Regression Result of 217663_at

Produced by Partek Genomics Suite

Model Information

Test	Chi Square	DF	p-value
Likelihood Ratio	28.575278	2	6.23673e-07
Wald	31.412603	2	1.50952e-07
Score	32.678824	2	8.01463e-08

Coefficient Information

Name	DF	Estimate	Std Error	W (Wald Chi Square)	p-value (W)	Hazard Ratio	LowCl	UpCI
217663_at	1	-6.201632	2.141665	8.385116	0.00378306	0.00202612	3.04539e- 05	0.134799
tumor size (mm)	1	0.0511531	0.0101721	25.288328	4.93687e- 07	1.052484	1.031708	1.073678

Model Fit Statistics

Without Predictor(s) With Predictor(s) -2logL 570.209215 541.633937 AIC 570.209215 545.633937 SBC 570.209215 549.648603

Figure 16. Cox Regression HTML report

Additional Assistance

If you need additional assistance, please visit our support page to submit a help ticket or find phone numbers for regional support.

	•.	

go to top

<u>go to top</u>

go to top



Your Rating: ☆☆☆☆☆ Results: ★★★★ 34 rates