

また上図右の設定 (16) では、RR 比較図⑬の表示域、P 値比較図⑭の表示域を設定できます。また推定生存曲線⑮に対しては、それを、2 値型の説明変数を指定して、分割表示できます。上図左は、その機能で群分けした例です。

**例題 16**：肺癌 137 例について、治療法、組織型の 4 型、肺機能、罹病期間、年齢、前治療の有無と、生存期間、生死判定との関連を調べたい。組織型については、扁平上皮癌を基準カテゴリーとして 3 つのダミー変数（小細胞癌、腺癌、大細胞癌）を作成している

- 1) Cox 回帰分析により、生存期間に関係する因子を調べよ
- 2) 肺機能（0～100）が -20、年齢が +10 才の変化に対する相対リスクを求めよ。



治療法	組織型	小細胞癌	腺癌	大細胞癌	肺機能	罹病期間	年齢	前治療	生存日数	生死
1	扁平上皮癌	0	0	0	20	21	65	1	1	1
1	扁平上皮癌	0	0	0	50	7	35	0	1	1
1	小細胞癌	1	0	0	40	36	44	1	2	1
0	腺癌	0	1	0	30	3	43	0	3	1
0	小細胞癌	1	0	0	40	2	35	0	4	1
0	小細胞癌	1	0	0	50	7	72	0	7	1
1	小細胞癌	1	0	0	20	11	66	0	7	1
1	腺癌	0	1	0	40	4	58	0	7	1
0	扁平上皮癌	0	0	0	40	58	63	1	8	1
0	腺癌	0	1	0	20	19	61	1	8	1
1	小細胞癌	1	0	0	80	2	68	0	8	1
1	腺癌	0	1	0	50	5	66	0	8	1
-	<以下省略>	-	-	-	-	-	-	-	-	-

治療法: 1=新法 0=標準 生死: 1=死亡 0=生存

n=137: Kleinbaum: Survival Analysis, Springer, 1996 より改変して引用

データは、[SampleData] フォルダの [DB 型 肺癌の生存分析.SFD6] です。読み込むとデータシートは次頁のようになっています。目的変数に観察期間と生死判定を指定し、残りの変数のうち、3 つのダミー変数と年齢を自動除外しないようにチェックをつけ、限界 P 値を 0.1 として、**変数自動選択実行** を行うと、結果は次のようになります。

次数	変数名	$\beta$	SE( $\beta$ )	z	P	相対リスク	95%CI下限	95%CI上限
1	小細胞癌	0.7208	0.2529	2.8498	0.00437	2.056124	1.25242	3.37559
2	腺癌	1.164	0.2936	3.9654	0.00007	3.203826	1.80192	5.69643
3	大細胞癌	0.3215	0.2766	1.1623	0.24512	1.379161	0.80202	2.37163
4	年齢	-5.899E-3	9.057E-3	.6513	0.51486	0.994118	0.97663	1.01192
5	肺機能	-0.03183	5.401E-3	5.8938	0.00000	0.96867	0.95847	0.97898

**出力例**

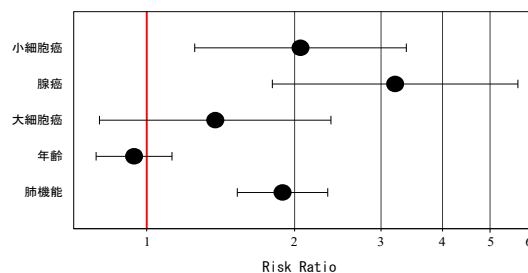
<< Coxの比例ハザード回帰 >>  
 目的変数: 生存日数, 生死 有効データ数 = 137

次数	変数名	$\beta$	SE( $\beta$ )	z	P	相対リスク	95%CI下限	95%CI上限
1	小細胞癌	0.7208	0.2529	2.8498	0.00437	2.056124	1.25242	3.37559
2	腺癌	1.164	0.2936	3.9654	0.00007	3.203826	1.80192	5.69643
3	大細胞癌	0.3215	0.2766	1.1623	0.24512	1.379161	0.80202	2.37163
4	年齢	-5.899E-3	9.057E-3	.6513	0.51486	0.994118	0.97663	1.01192
5	肺機能	-0.03183	5.401E-3	5.8938	0.00000	0.96867	0.95847	0.97898

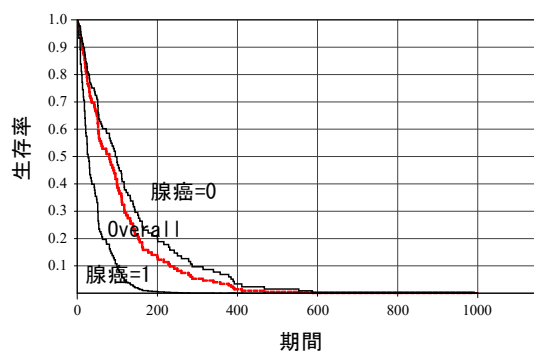
回帰の適合度指標  
 AIC=962.57860

ここで、RR の値を適正化するため、有意ではないが制御変数の年齢の変化量を+10、最も有意性の高い肺機能の変化量を -20 としたときの、分析結果を「相対リスク計算の変化量の指定」で行うと次のようになります。

Ord	Var	$\Delta$	$\beta$	SE( $\beta$ )	P-val	RR	CI-LL	CI-UL
1	小細胞癌	1	0.721	0.253	0.0044	2.06	1.25	3.38
2	腺癌	1	1.164	0.294	0.0001	3.20	1.80	5.70
3	大細胞癌	1	0.321	0.277	0.2451	1.38	0.80	2.37
4	年齢	10	-0.006	0.009	0.5149	0.94	0.79	1.13
5	肺機能	-20	-0.032	0.005	0.0000	1.89	2.34	1.53



この分析結果をもとに作成された生存曲線は、右上のページタブ「推定生存曲線」をクリックすると、次のようになります。初期作図では、2 値型で一番有意性の高かったダミー変数腺癌について、それが 0 と 1 の場合の比較図と、腺癌を考慮しない場合の曲線（赤色）が表示されています。設定タブでは、任意の説明変数について、その境界値を指定して、境界値以下と以上に対する推定曲線を作成できます。

R による分析結果<sup>17</sup>

```
> library(survival)
> dat = read.csv("DB型 肺癌の生存分析.csv", header=T)
> res = coxph(data=dat, Surv(生存日数, 生死)~小細胞癌+腺癌+大細胞癌+年齢+肺機能, method="breslow")
> summary(res)
Call:
coxph(formula = Surv(生存日数, 生死) ~ 小細胞癌 + 腺癌 + 大細胞癌 +
      年齢 + 肺機能, data = dat, method = "breslow")

      n= 137, number of events= 128

              coef exp(coef)  se(coef)      z Pr(>|z|)
小細胞癌  0.720823  2.056124  0.252937  2.850  0.00437 **
腺癌      1.164346  3.203826  0.293625  3.965  7.33e-05 ***
大細胞癌  0.321475  1.379161  0.276587  1.162  0.24512
年齢     -0.005899  0.994118  0.009057  -0.651  0.51486
肺機能   -0.031831  0.968670  0.005401  -5.894  3.77e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              exp(coef) exp(-coef) lower .95 upper .95
小細胞癌    2.0561    0.4864    1.2524    3.376
腺癌        3.2038    0.3121    1.8019    5.696
大細胞癌    1.3792    0.7251    0.8020    2.372
年齢        0.9941    1.0059    0.9766    1.012
肺機能      0.9687    1.0323    0.9585    0.979

Concordance= 0.735 (se = 0.03 )
Rsquare= 0.351 (max possible= 0.999 )
Likelihood ratio test= 59.19 on 5 df,  p=1.787e-11
Wald test               = 59.67 on 5 df,  p=1.419e-11
Score (logrank) test = 63.22 on 5 df,  p=2.617e-12
```

<sup>17</sup>注: cox 回帰には、同じ時点で複数のイベントが発生した場合 (tie)、それに対する補正法が幾つかあります。大部分の統計ソフトでは、"breslow"を利用しています。R の default 設定 (methodo を指定しない場合) は、"efron"法となります。