Prognosis of survival for breast cancer patients

Ken Ryder

Breast Cancer Unit Data Section

Guy's Hospital

Patrick Royston,

MRC Clinical Trials Unit

London

Outline

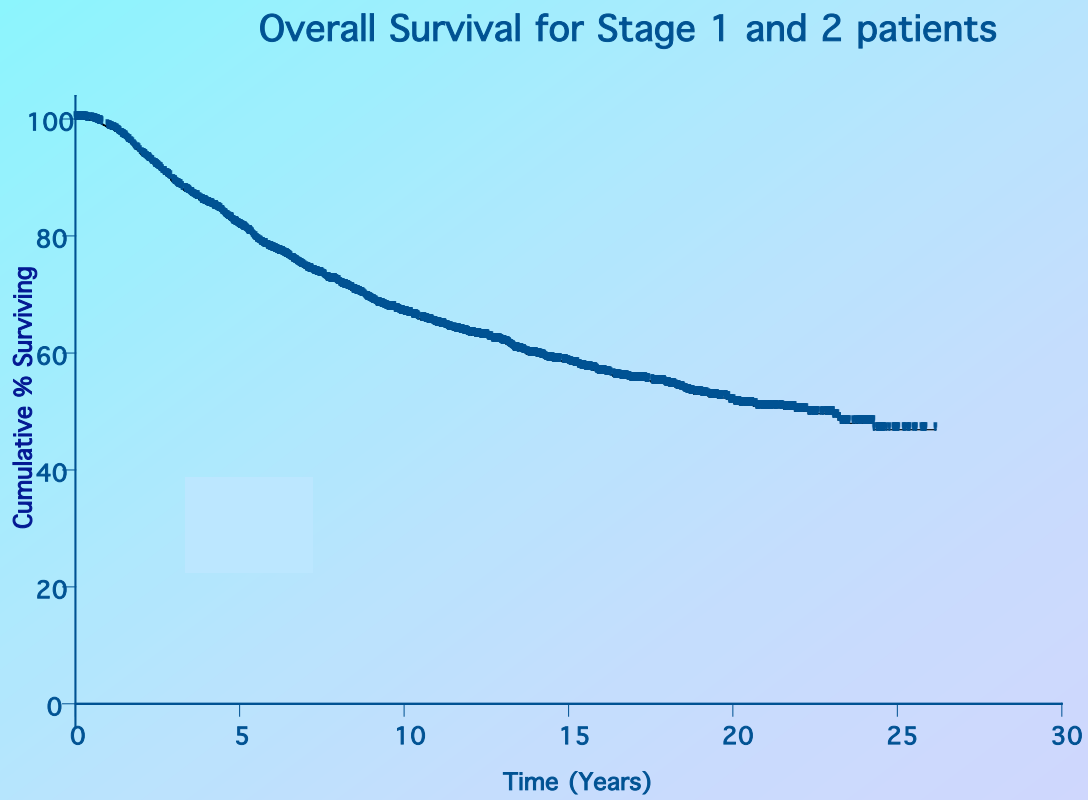
- **Introduce the data and outcomes requested by the medics.**
- **Work through the stages of the analyses.**
  - **Imputation**
  - **Fractional Polynomials**
  - **Producing a simple score**
  - **Fitting Splines**
- **Illustrate the menu call developed for everyday use.**

The Problem

- Over 8000 patients with invasive Breast Cancer with 3083 operable stage 1 and 2 disease.
- Time of relapse-free survival
- Time to a distant relapse
- Overall survival
  - Death from breast cancer
  - All-cause mortality

The Problem

- **Kaplan-Meier Curves**



The Problem
Factors affecting outcome

- **Histological type (x_4)**
  Ductal (Grade I, II & III), Lobular, others (Mucoid, Medullary, Tubular)
- **Tumour Size (x_3)**
  (≤ 2.0 cms, 2.0 - 5.0 cms, >5.0 cms)
- **Axilla Nodal status (x_5)**
  (All negative, 1-3, 4 - 9, ≥10 positive)
- **Age (x_1)**
  (Under 40, 41-50, 51-60, 61-70, Over 70)

The Problem
Factors affecting outcome

- Postmenopausal status (x_2)
- Type of surgery (x_{11})
- Progeterone receptor status (x_{6a})
- Oestrogen receptor status (x_{7a})
- Adjuvant treatment
  Chemotheraphy (x_8)
- Hormonal therapy (x_9)
- Radiotherapy (x_{10})

The Task

- **Given the information on the prognostic factors for a new patient**

- **To provide estimates of the probabilities of being relapse-free or surviving after 5, 10, 15, 20 years.**

Prognostic Factors

	Code	Mean	% missing
Age (years)	x_1	55	0
Postmenopausal status	x_2	0.55	2.0
Clinical tumour size (mm)	x_3	28	1.1
Histology Ductal Grade 2	x_{4a}	0.39	0
Histology Ductal Grade 3	x_{4b}	0.34	0
Histology Lobular	x_{4x}	0.13	0
Histology Other	x_{4y}	0.03	0
No of positive lymph nodes	x_5	2.8	0
Progesterone receptor positive	x_{6a}	0.52	13.7
Oestrogen receptor positive	x_{7a}	0.70	11.5
Chemotherapy (any type)	x_8	0.20	0
Hormonal therapy	x_9	0.35	0
Radiotherapy	x_{10}	0.42	0
Impalable tumour	x_{3x}	0.07	1.1

Imputation

van Buuren S., H. C. Boshuizen and D. L. Knook. 1999. Multiple imputation of missing blood pressure covariates in survival analysis. Statistics in Medicine 18:681-694.

mvisamp imputes missing values using "switching regression", an iterative multivariable regression technique.

mvisamp varlist [if exp] [in range] [weight], gen(name)
[cc(ccvarlist) cmd(cmdlist) seed(#)]

Imputation

Impute x6a and x7a on complete variables and transformed year and survival time

```
mvisamp x6a x7a   x1 x4a x4b x4x x4y x5 yr_1 yr_2
yr_3 lnt _d, gen(z1) seed(101)
```

Impute x3 on these variables and completed x6a, x7a

```
mvisamp x3  z1_1 z1_2 x1 x4a x4b x4x x4y x5 yr_1
yr_2 yr_3 lnt _d if x3>0,   gen(z2) seed(103)
```

Cox's Proportional Hazards model
Fractional Polynomials

- For the continuous variables, the Fractional Polynomials, based on X^{-2}, X^{-1}, $X^{-0.5}$, $X^{0.5}$, X^{2}, X^{3},and $\log_e(X)$ are compared with the linear polynomial to determine if one or more of these polynomials provides a better fit.

- All other factors are (0,1) variables, including the interaction of Hormone Therapy with ER status.

Cox's Proportional Hazards model
Fractional Polynomials

mfracpol selects the fractional polynomial (FP) model which best predicts the outcome variable, yvar, from the RHS variables, xvarlist. After execution, mfracpol leaves variables in the data named Ivar1__1, Ivar2__2, ...Ivar2_1, Ivar2_2 etc. The new variables contain the best-fitting fractional polynomial powers of var1, var2, .

mfracpol regression_cmd yvar xvarlist [in range] [if exp] [weight] [, alpha(alpha_list) df(df_list) select(select_list) regression_cmd_options]

Cox's Proportional Hazards model
Fractional Polynomials

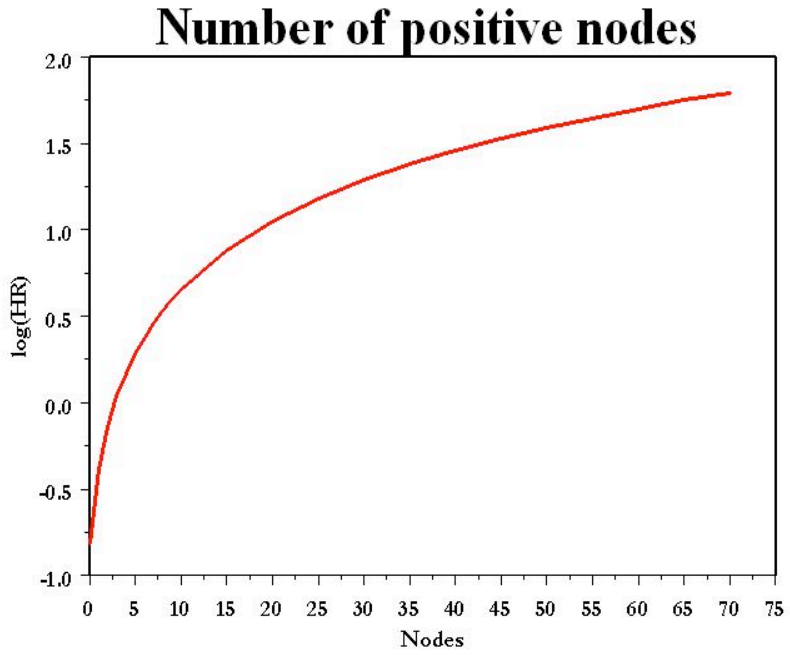
* Force in ER and hormon, since know ER x hormon is significant.

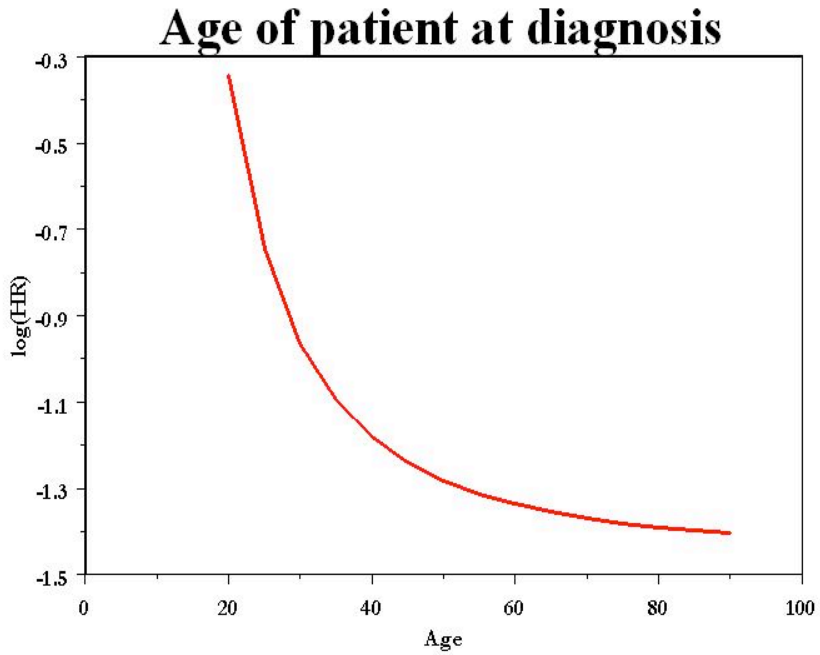
mfracpol cox _t x1 ix2 ix3 ix3x x4a x4b x4x x4y x5 ix6a ix7a x7ah hormon chemo rt yr, dead(_d) sel(.05, ix7a hormon x7ah:0.9999) df(ix3 x5:2)

Cox's Proportional Hazards model
Fractional Polynomials

```
-> gen double lx1__1 = X^-2-.0327 if e(sample)
   (where: X = x1/10)
-> gen double lix3__1 = ln(X)+1.243 if e(sample)
   (where: X = (ix3+1)/100)
-> gen double lx5__1 = ln(X)+.974 if e(sample)
   (where: X = (x5+1)/10)
```

Cox's Proportional Hazards model
Fractional Polynomials





Cox's Proportional Hazards model
Fractional Polynomials

Final multivariable fractional polynomial model for _t

Variable	Initial df	Initial Select	Initial Alpha	Final Status	Final df	Final Powers
x1	4	0.0500	0.0500	in	2	-2
ix2	1	0.0500	0.0500	out	0	
ix3	2	0.0500	0.0500	in	2	0
ix3x	1	0.0500	0.0500	in	1	1
x4a	1	0.0500	0.0500	in	1	1
x4b	1	0.0500	0.0500	in	1	1
x4x	1	0.0500	0.0500	in	1	1
x4y	1	0.0500	0.0500	out	0	
x5	2	0.0500	0.0500	in	2	0
ix6a	1	0.0500	0.0500	in	1	1
ix7a	1	0.9999	0.0500	in	1	1
x7ah	1	0.9999	0.0500	in	1	1
hormon	1	0.9999	0.0500	in	1	1
chemo	1	0.0500	0.0500	in	1	1
rt	1	0.0500	0.0500	out	0	
yr	4	0.0500	0.0500	out	0	

Cox's Proportional Hazards model
Fractional Polynomials

_t / _d	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Ix1__1	4.707014	1.439364	3.27	0.001	1.885911	7.528116
Iix3__1	.5143961	.0655541	7.85	0.000	.3859123	.6428798
ix3x	1.320571	.2790599	4.73	0.000	.7736239	1.867518
x4a	1.122357	.1557526	7.21	0.000	.8170873	1.427626
x4b	1.420355	.1588664	8.94	0.000	1.108983	1.731728
x4x	1.028325	.1704158	6.03	0.000	.6943161	1.362334
Ix5__1	.6070695	.0299867	20.24	0.000	.5482966	.6658423
ix6a	-.1314727	.066974	-1.96	0.050	-.2627393	-.000206
ix7a	.1902246	.0781319	2.43	0.015	.037089	.3433603
x7ah	-.708684	.1504531	-4.71	0.000	-1.003567	-.4138013
hormon	-.0381773	.1254731	-0.30	0.761	-.2841	.2077454
chemo	-.3198936	.0782234	-4.09	0.000	-.4732086	-.1665787

Deviance:16524.958.

Cox's Proportional Hazards model
Fractional Polynomials

. global model `e(fp_fvl) '

. stcox $model, nolog sca(sca*) sch(sch*)

. stphtest, rank detail

Test of proportional hazards assumption

	rho	chi2	df	Prob>chi2
Ix1__1	**-0.08739**	**10.11**	**1**	**0.0015**
Ix3_1__1	0.03250	1.23	1	0.2666
x4a	**-0.06038**	**4.17**	**1**	**0.0412**
x4b	**-0.11912**	**16.27**	**1**	**0.0001**
x4x	-0.01873	0.40	1	0.5268
Ix5__1	**-0.14572**	**23.72**	**1**	**0.0000**
ix6a	**0.06161**	**4.59**	**1**	**0.0321**
ix7a	**0.08467**	**8.35**	**1**	**0.0039**
x7ah	0.03161	1.14	1	0.2847
hormon	-0.01300	0.19	1	0.6609
chemo	**0.11522**	**16.23**	**1**	**0.0001**
global test		**145.80**	**11**	**0.0000**

Simple scoring system

- **A scoring system was developed so that a unit increase in the score was approximately an increase of 0.1 in the $\log_e$(HR).**
- **For the (0,1) binary variables the scores were taken to the nearest 0.1 of the $\log_e$(HR), but for the continuous variables the cut-off points were selected to have an increment of approximately 0.2 in the $\log_e$(HR).**

Simple scoring system

Node status	Score	Tumour size	Score
0	0	≤11mm	0
1	4	12 – 19	2
2	6	20 – 27	4
3 – 4	8	28 – 38	6
5 – 6	10	≥ 39	8
7 – 9	12		
10 – 14	14		
15 – 20	16	ER^+ and HT	0
21 – 29	18	ER^+ and no HT	8
30 – 42	20	ER^-	7
43 – 52	22		
≥ 53	24		

Simple scoring system
Fitting splines

stpm [varlist] [if exp] [in range] , model_complexity scale(hazard|normal|odds) stratify(strat_varlist)

stpm fits spline-based distributional models to right- or interval-censored survival-time data.

varlist is a set of covariates.

model_complexity given by defining the number of degrees of freedom to use or by defining the knot positions

scale specifies the scale of model

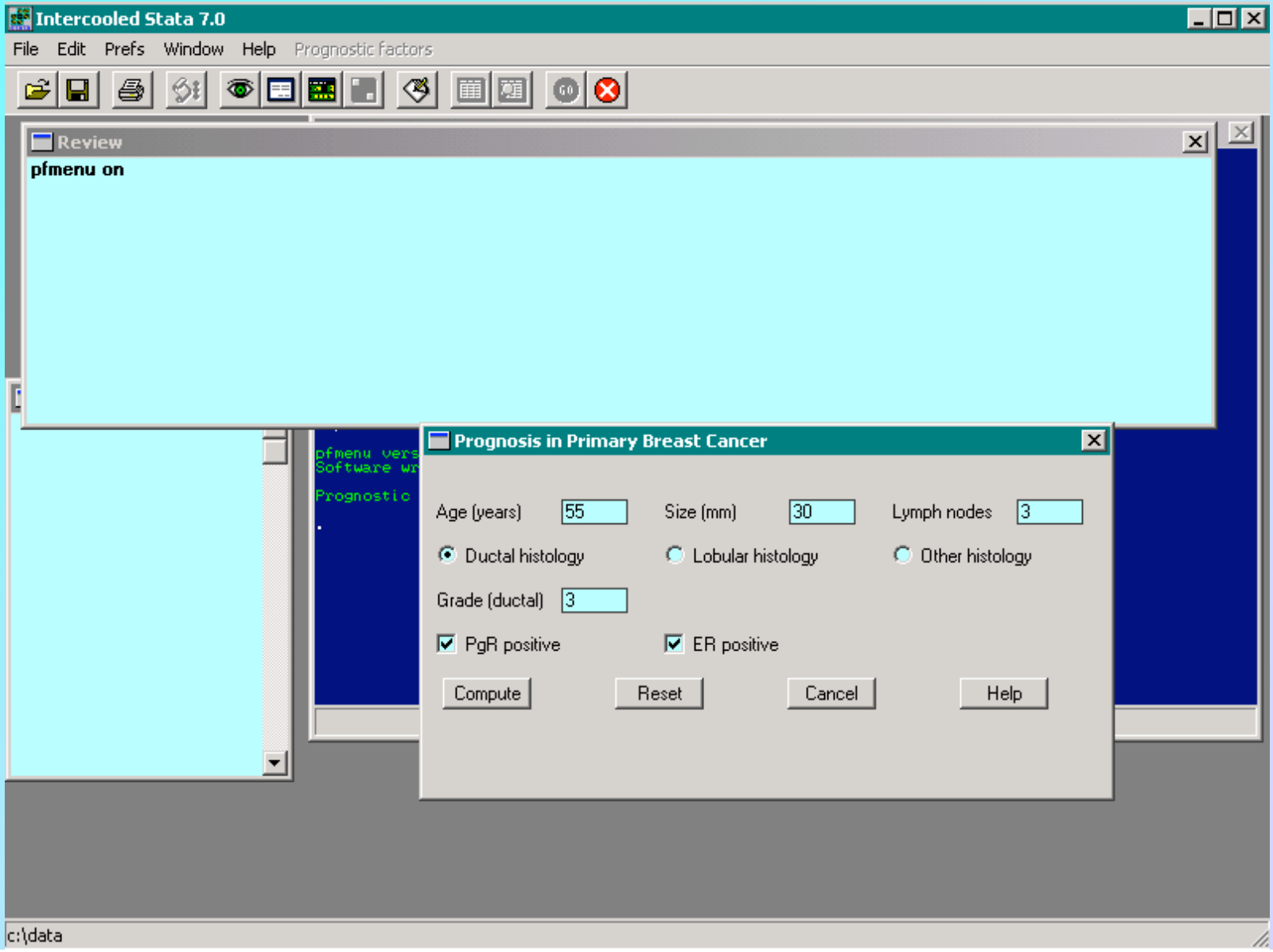
stratify

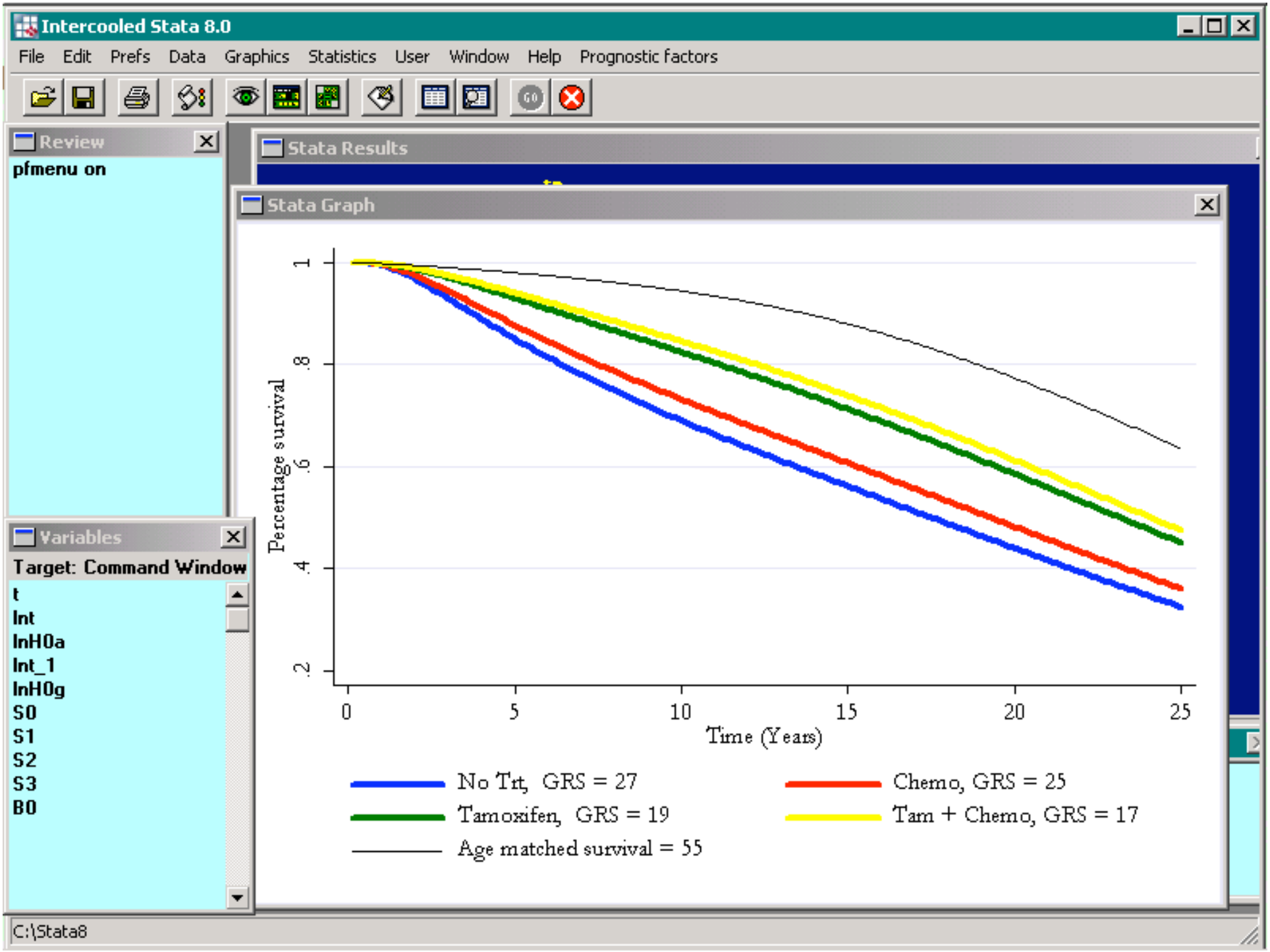
stpm guysrisk, st(guysrisk) sc(h) knots(0.643 1.614) nolog

Simple scoring system

Score	1	2	5	10	20
10	0.989	0.973	0.942	0.914	0.881
14	0.984	0.960	0.914	0.873	0.827
18	0.976	0.940	0.873	0.815	0.751
22	0.964	0.911	0.815	0.735	0.650
26	0.947	0.869	0.735	0.628	0.522
30	0.921	0.809	0.629	0.497	0.376
34	0.884	0.727	0.498	0.348	0.229
38	0.830	0.619	0.350	0.204	0.109
42	0.755	0.485	0.205	0.092	0.035
46	0.656	0.337	0.092	0.027	0.006

Menu system


Intercooled Stata 7.0
File Edit Prefs Window Help Prognostic factors
Review
pfmenu on
Prognosis in Primary Breast Cancer
Age (years) 55
Size (mm) 30
Lymph nodes 3
Ductal histology
Lobular histology
Other histology
Grade (ductal) 3
PgR positive
ER positive
Compute
Reset
Cancel
Help
c:\data


Intercooled Stata 8.0
File Edit Prefs Data Graphics Statistics User Window Help Prognostic factors
Review
pfmenu on
Stata Results
Stata Graph
Variables
Target: Command Window
t
lnt
lnH0a
lnt_1
lnH0g
S0
S1
S2
S3
B0
Percentage survival
.2
.4
.6
.8
1
0
5
10
15
20
25
Time (Years)
No Trt, GRS = 27
Chemo, GRS = 25
Tamoxifen, GRS = 19
Tam + Chemo, GRS = 17
Age matched survival = 55
C:\Stata8