

## **DataSet: NCCTG Lung Cancer Data**

### **Description**

Survival in patients with advanced lung cancer from the North Central Cancer Treatment Group.

Performance scores rate how well the patient can perform usual daily activities.

### **Format**

inst: Institution code

time: Survival time in days

**status: censoring status 1=censored, 2=dead**

age: Age in years

sex: Male=1 Female=2

ph.ecog: ECOG performance score as rated by the physician. 0=asymptomatic, 1= symptomatic but completely ambulatory, 2= in bed <50% of the day, 3= in bed > 50% of the day but not bedbound, 4 = bedbound

ph.karno: Karnofsky performance score (bad=0-good=100) rated by physician

pat.karno: Karnofsky performance score as rated by patient

meal.cal: Calories consumed at meals

wt.loss: Weight loss in last six months (pounds)

### **Note**

The use of 1/2 for alive/dead instead of the usual 0/1 is a historical footnote. For data contained on punch cards, IBM 360 Fortran treated blank as a zero, which led to a policy within the section of Biostatistics to never use "0" as a data value since one could not distinguish it from a missing value. The policy became a habit, as is often the case; and the 1/2 coding endured long beyond the demise of punch cards and Fortran.