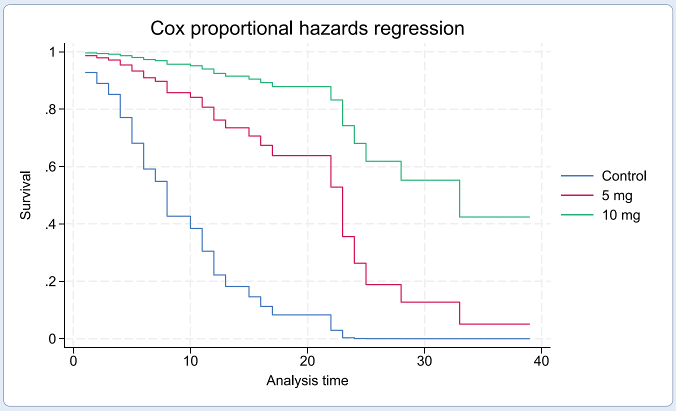
Survival analysis

From Kaplan–Meier estimates of the survivor function to the Cox proportional hazards model, from competing-risks regression to multilevel survival models, Stata has everything you need to analyze your survival- or event-time data.



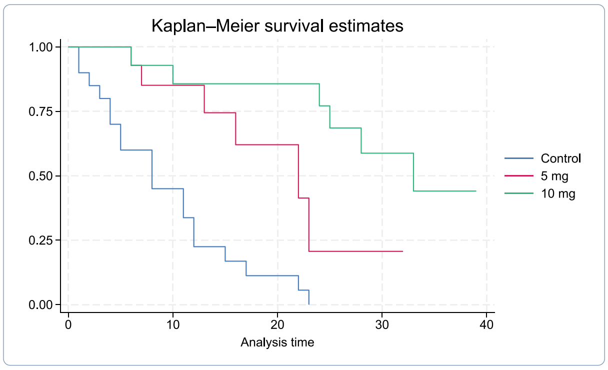
- **Survival-time data**
  - Single failure or multiple failures; right-, left-, and interval-censoring; left-truncation; gaps
  - Multiple events New
  - Support for complex survey designs
- **Life tables**
  - Tables and graphs with CIs
  - Tests for equality of survivor functions
  - Tests for trend
- **Graph survivor, cumulative hazard, and other functions** Updated
- **Cox proportional hazards model**
  - Stratified estimation
  - Time-varying covariates
  - Shared frailty models
  - Harrell's *C*, Somer's *D*, Gönen and Heller's *K*
  - Tests for proportional hazards
  - Goodness-of-fit plots Updated
- **Parametric survival models**
  - Weibull, exponential, Gompertz, lognormal, loglogistic, and generalized gamma
  - Stratified models
  - Individual or shared frailty
  - Predictions of mean or median time to failure, survival probabilities, and hazards
  - Goodness-of-fit plots Updated
  - Bayesian estimation
  - Finite mixture models
- **Competing risks model**
  - Fine and Gray proportional subhazards model
  - Graph cumulative subhazard and cumulative incidence
- **Multilevel survival models**
  - Weibull, exponential, lognormal, loglogistic, and gamma
  - Marginal predictions and marginal means
- **Structural equation models**
  - Weibull, exponential, lognormal, loglogistic, and gamma models
  - Survival outcomes with other outcomes
  - Path models, growth curve models, and more
- **Additive models of relative risk**
  - Cox, parametric survival, interval-censored Cox, and interval-censored parametric survival models
- **Power analysis**
  - Log-rank test of survival curves, Cox models, exponential regression
- **Sample-size analysis for group sequential designs (GSDs)**
  - Log-rank test of survival curves
- **Causal inference (treatment-effects estimation)**
  - Regression adjustment, inverse-probability weighting (IPW), and doubly robust methods
  - Average treatment effects (ATEs) and ATEs on the treated (ATETs)
- **Lasso and elastic net for Cox model**
  - Select predictors via cross-validation, adaptive lasso, or BIC
  - Penalized and postselection predictions
  - Graph survivor and other functions

We begin by specifying that we have survival data using **stset**. Here **studytime** records the time of failure or censoring, and the variable **died** indicates whether the subject died or was censored.

```
. stset studytime, failure(died)

Survival-time data settings

         Failure event: died!=0 & died<.
Observed time interval: (0, studytime]
     Exit on or before: failure

--------------------------------------------------------------------------
         48  total observations
          0  exclusions
--------------------------------------------------------------------------
         48  observations remaining, representing
         31  failures in single-record/single-failure data
        744  total analysis time at risk and under observation
                                                At risk from t =         0
                                     Earliest observed entry t =         0
                                          Last observed exit t =        39
```

We are now ready to graph the survivor function for each level of our treatment,

```
. sts graph, by(dose)
```



or to fit a Cox proportional hazards model,

```
. stcox age i.dose

        Failure _d: died
  Analysis time _t: studytime

Iteration 0:  Log likelihood = -99.911448
Iteration 1:  Log likelihood = -82.331523
Iteration 2:  Log likelihood = -81.676487
Iteration 3:  Log likelihood = -81.652584
Iteration 4:  Log likelihood = -81.652567
Refining estimates:
Iteration 0:  Log likelihood = -81.652567

Cox regression with Breslow method for ties

No. of subjects =  48                                  Number of obs =     48
No. of failures =  31
Time at risk    = 744
                                                       LR chi2(3)    =  36.52
Log likelihood = -81.652567                            Prob > chi2   = 0.0000

------------------------------------------------------------------------------
          _t | Haz. ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
         age |   1.118334   .0409074     3.06   0.002     1.040963    1.201455
             |
        dose |
        5 mg |   .1805839   .0892742    -3.46   0.001     .0685292    .4758636
       10 mg |   .0520066    .034103    -4.51   0.000     .0143843    .1880305
------------------------------------------------------------------------------
```

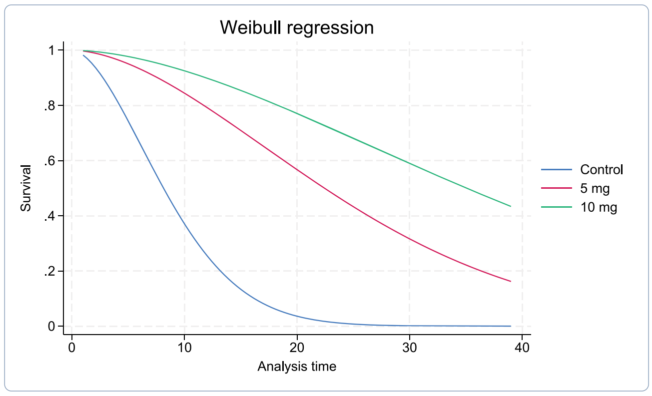
and test for violations of the proportional-hazards assumption.

```
. estat phtest, detail

Test of proportional-hazards assumption

Time function: Analysis time
--------------------------------------------------------
             |      rho     chi2       df     Prob>chi2
-------------+------------------------------------------
         age |  -0.06752    0.13        1       0.7199
     1b.dose |         .       .        1            .
      2.dose |   0.13698    0.50        1       0.4796
      3.dose |  -0.07380    0.16        1       0.6921
-------------+------------------------------------------
 Global test |              1.01        3       0.7994
--------------------------------------------------------
```

We can fit a Weibull model,

```
. streg age i.dose, distribution(weibull)
```

and plot the estimated survivor function for each dosage level,



or compute the marginal predictions of mean survival time for each dosage.

```
. margins dose, predict(mean time)

Predictive margins                                       Number of obs = 48
Model VCE: OIM

Expression: Predicted mean _t, predict(mean time)

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
        dose |
     Control |   9.574003   1.336857     7.16   0.000     6.953812    12.19419
        5 mg |   26.36447   6.992949     3.77   0.000     12.65854     40.0704
       10 mg |   41.19108   10.31993     3.99   0.000     20.96438    61.41778
------------------------------------------------------------------------------
```

And that's just the beginning.