The Stata Journal (2004)
4, Number 3, pp. 257–264

Compliance-adjusted intervention effects in survival data

Lois G. Kim
MRC Biostatistics Unit

Ian R. White
MRC Biostatistics Unit

Abstract. Survival data are most frequently analyzed by the intention-to-treat principle. However, presenting a compliance-adjusted analysis alongside the primary analysis can provide an insight into the effect of the treatment for those individuals actually complying with their randomized intervention. There are a number of methods for this type of analysis. Loeys and Goetghebeur (2003) use proportional hazards techniques to provide an estimate of the treatment effect for compliers when compliance is measured on an all-or-nothing scale. This methodology is here made available through a new Stata command, `stcomply`.

Keywords: st0068, stcomply, compliance, proportional hazards

1 Introduction

Time-to-event endpoints are a common outcome of interest in randomized clinical trials. In these situations, the primary analysis is most often done by the intention-to-treat principle, which gives an indication of the effectiveness of the intervention in the whole population. However, the effectiveness of the intervention for a specific individual choosing to undertake the intervention regime is also of interest. Such results are becoming increasingly important as patient decisions based on informed choice in health care become more widespread.

Effectiveness is defined as the benefit of intervention as actually applied and is estimated from simple all-or-nothing compliance data. Efficacy, on the other hand, is the benefit of intervention under ideal circumstances and requires more complex compliance data. Intervention effectiveness and efficacy after accounting for noncompliance can be estimated in a number of ways (Loeys and Goetghebeur 2003; Sommer and Zeger 1991; Cuzick, Edwards, and Segnan 1997; and Robins and Tsiatis 1991), some of which have already been implemented in Stata (e.g., `strbee` [White, Walker, and Babiker 2002]). A recent publication by Loeys and Goetghebeur (2003) provides new methodology using proportional hazards techniques in survival data where compliance is all or nothing in the intervention arm and perfect in the control arm. We implement their method in Stata. The output is in the form of a hazard ratio (with confidence intervals) for the effectiveness of intervention, adjusted for the observed adherence to intervention in the treated group.

2 Methodology

The method for calculating compliance-adjusted intervention effects used here was developed by Loeys and Goetghebeur (2003), and full details of the methodology can be found in their paper. Individuals in the control arm are classified as compliers and noncompliers according to how they would have behaved if they had been randomized to the intervention group. The proportion of noncompliers, α, is the same in both arms. Denote the probability of survival to time t as $S_{n0}(t)$ and $S_{c0}(t)$ for noncompliers and compliers randomized to control, and as $S_{n1}(t)$ and $S_{c1}(t)$ for noncompliers and compliers randomized to intervention, so that for each arm $j = 0, 1$:

$$S_j(t) = \alpha S_{nj}(t) + (1 - \alpha) S_{cj}(t)$$

Assume that allocation to intervention has no effect on noncompliers and has hazard ratio ψ for compliers:

$$\begin{aligned} S_{no}(t) &= S_{n1}(t) \\ S_{c0}(t) &= S_{c1}(t)^{1/\psi} \end{aligned}$$

Estimation of the compliance-adjusted intervention effect, $\widehat{\psi}$, is achieved by using Kaplan–Meier estimates of $S_{n1}(t)$ and $S_{c1}(t)$ to estimate the survivor function in the control arm:

$$\widehat{S}_0^*(t|\psi) = \widehat{\alpha}\widehat{S}_{n1}(t) + (1 - \widehat{\alpha})\widehat{S}_{c1}(t)^{1/\psi}$$

A value of ψ is found at which this quantity matches the observed survival in the control arm, in the sense that it predicts the correct number of events in the control arm. Define

$$\begin{aligned} \widehat{\Lambda}_0^*(t|\psi) &= -\log \widehat{S}_0^*(t|\psi) \\ G_0^*(\psi) &= \sum_j \left[\widehat{\Lambda}_0^*(T_j|\psi) - \delta_j \right] \end{aligned}$$

where the sum is over all individuals in the control group, T_j is the censoring/event time for the jth individual, and δ_j is the failure indicator for the jth individual. $G_0^*(\psi)$ can be understood as the difference between observed and expected events in the control arm, based on predictions from the intervention arm if the hypothesized ψ is correct. The value of ψ that represents the final estimate of the compliance-adjusted intervention effect is found by solving $G_0^*(\psi) = 0$.

Loeys and Goetghebeur (2003) show that $G_0^*(\psi) \sim N\{0, s(\psi)^2\}$ approximately, where $s(\psi)^2 = 2\sum_j \widehat{\Lambda}_0^*(T_j|\psi)$. Confidence limits for ψ are found by solving $G_0^*(\psi) = \pm z_{\text{crit}} s(\psi)$, where z_{crit} is the critical value for the appropriate significance level. In the Stata program, estimation of the point estimate and confidence limits is achieved using a loop employing interval bisection. First, the target value is set as 0, $-z_{\text{crit}}s$ or $+z_{\text{crit}}s$, depending whether the point estimate, lower confidence limit, or upper confidence limit is being estimated. Then, minimum and maximum values of ψ are initialized. At the start of each run of the loop, ψ is defined as the midpoint of the current minimum and

maximum values, and $G_0^*(\psi)$ is calculated. If $G_0^*(\psi)$ is greater than the target value, the minimum is reset to the value of ψ used in this run, or if it is less than the target, the maximum is reset to ψ. The loop is then run again, applying these new minimum and maximum values, unless the difference between them is less than a user-defined value (option `convcrit()`, default value 0.01).

In the presence of censoring, the method assumes that $S_{c1}(t)$ and $S_{n1}(t)$ are consistently estimated. This happens if censoring in the intervention arm is noninformative, conditional on compliance. The estimation procedure further assumes that censoring in the control arm is noninformative unconditionally. By contrast, Frangakis and Rubin (1999) have described methods that require censoring in both arms to be noninformative conditional on compliance, which they term "latent ignorability".

3 Description of stcomply

We have developed a command, `stcomply`, that produces the estimated compliance-adjusted intervention effect and confidence intervals from user input indicating the location of intervention and compliance data. The `stcomply` command is intended for use with survival data and should be preceded by the `stset` command.

3.1 Syntax

`stcomply` *group comply* [`if` *exp*] [`in` *range*] [`, data graph`[`(`*graph_options*`)`]
`grfit`[`(`*graph_options*`)`] `level(#) convcrit(#)`]

where *group* is the name of the variable holding the intervention assignment, and *comply* is the name of the variable holding the compliance data. Both variables must contain binary data (i.e., the command is valid for two-arm trials only and all-or-nothing compliance only), and compliance data must only apply to one arm of the trial (that with the higher value of group), with the higher value indicating compliance. Compliance must be set to missing for the other arm of the trial.

3.2 Options

`data` stores the values of the adjusted hazard ratios and the corresponding standard deviations from the observed hazard ratio, in matrix `psi_z`.

`graph`[`(`*graph_options*`)`] graphs the standardized test statistic $G_0^*(\psi)/s(\psi)$ against the hazard ratio ψ, where $G_0^*(\psi)$ is the difference between the number of deaths resulting from the predicted survival function for the control arm and the number of deaths actually observed in the control arm. Where this function does not intersect the upper/lower horizontal lines, a confidence interval for ψ cannot be calculated. *graph_options* are options of [G] **graph twoway line** that specify details for the graph.

`grfit`[`(`*graph_options*`)`] provides Kaplan–Meier plots of predicted (based on potential compliers and noncompliers) and observed survival in the control arm. *graph_options* are options of [G] **graph twoway line** that specify details for the graph.

`level(#)` specifies the confidence level, as a percentage, for the confidence intervals of the Loeys estimate. The default is the current value of `c(level)`.

`convcrit(#)` specifies the accuracy of convergence of ψ to the estimate and confidence limits. The default is within 0.01.

4 Example: Application to a large, randomized trial

The Multicentre Aneurysm Screening Study Group (2002) (MASS) is a large, randomized trial involving 68,000 men aged 65–74 in five UK centers. Individuals were randomized to receive an invitation to attend ultrasound screening for abdominal aortic aneurysms (AAAs) or to the control group with whom no contact was made. Individuals with an aneurysm detected at screening were considered for elective surgery; without surgery, an AAA might grow and eventually rupture. A ruptured AAA requires higher risk emergency surgery to prevent death. The primary outcome of the trial was AAA-related mortality (including postoperative mortality), and the results showed a significant decrease in risk for those in the "invited to screening group" when analyzed by intention-to-treat (hazard ratio 0.58 [95% CI: 0.42, 0.78]).

Among those in the "invited to screening group", 80% attended an initial screening. In addition to the intention-to-treat analysis, the trial also reported an intervention effect for attenders (hazard ratio 0.47 [95% CI 0.36, 0.70]), which was calculated using the Loeys and Goetghebeur (2003) methodology described above. The compliance in the invited group was treated as all or nothing according to attendance at the initial scan; those failing to attend subsequent follow-up scans after attending an initial scan were considered to be compliers.

The example given here uses a random subset of the MASS data, containing 15% of the individuals ($n = 10123$) from the full analysis. This subset shows an estimated intention-to-treat intervention effect of 0.74 (95% CI 0.36, 1.51) and a compliance-adjusted intervention effect of 0.66 (95% CI 0.30, 1.74):

(*Continued on next page*)

```
. use mass, clear
(MASS trial subset, supplied as example with stcomply command)
. stset timeout, id(id) origin(dateran) fail(aaadeath) scale(365.25)
                id:  id
     failure event:  aaadeath != 0 & aaadeath < .
obs. time interval:  (timeout[_n-1], timeout]
 exit on or before:  failure
    t for analysis:  (time-origin)/365.25
            origin:  time dateran
--------------------------------------------------------------------------
    10123  total obs.
        0  exclusions
--------------------------------------------------------------------------
    10123  obs. remaining, representing
    10123  subjects
       31  failures in single failure-per-subject data
 39433.19  total analysis time at risk, at risk from t =         0
                             earliest observed entry t =         0
                                  last observed exit t =  5.185489
. stcox status
         failure _d:  aaadeath
   analysis time _t:  (timeout-origin)/365.25
             origin:  time dateran
                 id:  id
Iteration 0:   log likelihood = -278.71179
Iteration 1:   log likelihood = -278.36571
Iteration 2:   log likelihood = -278.36569
Refining estimates:
Iteration 0:   log likelihood = -278.36569
Cox regression -- no ties
No. of subjects =        10123                     Number of obs   =     10123
No. of failures =           31
Time at risk    =  39433.19097
                                                   LR chi2(1)      =      0.69
Log likelihood  =   -278.36569                     Prob > chi2     =    0.4054
------------------------------------------------------------------------------
          _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      status |   .7401218   .2693876    -0.83   0.408     .3626478    1.510502
------------------------------------------------------------------------------
. stcomply status scanned, graph grfit
Loeys-Goetghebeur estimates of effect of treatment actually received
--------------------------------------------------------------------------
Estimate of effect adjusted for compliance = 0.6609
95% confidence interval = 0.3035 , 1.7357
--------------------------------------------------------------------------
```

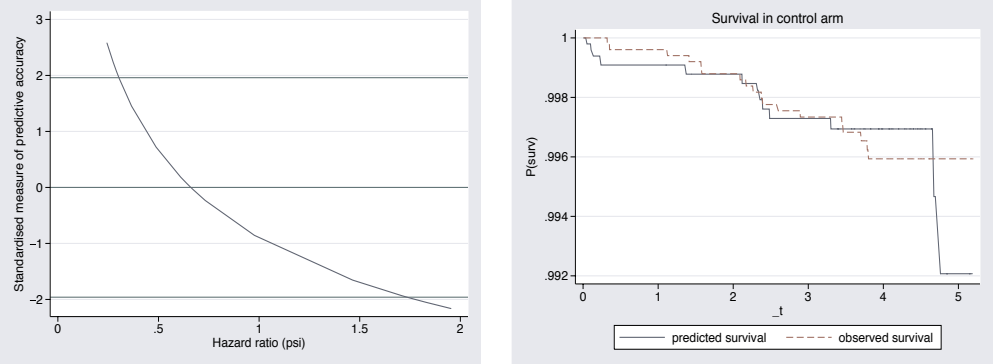

Figure 1: (Left) Plot of test statistic against ψ for the MASS dataset, produced by the `graph` option of `stcomply` command. (Right) Kaplan–Meier plots of predicted and observed survival in the control arm for the MASS dataset, produced by the `grfit` option.

`timeout` gives the event/censoring time, `dateran` the date of randomization into the MASS study, and `aaadeath` is the binary indicator of an event of interest (AAA-related death). The variable `status` holds the intervention allocation (1 = control, 2 = invited to screening), and `scanned` is the compliance indicator for those invited to screening (0 = unscreened, 1 = screened). Figure 1 shows the plot of z-values against values of ψ, as produced by the `graph` option of the `stcomply` command, and the Kaplan–Meier plots of predicted and observed survival in the control arm, as given by the `grfit` option.

5 A negative weighting method

We also describe a simple alternative method that may be implemented using standard Stata code and may easily be applied in other situations. For simplicity, we assume that the two arms have equal size. To compare compliers in the intervention arm with compliers in the control arm, noncompliers must be removed from the control arm. As before, this is achieved by assuming that the noncompliers in the control arm are comparable to the noncompliers in the intervention arm. Therefore each noncomplier in the intervention arm is subtracted from the control arm by including them in the control arm with weight (-1); see figure 2. When sample sizes n_0, n_1 in the two arms are unequal, the weight for noncompliers in the intervention arm is $(-n_0/n_1)$.

Intervention arm	**Control arm**
Noncompliers $(-n_0/n_1)$	Noncompliers (1)
Compliers (1)	Compliers (1)
Observed compliance	Potential compliance

Figure 2: Weightings applied in compliance-adjusted Cox regression (weightings shown in parentheses).

This approach may be implemented as follows. The `cox` command is used, since `stcox` does not allow `iweight`s. Confidence intervals must be computed by bootstrapping, and setting the version to 7 is required since Stata 8 does not allow bootstrapping a command involving weights. For the MASS dataset example,

```
. set seed 100
. gen wt=1
. replace wt=-5129/4994 if status==2 & scanned==0
(995 real changes made)
. gen exposed=status==2 & scanned==1
. cox timeout exposed [iw=wt], dead(aaadeath) nohr
Iteration 0:   log likelihood = -204.83606
Iteration 1:   log likelihood = -204.36678
Iteration 2:   log likelihood = -204.36666
Refining estimates:
Iteration 0:   log likelihood = -204.36666
Cox regression -- no ties
Entry time 0                                      Number of obs   =       8106
                                                  LR chi2(1)      =       0.94
                                                  Prob > chi2     =     0.3326
Log likelihood = -204.36666                       Pseudo R2       =     0.0023

------------------------------------------------------------------------------
     timeout |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     exposed |  -.4099577   .4278969    -0.96   0.338     -1.24862    .4287049
------------------------------------------------------------------------------
. version 7: bs "cox timeout exposed [iw=wt], dead(aaadeath)" "_b[exposed]", re
> ps(1000)
command:     cox timeout exposed [iw=wt], dead(aaadeath)
statistic:   _b[exposed]
(obs=10123)
```

```
Bootstrap statistics
Variable |  Reps   Observed       Bias    Std. Err.   [95% Conf. Interval]
---------+--------------------------------------------------------------
     bs1 |  1000  -.4099577   .6362536    20.48761   -40.61365  39.79374  (N)
         |                                           -1.475853  .7481149  (P)
         |                                           -1.468809  .8002678 (BC)
---------------------------------------------------------------------------
                                N = normal, P = percentile, BC = bias-corrected
```

The normal bootstrap interval is influenced considerably by a small number of outlying bootstrap estimates; it is therefore important to use the bias-corrected estimate of the confidence interval. This gives a hazard ratio of 0.66 (95% BC-CI 0.23, 2.23) for the treatment effect in compliers, which is comparable to the estimate provided by `stcomply` of 0.66 (95% CI 0.30, 1.74).

Considering separately the survival of potential compliers and noncompliers in the control arm, this method is valid under the assumption of latent ignorability.

6 References

Cuzick, J., R. Edwards, and N. Segnan. 1997. Adjusting for noncompliance and contamination in randomized clinical trials. *Statistics in Medicine* 16: 1017–1029.

Frangakis, C. E. and D. B. Rubin. 1999. Addressing complications of intention-to-treat analysis in the combined presence of all-or-none treatment-noncompliance and subsequent missing outcomes. *Biometrika* 86: 365–379.

Loeys, T. and E. Goetghebeur. 2003. A causal proportional hazards estimator for the effect of treatment actually received in a randomized trial with all-or-nothing compliance. *Biometrics* 59: 100–105.

Robins, J. M. and A. A. Tsiatis. 1991. Correcting for noncompliance in randomized trials using rank preserving structural failure time models. *Communications in Statistics Theory and Methods* 20: 2609–2631.

Sommer, A. and S. L. Zeger. 1991. On estimating efficacy from clinical trials. *Statistics in Medicine* 10: 45–52.

The Multicentre Aneurysm Screening Study Group. 2002. The Multicentre Aneurysm Screening Study (MASS) into the effect of abdominal aortic aneurysm screening on mortality in men: a randomized controlled trial. *Lancet* 360: 1531–1539.

White, I. R., S. Walker, and A. Babiker. 2002. strbee: Randomization-based efficacy estimator. *Stata Journal* 2(2): 140–150.

About the Authors

Lois Kim is a Research Associate at the MRC Biostatistics Unit, Cambridge, UK.

Ian White is a Senior Scientist at the MRC Biostatistics Unit, Cambridge, UK.