Sensitivity Analysis - Ovarian

Florian Stijven

29-3-2022

```
##
##       0   1
##   0 215   0
##   1  26 950
```

```
## [1] 200
```

```
## [1] 950
```

Recode the data is interval censored data. So we assume (possibly wrongly) that no patient survives more than 20 years, and that progression occurs within 10 years.

Introduction

The `Ovarian` data set contains the progression-free survival times (PFS), and overall survival times (OS) for 1192 ovarian cancer patients. One observation is left out because the PFS is longer than OS, which is not possible. Both PFS and OS can be censored. Due to the nature of the endpoints, PFS is always smaller than or equal to OS. For 200 out of the 1191 observations, the observed times for PFS and OS are equal. With the copula based approach, this atomicity is however ignored. Moreover, the copula based approach does not impose ordering restrictions. Depending on the data, this could lead to issues.

Also note that due to the nature of the endpoints, PFS and OS are by definition associated. Indeed, PFS is a composite endpoint combining time-to-progression and time-to-death, whichever comes first.

Several conventional parametric models were considered for modelling the marginal survival functions: Weibull, log-logistic and generalized gamma. However, these all fail to describe the data well. This is caused by the fact that in the data, the hazard is initially very high, but later becomes very small. This is seen on the marginal survival functions were intially there is a large drop, but after some time the survival functions become almost constant at about 0.1-0.2. Therefore, a Royston-Parmar spline model is fitted using a two-stage approach.
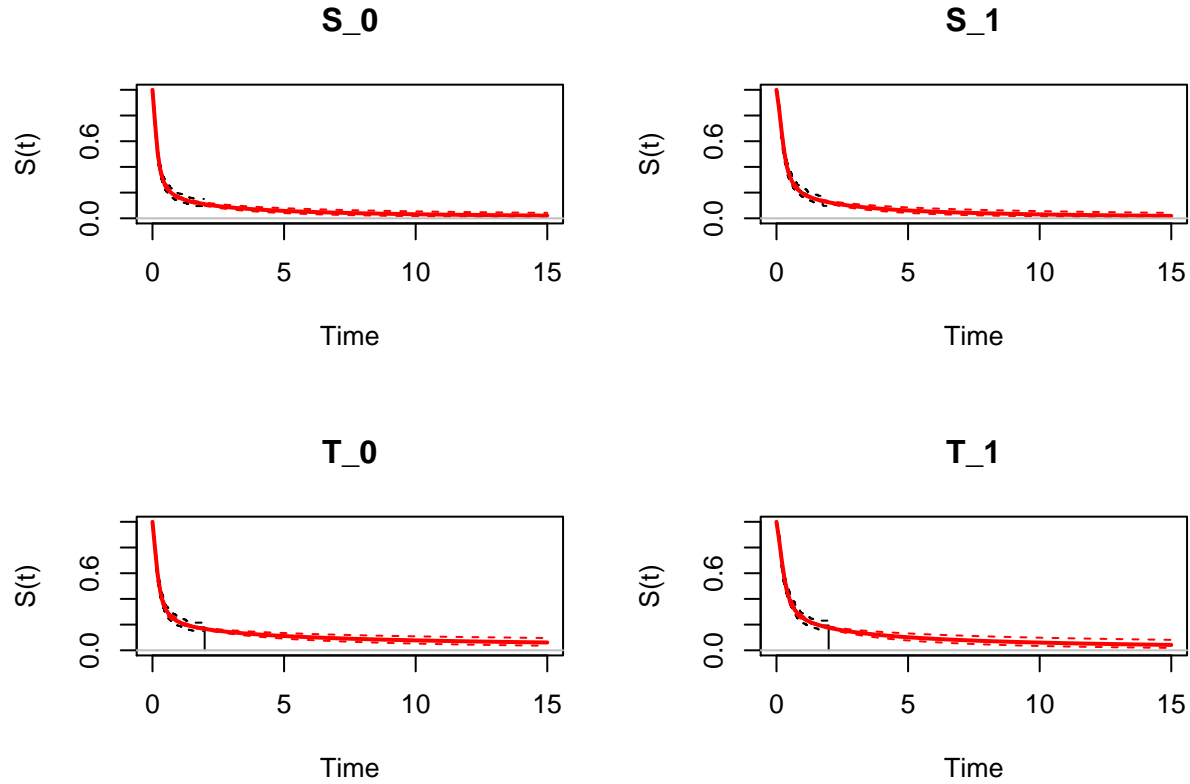
Model Fitting

Royston-Parmar Spline Model

The Royston-Parmar spline model is fitted for each potential outcome *separately*, so four times. A simplification would be fit it once for PFS and once for OS, and then include a treatment effect into the model. This however entails and additional PH-assumption which is not needed anywhere in the further approach.

All Royston-Parmar spline models are fitted with two knots using the `flexsurvspline` function from the `flexsurv` package. So each survival function is modelled by 4 parameters. As can be seen in the plots below, the KM-estimate and parametric estimate of the respective survival functions are very near, indicating that this parametric model is appropriate.

S_0

S_1

T_0

T_1
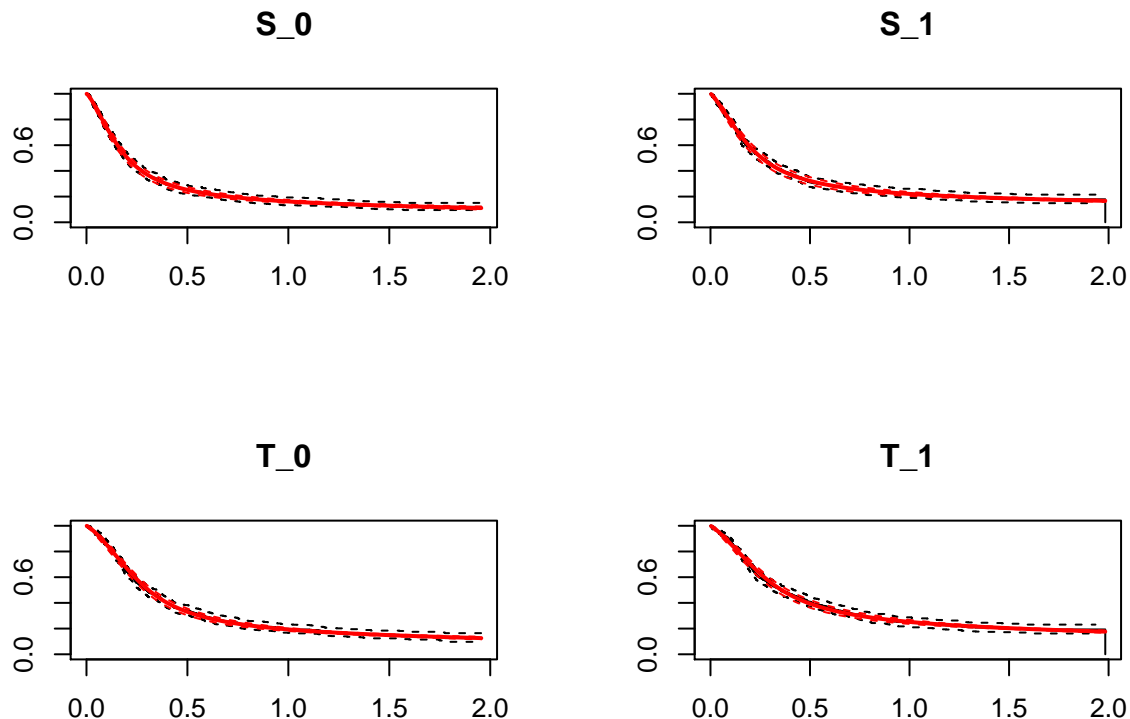
One particular issue with these data is that the maximum follow-up time is 2 years while a considerable proportion of patients survive past that point. The fitted parametric model however implies a survival and hazard function beyond two years. The survival and hazard function beyond those two years are however unobservable.

This is shown for S_0. The modelled survival function beyond the KM-estimate is implied by the model, but this is not verifiable. Still, this unverifiable part influences the metrics of surrogacy. However, it is not immediately clear how large this influence is.

S_0

S(t)

0.6

0.0

0   5   10   15

Time

S_1

S(t)

0.6

0.0

0   5   10   15

Time

T_0

S(t)

0.6

0.0

0   5   10   15

Time

T_1

S(t)

0.6

0.0

0   5   10   15

Time

Copula Model

The copula model is fitted in the second stage. To this end, the original PFS and OS times are transformed to uniform variables using the fitted survival functions. These uniform variables are then used to fit the copula model. Two parameters are fitted with this copula model: ρ_{S_0,T_0} and ρ_{S_1,T_1}. The other four correlation parameters are part of the sensitivity analysis.

```
## iter    2 value -204.304062
## iter    4 value -204.316913
## final  value -204.316913
## converged
```

```
## iter    2 value -228.440508
## iter    4 value -231.806467
## iter    6 value -231.806832
## final  value -231.806832
## converged
```

The observable correlations are 0.8851453 for the control group, and 0.9199036 for the treated group.

Goodness of fit

The goodness of fit is checked by sampling from the fitted model, and comparing this sample with the original data. However, the original data are censored which complicates a direct comparison. The distribution of

censoring indicators is therefore first estimated. Next, values are sampled from this distribution as to artificially censor the sample from the fitted model. In this way, the sample from the fitted model is comparable with the original data.

Fitted Model

Observed Data
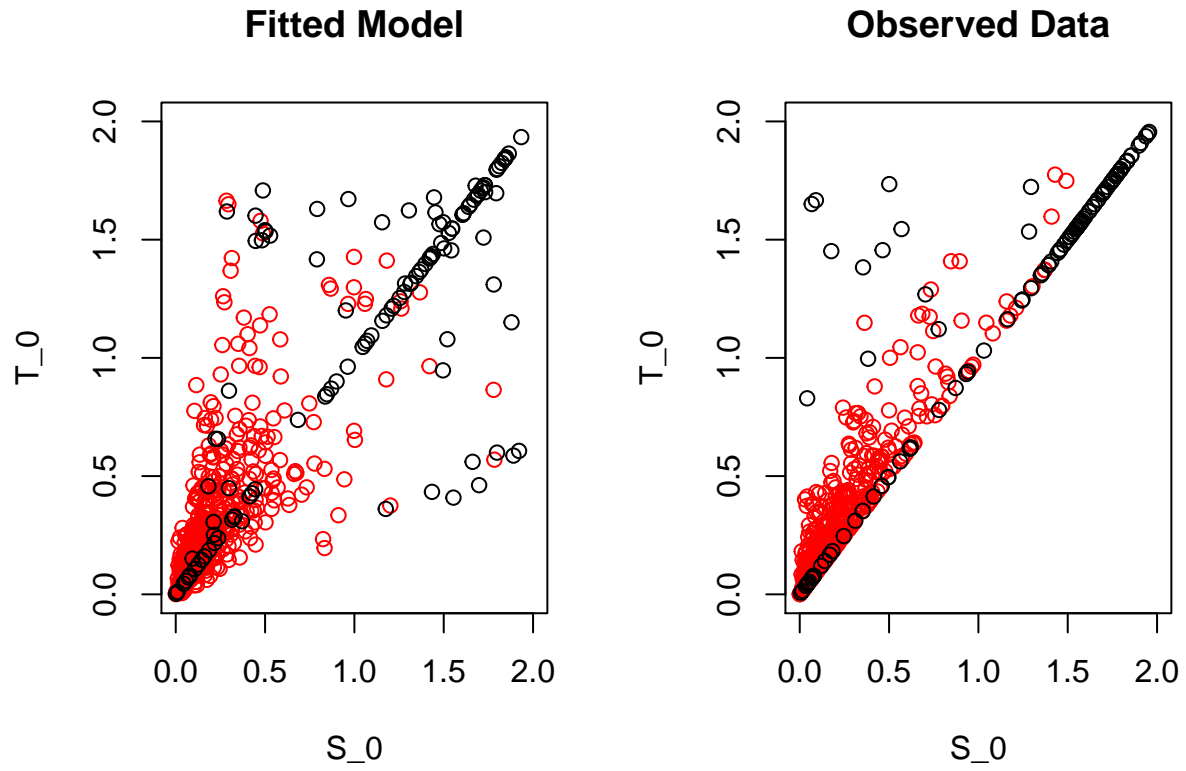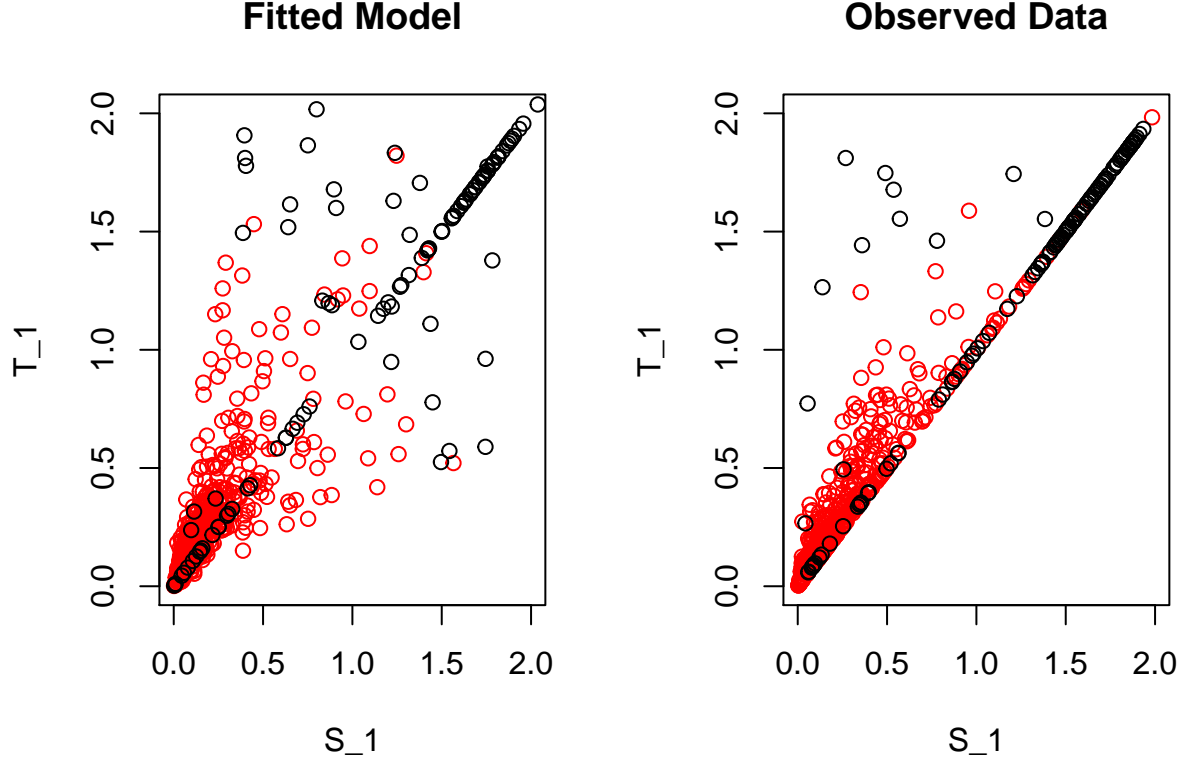
Sensitivity Analysis

The sensitivity analysis consists of sampling positive definite matrices. Given such a matrix, the measures of surrogacy can be derived. In this case, only ρ_Δ is considered as a measure of surrogacy because it easily follows from the normal copula parameters. ρ_Δ is defined as the Pearson correlation between the individual causal effects on the normal transformed variables of the surrogate and true endpoint, respectively: $\rho_\Delta = cor(q_{S_1} - q_{S_0}, q_{T_1} - q_{T_0})$ where $q_{S_0} = \Phi^{-1}(F_{S_0}(S_0)) \sim N(0,1)$ etc. This measures directly follows from the correlation parameters of the Gaussian copula:

$$\rho_\Delta = \frac{\rho_{S_0,T_0} + \rho_{S_1,T_1} - \rho_{S_0,T_1} - \rho_{S_1,T_0}}{2\sqrt{(1 - \rho_{T_0,T_1})(1 - \rho_{S_0,S_1})}}$$

The interpretation of the Pearson correlation between individual causal effects on the normal transformed variables is however not clear. Nonetheless, this measures closely corresponds to the information theoretic measures, which has a clear interpretation. The advantage of the former is that it directly follows from the fitted model, whereas the information theoretic measures requires numerical integration.

Sensitivity Analysis for rho_delta
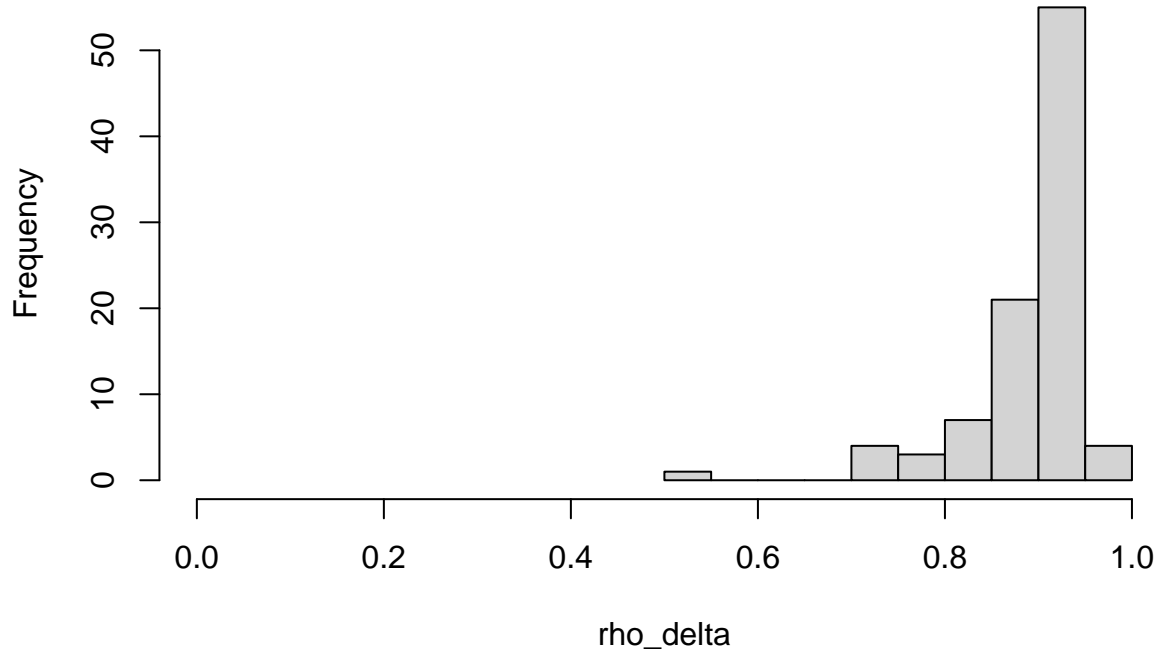
The values for ρ_Δ are mostly between 0.8 and 1. This is indicate of a strong association in most realities compatible with the observed data.

The information theoretic measure is the informational coefficient of correlation (ICC) between the individual causal effects on the surrogate and true endpoint, respectively: $R_h = 1 - \exp(-2 \cdot I(\Delta S, \Delta T))$. The mutual information requires solving the following integral:

$$I(\Delta S, \Delta T) = \int f(\Delta S, \Delta T) \log \left(\frac{f(\Delta S, \Delta T)}{f(\Delta S) \cdot f(\Delta T)} \right) d\Delta S d\Delta T$$

which is computed numerically by Monte Carlo integration.

Sensitivity Analysis for R_h

rho_delta versus R_h
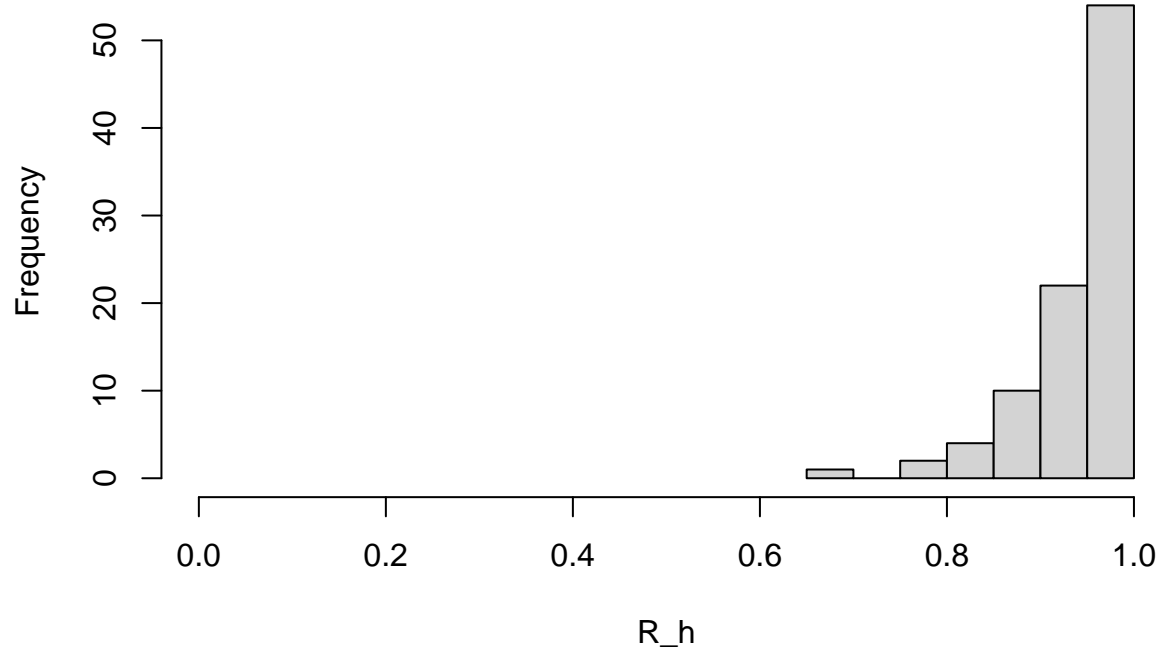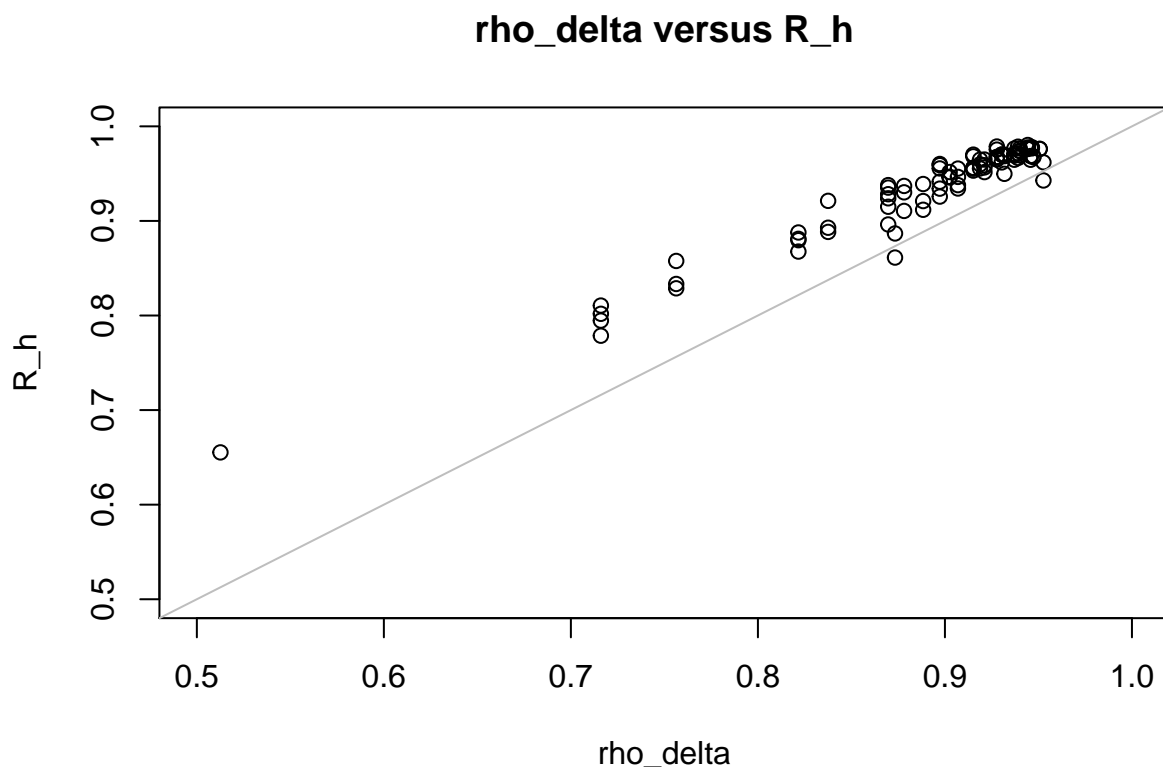
The qualitative conclusions for the strength of surrogacy based in the histogram of R_h is the same as for ρ_Δ. In the scatterplot, the two different measures of surrogacy are plotted against each other. The R_h tends to be larger, but there is a relatively close correspondence between both measures.

The histogram clearly indicates that across almost all realities compatible with the data, the ICA is very large (for both definitions of ICA). This indicates that PFS is a good surrogate for OS for the types of treatments and types of patients included in the data set. The histogram is based on 2085 positive definite matrices.

Use of interval Censoring

The same analysis is repeated, but now treating the right-censored observations as interval-censored data. Specifically, it is assumed that all patients progress before 15 years, and die before 20 years. These 2 values should however themselves be part of a sensitivity analysis. Moreover, input from literature might be helpful to determine whether such an assumption is plausible. One particular concern is that some patients might be cured. Progression would never occur in these patients, and death possibly only after a very long period.

The interval censored data is now fitted with 3 knots instead of 2. This gives generally a better fit in terms of the AIC.

These plots show how the survival function is modeled beyond the observable time period. With the data recoded into interval censored data, the long term survival probabilities are reduced by a lot as compared to the original analysis. Note however, that the correspondence to the KM-estimate is slightly off as compared to the original analysis.

S_0

S_1

T_0

T_1
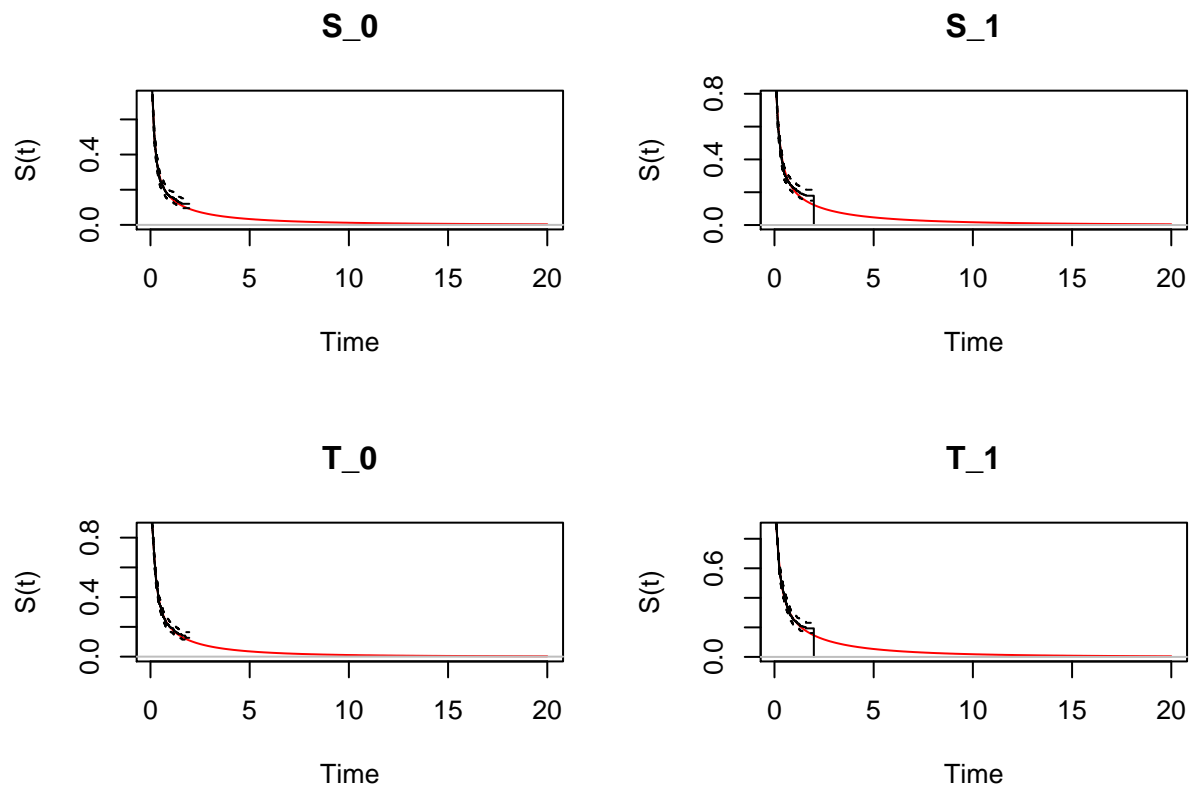

```
## iter    2 value -141.937865
## iter    4 value -141.939990
## final  value -141.939990
## converged

## iter    2 value -102.709533
## iter    4 value -102.769172
## final  value -102.769172
## converged
```
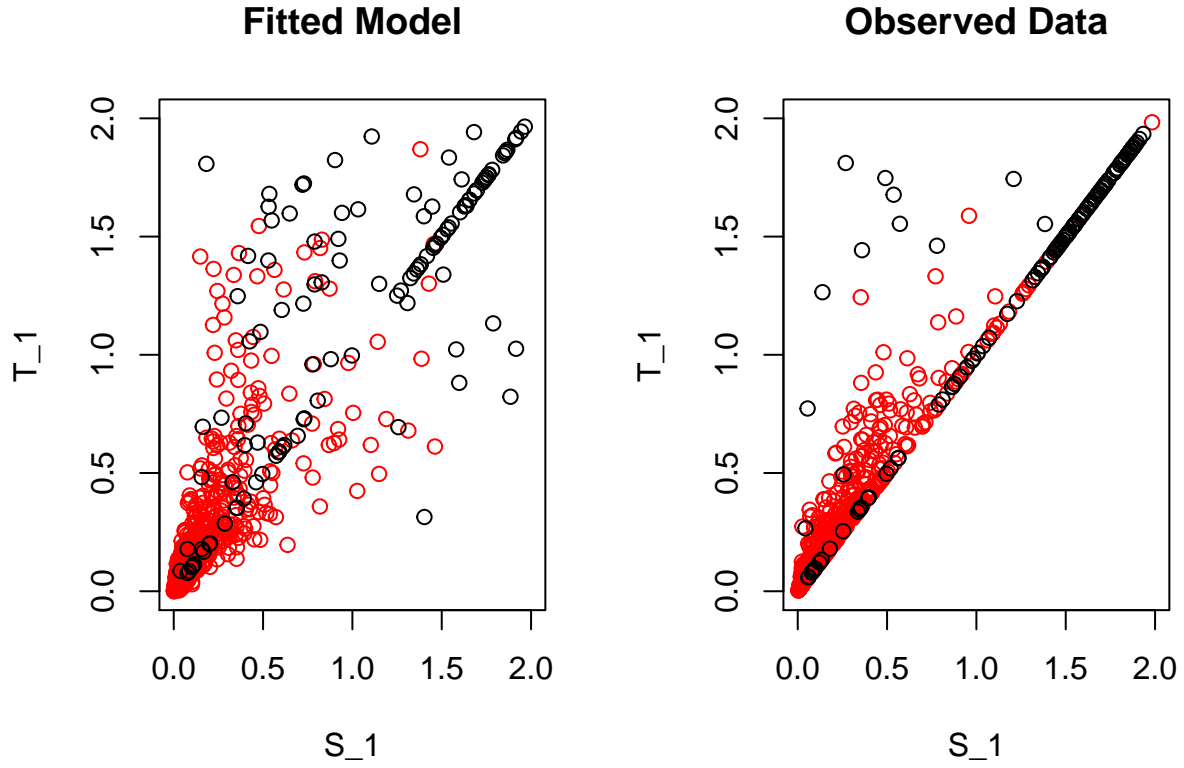
The observable correlations are 0.873929 for the control group, and 0.9005258 for the treated group. This is slightly smaller as when the data is treated as right-censored data.

Goodness of fit

The goodness of fit is checked by sampling from the fitted model as was done before.

Fitted Model   **Observed Data**
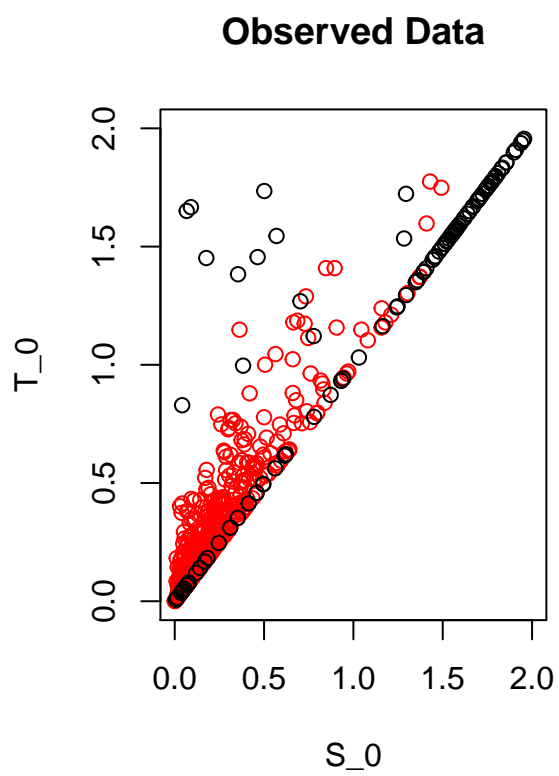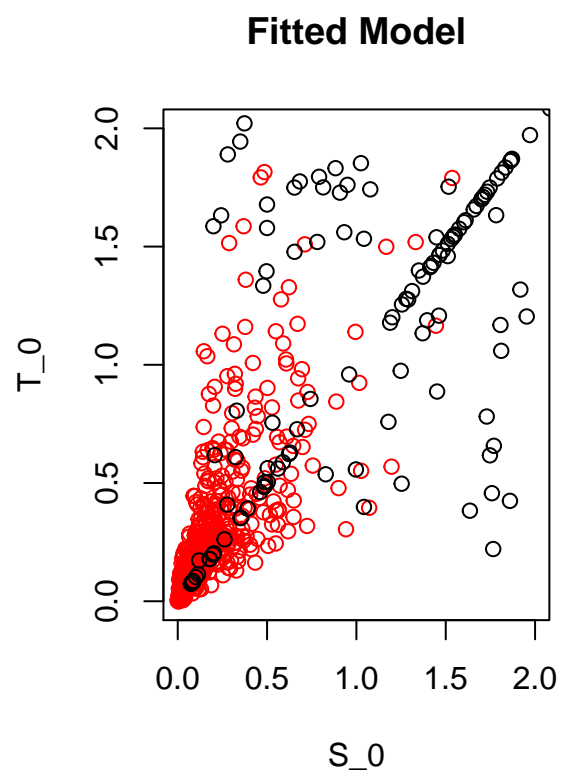
The same conclusions holds as when the data are treated as right-censored data.

Sensitivity Analysis

The results of the sensitivity analysis of the original right-censored data, and artificially interval-censored data are compared. Because the observable correlations are slightly smaller for the second case, there are more positive definite matrices sampled.

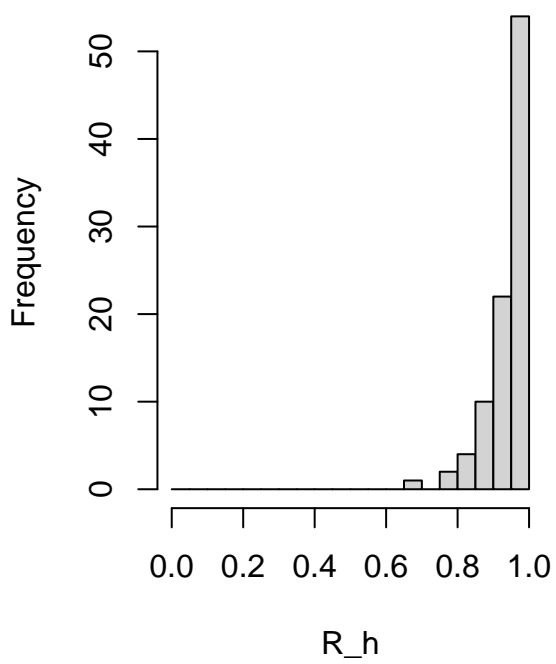There is a limited influence of the interval censoring on the distribution of ρ_Δ.

Interval Censoring

rho_delta

Right Censoring

rho_delta
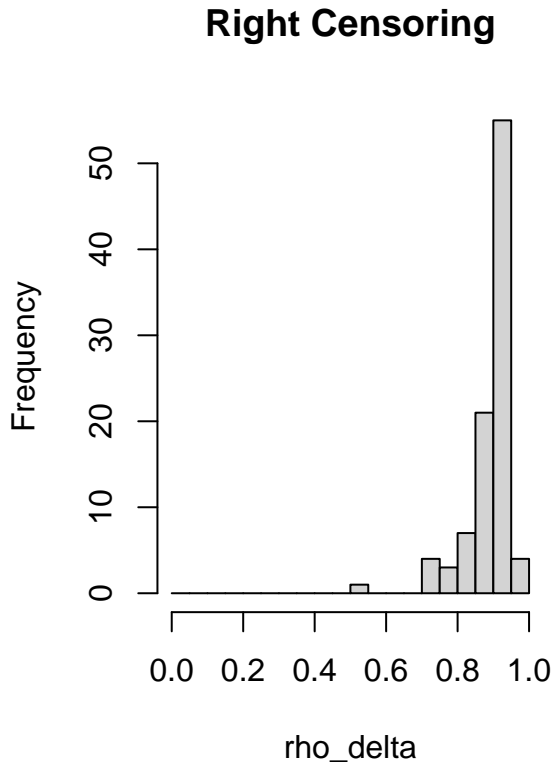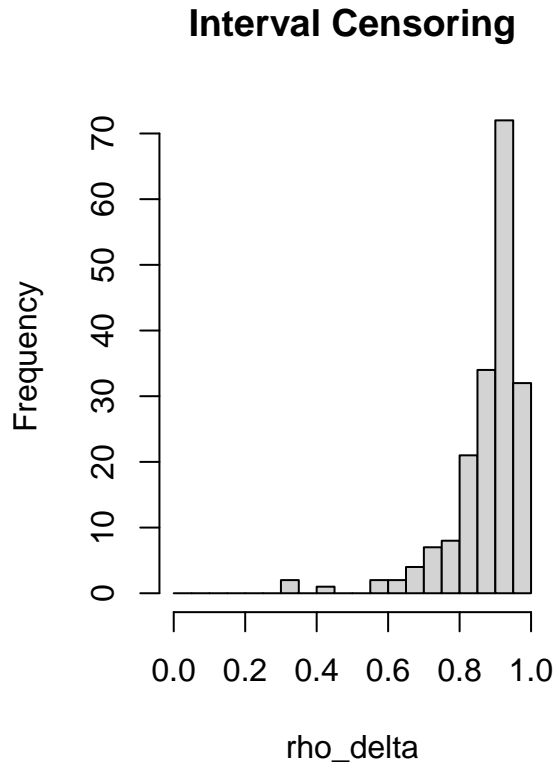
There is again a limited influence of the interval censoring on the distribution of R_h; of the same magnitude as for ρ_Δ. The same holds for the relation between R_h and ρ_Δ in both situations.
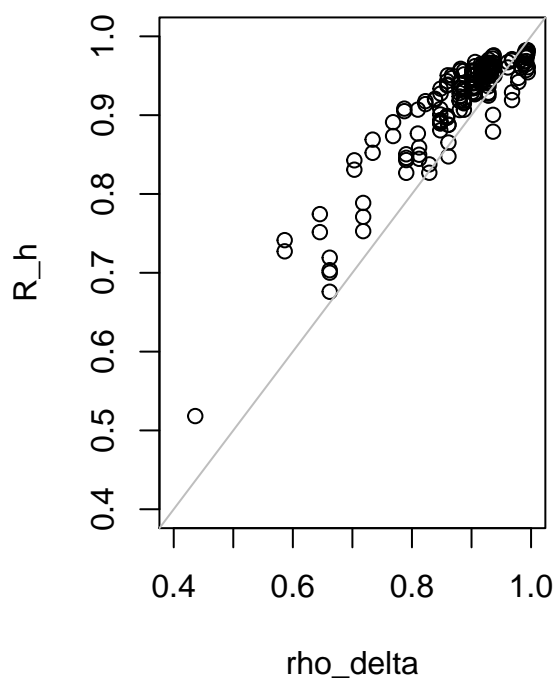
Interval Censoring

Frequency

R_h

Right Censoring

Frequency

R_h
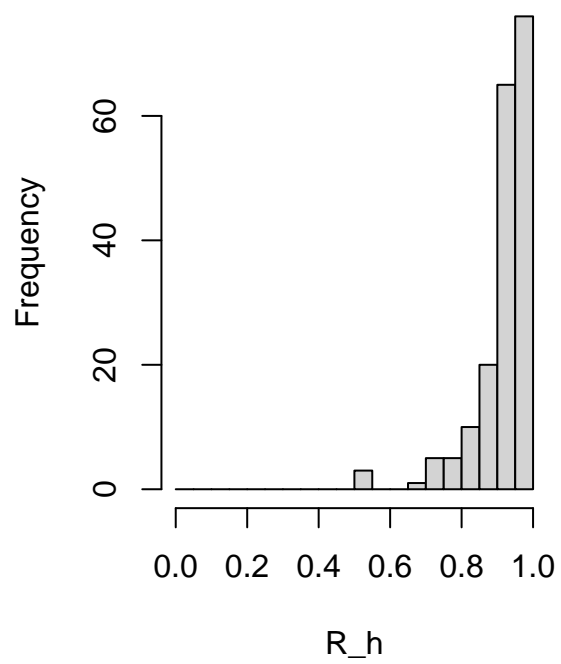
Interval Censoring

Right Censoring