

RMST: Restricted Mean Survival Time Methods

Jonathan Wessen (jonathan.wessen@astrazeneca.com)

In randomised clinical trials with a right-censored time-to-event outcome the hazard ratio is often used to compare the efficacy of two or more treatments, however the hazard ratio is only valid when the proportional hazards (PH) assumption is satisfied. The absence of PH means that the hazard ratio between two treatments will vary depending on the time. One solution to this might be to consider an average hazard ratio, however its results can be misleading[1] and therefore are potentially not a useful summary statistic.

The restricted mean survival time has been proposed as a better summary statistic for when the PH assumption has been violated; Royston P et al.[1] provide evidence in favour of its use as a primary endpoint in the absence of PH, and also suggest that it could be a useful secondary measure even when the PH assumption is satisfied. The advantages of RMST over traditional metrics, such as hazard ratio, include the following:

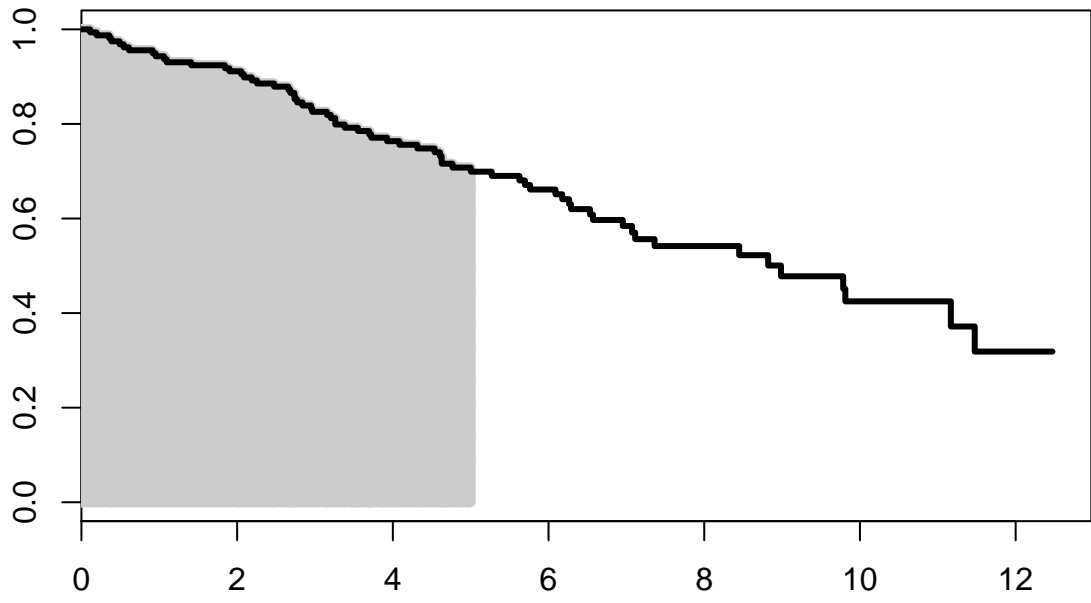
1. Its interpretation is straightforward and intuitive
2. The entire survival distribution up till the chosen time point is considered by integrating the survival function
3. Structural assumptions (PH) are minimal

The disadvantages however are:

1. The conclusion drawn from RMST can vary depending on the time point specified for analysis in the case of non-proportional hazards
2. Therefore, the time interval must be clinically motivated and prespecified in the trial protocol for use as a primary endpoint

There are three different approaches to calculating the RMST provided by this package. Each method takes a different approach at calculating the RMST and have their own merits. The RMST is defined as a measure of average survival from time 0 to a specific time point, and can be estimated by taking the area under the survival curve up to that point. The example data used in this document is contained within the `survRM2` package and can be called using `“rmst2.sample.data()”`. The Kaplan-Meier curve in the figure on the next page illustrates this concept by shading in the area under the curve up till the time point 5; this also represents one potential way to calculate the RMST.

For each method it is possible to specify the truncation time and the alpha level. The truncation time is the time in which to restrict the calculate of the mean to, and the alpha level controls what confidence intervals are reported (95% confidence intervals are reported by default).



Area under the Kaplan-Meier curve

The simplest implementation of RMST is to take the area under the Kaplan-Meier curve. The method “KM” in the `rmst` function, or the `rmstKM` function can be used to calculate this; these functions use the `rmst2` function from the package `survRM2` to calculate the area under the KM curve.

```
library(azRMST)
D <- rmst2.sample.data()
# Set the number of digits to be shown in outputs to be 3
options(digits = 3)
```

```
fit <- rmst(Surv(time,status) ~ arm, data=D, trunc=5, alpha=0.05, method="KM")
fit[[1]]
```

```
##           RMST      SE Lower CI Upper CI
## 1 arm=0 4.18 0.119      3.95      4.42
## 2 arm=1 4.30 0.106      4.09      4.51
```

The `azRMST` package calculates the RMST along with the standard error and 95% confidence interval. The between-group contrast can be seen with the command

```
fit[[2]]

##   Difference      SE Lower.CI Upper.CI p.value
## 1          0.12 0.159   -0.193   0.432  0.453
```

Pseudo-values

Pseudo-values as described by Anderson P et al.[2] can be used to investigate RMST. The idea behind Pseudovalues in this context is that if the data were complete, $f(X_i)$ would be observed for each individual i and the expected value $E(f(X))$ could be estimated by $\frac{1}{n} \sum_i f(X_i)$. In the case of censoring not all $f(X_i)$ are observed, however a well-behaved estimator, $\hat{\theta}$, for the expectation

$$\theta = E(f(X))$$

is available. The Pseudovalue for $f(X)$ for individual i where $i = 1, \dots, n$ is defined as

$$\hat{\theta}_i = n \cdot \hat{\theta} - (n - 1) \cdot \theta^{\hat{i}},$$

where $\theta^{\hat{i}}$ is the estimator applied to the sample of size $n - 1$ obtained by eliminating the i -th individual from the data set. The i -th psuedovalue can be viewed as the contribution of the individual i to the $E(f(X))$ estimate on the sample size n .

The pseudovalues $\hat{\theta}_i$ will then be used for all n subjects in place of the real data which has incomplete observations. Further information on these pseudovalues is given by Anderson P et al[2].

The result of the above is that the pseudovalues are constructed such that their mean is an estimate of the restricted mean survival time at t^* . They are approximately unbiased for each individual and therefore unbiased for the overall mean. They are a non-parametric method computed from the Kaplan-Meier estimate of the survival curve. The `rmst` function with the “pseudo” method, or the `rmstPseudo` function, can be used to calculate an estimate of the RMST using this method. The package `pseudo` is used to calculate the restricted pseudovalues for each subject using the command `pseudomean`; the RMST is then calculated as the mean of these pseudovalues.

```
fit <- rmst(Surv(time,status) ~ arm, data=D, trunc=5, alpha=0.05, method="pseudo")
fit[[1]]
```

```
##           RMST      SE Lower CI Upper CI
## 1 arm=0 4.18 0.119      3.95      4.42
## 2 arm=1 4.30 0.106      4.09      4.51
```

The standard errors are calculated using a Generalised Estimating Equation (GEE) using a robust sandwich estimator for the standard error of the parameters. The GEE is fit using the `geese` command from the package `geepack`. The between-group contrast and standard error is also calculated in the same model and can be seen as follows,

```
fit[[2]]
```

```
##   Difference      SE Lower.CI Upper.CI p.value
## 1           0.12 0.16   -0.193    0.432  0.453
```

Royston-Parmar model fitting

The Royston-Parmar model is a flexible parametric method of fitting a survival model to data and is described in the paper by Royston, P. and Parmar, K.B.M. [1]. The distribution of the integrated hazard is not assumed to follow a specific distribution. Instead, the baseline function is modelled using splines which allows for a more flexible and often more accurate fit. Time-dependent covariates can be introduced when the proportional hazards assumption is violated to allow the hazard function for each group to differ.

The `rmst` function with the “RP” method, or the `rmstRP` function, can be used to calculate an estimate of the RMST using this method. This method uses the function `flexsurvspline` from the package `flexsurv` to fit the model, and then builds upon this package with the Delta method to calculate estimates of the standard error. The Royston-Parmar method introduces an additional parameter, the number of knots, which can be passed to the command. This parameter controls the number of internal spline knots.

```
fit <- rmst(Surv(time,status) ~ arm, data=D, trunc=5, alpha=0.05, method="RP", knots=2)
fit[[1]]
```

```
##           RMST      SE Lower CI Upper CI
## 1 arm=0 4.25 0.0985      4.05      4.44
## 2 arm=1 4.28 0.0982      4.09      4.47
```

The standard errors are calculated using the Delta method on the Royston-Parmar model. The between-group contrast and standard error is also calculated using the Delta method, and can be display as follows,

```
fit[[2]]
```

```
##   Difference      SE Lower.CI Upper.CI p.value
## 1      0.0332 0.118    -0.198    0.265    0.779
```

Delta method

The Delta method provides a much quicker way to estimate the standard error compared to bootstrapping an estimate. It utilises the first-order term in a Taylor series expansion to estimate the variance of a function. To approximate the variance of a multi-variable function, first consider the vector of partial derivatives of the function with respect to each parameter in turn evaluated at the point of interest, right-multiply this vector by the variance-covariance matrix, and right-multiply the resulting product by the transpose of the original vector of partial derivatives [5][6].

This method can be applied to a parametric model such as the Royston-Parmar model as follows. First define a function which calculates the RMST as a function of the coefficients of your model. Then fit the model and extract the information required such as the coefficients of the model, the variance-covariance matrix etc... Following this, calculate the gradient of the RMST function at the points defined by the coefficients of the model between the interval of interest, in this case the RMST is calculated between 0 and 5. These values can then be used to estimate the standard error of the RMST between 0 and 5 by taking the square root of GVG^T where G is a matrix of the gradients calculated, and V is the variance-covariance matrix.

p-values

The p-value for the null hypothesis of no difference can be calculated by taking the square of the difference in RMST divided by the standard error and then comparing this test statistic to a chi-square distribution with 1 degree of freedom [7]; the following is the equation used to implement this,

$$\gamma = \frac{\text{Estimate}}{\text{SE}^2} \sim \chi_1^2$$

$$\text{p-value} = 1 - \gamma$$

Acknowledgements

This package was written and developed as part of an AstraZeneca workstream.

References

- [1] Royston, P., Parmar, K.B.M., “The use of restricted mean survival time to estimate the treatment effect in randomized clinical trials when the proportional hazards assumption is in doubt.” *Statistics in medicine*, 30(19), 2409-2421, May 2011.
- [2] Anderson P.K, Hansen, M.G., Klein, J.P., “Regression analysis of restricted mean survival time based on Pseudo-observations.” *Lifetime Data Anal*, 10(4), 335-350, Dec 2004.
- [3] Klein, J.P., Gerster, M., Andersen, P.K, “SAS and R functions to compute Pseudo-values for censored data regression.” *Computer Methods and Programs in Biomedicine*, 89(3), 289-300, March 2008.
- [4] Dewar, R., Khan, I., “A new SAS Macro for flexible parametric survival modelling: Applications to clinical trials and surveillance data.” *Computer Methods and Programs in Biomedicine*, Dec 2015.
- [5] Cooch, G., White, G., “Program MARK: A Gentle Introduction.”, Appendix B, *14th edition*.
- [6] Lui, X., “Methods and Applications of Longitudinal Data Analysis.”, Appendix B, *Elsevier, 2015*.
- [7] Royston, P., Parmar, M., “Augmenting the logrank test in the design of clinical trials in which non-proportional hazards of the treatment effect may be anticipated.” *BMC Med Res Methodol.*, 16(16), 2016.