Following Royston and Parmar [21], the RMST, μ(τ), of a random variable T is the mean of the survival time X = min (T, τ) limited to some horizon τ > 0. It equals the area under the survival curve S (t) from 0 to τ:

$$\mu(\tau) = E(X) = E[\min(T, \tau)] = \int_0^\tau S(t)dt$$

When T is years to death, we may think of μ(τ) as the 'τ -year life expectancy'. The variance, var(X), of the restricted survival time X, is calculated as

$$var(X) = RSDST^2 = E(X^2) - [E(X)]^2 = 2\int_0^\tau tS(t)dt - \left[\int_0^\tau S(t)dt\right]^2$$

The restricted standard deviation (RSDST) is $\sqrt{var(X)}$. In a two-arm clinical trial with survival functions S_0(t) and S_1(t) for the control and treatment arms, respectively, the difference in RMST between arms (treatment – control), Δ(τ), is given by

$$\Delta(\tau) = \int_0^\tau S_1(t)dt - \int_0^\tau S_0(t)dt = \int_0^\tau (S_1(t) - S_0(t))\,dt$$

i.e., Δ(τ) is the area between the survival curves. The delta method can be used to calculate the variance $var(\Delta(\tau))$ and then construct the confidence interval for the RMST difference Δ(τ) using the normal approximation. The treatment effect estimation that corresponds to the RMST based test can be performed by constructing a pointwise two-sided 1-α interval based on the standard normal approximation as

$$\Delta(\tau) \pm z_{1-\alpha/2}\sqrt{var(\Delta(\tau))}$$

where $z_{1-\alpha/2}$ is the 100(1−α/2)-th percentile of the standard normal distribution. Note that a simultaneous confidence interval for RMST differences may also be constructed using a similar procedure proposed by Zhao et al. [18].

The ratio of RMST between arms, $\theta(\tau)$, is given by

$$\theta(\tau) = \frac{\int_0^\tau S_1(t)dt}{\int_0^\tau S_0(t)dt}$$

Similarly, the confidence interval for $\theta(\tau)$ can be constructed by calculating the confidence interval for $\log(\theta(\tau))$ using delta method and then transforming back. In theory, a simple parametric model such as the Weibull, or Gompertz, or three parameter Gamma can be used to estimate RMST. However, a simple parametric model may not well fit the complex survival curve often seen in recent IO development. Several methods of estimating RMST are available [22] and discussed in literature, including direct integration of Kaplan-Meier survival curves, a jackknife method, and the Royston and Parmar's modelling using a smoothing spline for the log-hazard function, and more recently the trapezoidal rule approach [23]. Note that KM uses a step hazard rate function. As pointed out in Royston and Parmar's paper, the direct integration of Kaplan-Meier curves may be unreliable. The jackknife method has the advantage of being non-parametric but the drawback of being relatively slow to compute, which makes it cumbersome when simulation with many replicates is needed. Tian's presentation [24] indicated that the performance of these RMST methods depends on the selection of τ. However, pre-specification of τ is not always easy. Using a large τ may not always be better because the separation of survival curves may not increase over time in long-term trials. In addition, a KM survival curve is not defined beyond the largest follow-up time. Therefore, in applications we can only pre-specify the τ as the minimum of the longest follow-up times for treatment groups, which will be data dependent.

To avoid these difficulties, we consider a flexible survival function as defined from the mixture of three components of Weibull [20]

$$S(t) = p_1 \exp\left[-\left(\frac{t}{\lambda_1}\right)^{k_1}\right] + p_2 exp\left[-\left(\frac{t}{\lambda_2}\right)^{k_2}\right] + (1 - p_1 - p_2) exp\left[-\left(\frac{t}{\lambda_3}\right)^{k_3}\right]$$

Liao and Liu [20] have demonstrated that the mixture model with 3 components of Weibull distribution fulfills the needs for modeling the delay effect or survival sudden drops, cure rate, or long term survival which are often observed in recent IO development. The advantage of using the mixture Weibull models includes: a) it is flexible and can produce a survive curve almost the same as the KM fitting; b) it is fully parametric which allows predicting future events, survival probability, and hazard function. In addition, the estimated hazards and survival curves are smooth functions as compared to the step functions from the Cox model or nonparametric estimators such as the KM method. In real applications, if a prior knowledge is available for the number of subgroups/components based on composition of the study population, then this knowledge should be used to determine the number of components for the mixture distribution. Huang, et. al. [25] used this mixture model to estimate the progression free survival (PFS) and overall survival (OS) - functions based on an interim dataset and showed that the models can predict the final PFS and OS survival curves very well.

With the parametric components, the RMST across the entire time space, i.e., $(0, \infty)$, can be estimated directly from the mixture Weibull estimates using $\int_0^\infty e^{-ax^b}\,dx = \frac{1}{b}a^{-\frac{1}{b}}\Gamma\left(\frac{1}{b}\right)$ and $\int_0^\infty x^n e^{-ax^b}\,dx = \frac{1}{b}a^{-\frac{n+1}{b}}\Gamma\left(\frac{n+1}{b}\right)$, where $\Gamma(p)$ is a gamma function. In fact, the RMST $\mu(\tau)$ can be estimated for any given τ using the incomplete gamma function $\gamma(s, x) = \int_0^x t^{s-1}e^{-t}\,dt$, which can be computed using the r-function. The standard error and pointwise confidence intervals for $\mu(\tau)$ can also be obtained from delta method.

In applications, the estimated $\mu(\tau)$ and its pointwise confidence intervals provide dynamic views for the treatment effects over a different time window τ. As compared to the KM-based RMST analysis, this dynamic RMST analysis provides several advantages such as 1) a straightforward estimate calculation, 2)

a better control on variability since the KM method may have large variance towards the tail, 3) easy implementation when adjusted for covariates, and 4) the availability of using RMST across the entire time space. Specifically, the last point of advantages would be useful for checking whether the follow-up time is long enough to demonstrate a treatment difference or to reach the maximum of the treatment effect (e.g., the estimated difference can be better later if follow up is longer) by checking the direction of the dynamic RMST difference or RMST ratio for a stabilized treatment effect; and it is useful for determining a time point for interim analysis by picking the time where the dynamic RMST difference or RMST ratio crossing a pre-defined acceptable treatment effect or for predicting future trends. In general, the timepoint selected for the analysis is very critical and is not a purely statistical issue. It should be chosen based on clinical consideration for the treatment and disease area such as how long we should follow the patients (for example, 3 years) to obtain enough information to assess the treatment benefit and harm. The tools mentioned here can be used to help understanding how feasible and reasonable of this choice. This piece of information will be explored in the three real datasets in the next section.

It should be noted that although the RMST curves based on the mixture models can be calculated over the entire time space in theory but we may not want to extend the estimated RMST too far away from the study follow-up to avoid too much extrapolation. Huang, et. al. [25] extrapolated the PFS and OS survival curves using the mixture model based on the interim data and predicted the final PFS and OS survival curves very successfully. Based on the data maturity on which the mixture model was built, the amount of extrapolation can be varied. As pointed out in Liao and Liu [20], some practical guidelines are provided by Pocock et al. [26] and Gebski et al. [27] about curtailing the KM plot when the risk set is too small in applications. We apply the methods to a few real oncology trials in the next section to illustrate the flexibility and advantage of the mixture-model-based RMST analysis as compared to KM-based RMST.