Journal of Biostatistics and Biometric Applications
Volume 10 | Issue 1
ISSN: 2455-765X

Research Article Open Access

The Win Ratio and Novel Continuous Win Ratio and Win Difference Statistics with a Focus on Oncology Trials with Two Prioritized Outcomes

Revathi Ananthakrishnan[1,*], Yodit Seifu[1], Zhi Yang[2], and Michael LaValley[3]

[1]GBDS, Bristol-Myers Squibb, 5 Giralda Farms, Madison, NJ, 07940, United States of America
[2]Bristol-Myers Squibb, Seattle, Washington, 98109, United States of America
[3]Department of Biostatistics, Boston University School of Public Health, Boston, MA, 02118, United States of America

[*]**Corresponding Author:** Revathi Ananthakrishnan, GBDS, Bristol-Myers Squibb, 5 Giralda Farms, Madison, NJ, 07940, United States of America, E-mail: revathia@gmail.com

Citation: Revathi Ananthakrishnan, Yodit Seifu, Zhi Yang, Michael LaValley et al. (2025) The Win Ratio and Novel Continuous Win Ratio and Win Difference Statistics with a Focus on Oncology Trials with Two Prioritized Outcomes, J Biostat Biometric App 10(1): 103

Received Date: April 14, 2025 **Accepted Date:** May 14, 2025 **Published Date:** May 19, 2025

Abstract

The win ratio is an estimate of the treatment effect used in situations where composite endpoints are employed and where the events that make up the composite endpoint can be prioritized. We study the win ratio in the context of oncology trials with the events of progression and death. We then propose two new measures related to the win ratio, called the continuous win ratio and continuous win difference. We define the continuous win ratio to be the ratio of the length of time the treatment arm wins by when it wins to the length of time it loses by when it loses, while the continuous win difference is the difference between these two quantities scaled by the number of pairs. The win ratio, continuous win ratio, continuous win difference, net benefit, and win odds are compared using simulations in the setting where death is considered the highest-priority or primary event, and progression is considered the secondary event, and when the priorities are switched. We also estimate the restricted mean survival time (RMST) difference and ratio and pairwise win time for various scenarios. Using exponential and Weibull distributions to simulate times to progression and death, we consider both the proportional and non-proportional hazards situations. Finally, we present an oncology case study to estimate the win ratio, continuous win ratio, continuous win difference, net benefit, win odds, RMST difference and ratio and pairwise win time. The continuous win ratio, continuous win difference and RMST difference and pairwise win time are measures of how much benefit in terms of time the study drug provides to the patient compared to the control, and thus may help doctors and payers better understand the effect of the study drug.

Keywords: Win Ratio; Continuous Win Ratio; Continuous Win Difference; Win Odds; Net Benefit; RMST Difference and Ratio; Pairwise Win Time

Introduction

Progression free survival (PFS) is a common endpoint in late phase oncology clinical trials. It is a composite endpoint of progression or death (defined as the earliest of progression (PD) or death). The log-rank test is generally used to determine whether the progression-free survival curves of the treatment arm and the control arm are different. The log rank test is optimal under the proportional hazards (PH) assumption, which means that the hazard ratio (HR) for treatment on PFS does not change over time.

The win ratio (WR) is an alternate estimate of the treatment effect in situations where composite endpoints are employed. The win ratio estimates the ratio of the probability that a randomly chosen treatment arm patient has a better outcome ("winning") than a randomly chosen control arm patient, to the probability that the control arm patient has a winning outcome over the treatment arm patient. In the win ratio approach, the events that make up the composite endpoint can be prioritized and evaluated sequentially from highest to lowest priority to determine the winner. Under PH, the win ratio is equal to the reciprocal of the HR for a time to event endpoint with a single event [1, 2].

In this paper, we consider the win ratio in the context of oncology trials with the events of progression and death. We then propose two new measures based on the win ratio, called the continuous win ratio (CWR) and continuous win difference (CWD). The rationale for these new measures is to not only address the likelihood of treatment arm patients winning, but also to quantify the treatment benefit in terms of time, considering progression or death as the primary event. The CWR provides a ratio of how much time the treatment arm gains when it wins to how much time it loses when it loses. The continuous win difference is the difference between how much time the treatment arm gains when it wins to how much time it loses when it loses, scaled by the number of pairs. The win ratio, the CWR and CWD are compared in the setting where death is considered the highest-priority or primary event, and progression is considered the secondary event, and in the setting where progression is considered the primary event and death is the secondary event (Table 1).

Table 1: Pair Outcomes with Two Endpoints (Adapted from [3])

Endpoint with higher priority e.g. Death	Endpoint with lower priority e.g. Progression	Pair Outcome
Wins	Ignored	Wins
Loses	Ignored	Loses
Uninformative/Neutral	Wins	Wins
Uninformative/Neutral	Loses	Loses
Uninformative/Neutral	Uninformative/Neutral	Tied

We also calculate the established measures of net benefit, win odds, restricted mean survival time (RMST) ratio and RMST difference and a recently proposed measure called the pairwise win time to put our new measures in context. The RMST difference, CWD and pairwise win time have the same units (months or days) and provide measures of how much time is gained (or lost) from the treatment compared to the control. The RMST difference has been proposed as an appropriate measure of the treatment effect when the PH assumption is not met [4]. We propose using the CWR and CWD as complementary measures to the WR, win odds and net benefit to more fully understand the treatment effect. The win ratio, CWR, CWD, net benefit, win odds, RMST ratio and RMST difference and pairwise win time are defined in Table 2.

Table 2: Definitions of Win Ratio, Continuous Win Ratio, Continuous Win Difference and RMST Difference and Ratio

Parameter	Definition
Win Ratio (WR)	If a pair of patients is randomly picked, where one is from the treatment arm and the other is from the control arm, the win ratio is the ratio of the probability that a patient from the treatment arm wins to that a patient in the control arm wins.
Continuous Win Ratio (CWR)	The ratio of the length of time the treatment arm wins by when it wins to the length of time it loses by when it loses.
Continuous Win Difference (CWD)	The difference between the length of time the treatment arm wins by when it wins to the length of time it loses by when it loses, scaled by the number of pairs.
Net Benefit	The difference in win proportions.
Win Odds	The odds of win proportions where a tie results in a half win being assigned to the treatment arm and a half win to the control arm.
RMST Difference	The absolute gain or loss in event-free survival time due to treatment, where event times are restricted to be less than or equal to time τ. The RMST values are calculated up to a common time τ, which is the minimum of the largest observation times in the control and treatment arms.
RMST Ratio	The ratio of the area under the Kaplan-Meier curve for the treatment arm to the control arm, where the areas are calculated up to time τ.
Pairwise Win time	The pairwise average of the win time differences. A win time difference is the excess time that the patient in the treatment arm is in a more favorable (or unfavorable) state than the patient in the control arm in each pair over the effective common follow-up time (maximum of death times if both patients in the pair die and minimum of censoring time(s), if either patient is censored).

The WR, CWR, CWD, net benefit, win odds, RMST difference and ratio and pairwise win time were compared via the following four simulation scenarios:

1) With exponential distributions where the hazard ratios for progression and death are the same (PH assumption holds);

2) With exponential distributions where the hazard ratios for progression and death are in opposing directions (PH assumption holds);

3) With Weibull distributions with the same shape parameter and where the hazard ratios for progression and death are the same (PH assumption holds);

4) With Weibull distributions with different shape parameters for the control and treatment arms (PH assumption is violated). The fourth scenario is used to evaluate the impact of the critical PH assumption.

Finally, we use oncology trial data and estimate the win ratio, CWR, CWD, net benefit, and win odds in the setting where death is considered the highest-priority event, and progression is considered a secondary event, and if the priority is reversed. In addition, we estimate the RMST difference and ratio and pairwise win time. Further, we provide the inverse probability of censoring weighted (IPCW) WR to account for the occurrence of right-censoring in the time to event data (5), as well as the IPCW-adjusted CWR and CWD for these data.

2. Method: Estimating the WR, CWR, CWD, Win Odds, Net Benefit, RMST Ratio and RMST Difference

2.1 WR, CWR and CWD - unmatched analyses

Expanding on the WR algorithm in Pocock et al. [6], the derivation of the WR, CWR and CWD in the cancer setting with death as the primary event and progression as the secondary event are presented in Table 3. Without loss of generality, we will use months as the unit of time.

Table 3: Algorithm for the derivation of WR, net benefit, CWR and CWD

Description	Algorithm consideration
Pairing	Each patient in the control arm is paired with every patient in the treatment arm. Thus, if there are n_c patients in the control arm and n_t patients in the treatment arm, there are $N_p = n_c * n_t$ pairs.
Pairwise comparison for win vs. loss	Each one-to-one pair of patients is classified into one of five categories: a) The treatment arm patient dies first (including if the control arm patient does not die and is followed longer than the treatment arm patient); we calculate the difference in months between the control arm patient's death date or censoring date and the treatment arm patient's death date; b) The control arm patient dies first (including if the treatment arm patient does not die and is followed longer than the control arm patient); we calculate the difference in months between the treatment arm patient's death or censoring date and the control arm patient's death date; For cases c), d) and e), it is assumed death cannot be used to assess a win or a loss between the pair of patients being compared. c) The treatment arm patient progresses first (including if the control arm patient does not progress and is followed longer than the treatment arm patient); we calculate the difference between the treatment arm patient's PFS date and the control arm patient's progression date; d) The control arm patient progresses first (including if the treatment arm patient does not progress and is followed longer than the control arm patient); we calculate the difference between the control arm patient's PFS date and the treatment arm patient's progression date; e) None of the above, and the pair produces a tie. We assign 0 months to this pair. Based on death as the primary event and progression as the secondary event in our example, categories c) and d) are considered only if it is not known who dies first. Category e) contains patients who had neither death nor progression but will also include pairs where one of the patients had an event but the other patient's follow-up time was shorter and hence it could not be decided which arm won.
Win ratio	We denote the numbers of pairs in categories a), b), c), d), and e), respectively by N_a, N_b, N_c, N_d, and N_e. These numbers are used to quantify the treatment effect: $N_b + N_d = N_W$ the number of 'winners', and $N_a + N_c = N_L$ is the number of 'losers'. The win ratio treatment effect is then $WR = N_W / N_L$. A win ratio of 2 implies that treatment arm wins in twice as many pairs as it loses. It can also be interpreted as a patient in the treatment arm is 100% more likely to win than a patient in the control arm.

Continuous win ratio/continuous win difference	The trial's composite endpoint results are summarized by M_a, M_b, M_c, M_d, and M_e, which stand for the sums of the number of months from pairs in categories a), b), c), d), and e), respectively. $M_b + M_d = M_W$ is the number of months of advantage for 'winners'. Similarly, $M_a + M_c = M_L$ is the number of months of disadvantage for the 'losers'. $CWR=M_W/M_L$, the 'continuous win ratio', provides a ratio of how much time the treatment gains in the pairs where the treatment arm does better than the control arm to how much time the treatment loses in the pairs where the treatment arm does worse than the control arm. A CWR of 2 can be interpreted as the treatment winning by twice as much time as it loses. $CWD=(M_W-M_L)/N_p$ is the 'continuous win difference' (N_p is the total number of pairs). This measure provides an absolute difference of how much time the treatment gains in the pairs where the treatment arm does better than the control arm to how much time the treatment loses in the pairs where the treatment arm does worse than the control arm. A CWD of 2 months can be interpreted as winning two months more from being in the treatment arm than in the control arm. This represents the benefit (or loss) of treatment in the time to event, comparing a treated subject to a control subject.

In Table 3, in cases a) – d) when one of the patients in the pair is right censored, then the time advantage or deficit for the treated arm patient cannot be exactly specified (for e.g. if the first patient in the pair dies at day 50 and the second patient in the pair is alive and censored at day 70, the difference of 20 days is the minimum value of the difference). In our basic approach to calculate the CWR and CWD, which we evaluate in simulation studies, we take the time differences in these cases by treating the censoring times as if these were event times. This leads to minimizing the time advantage or deficit for such pairs. As with the basic unmatched win ratio calculations which are challenged by a high proportion of ties with a high proportion of right censored observations, so are the basic unmatched CWR and CWD calculations. We evaluate the impact on power as right censoring increases on the CWR and CWD for the four simulation scenarios we consider.

Note that an IPCW adjusted WR, with weights based on the Kaplan-Meier curve of the time-to-event data, can be calculated when the percentage of ties is high. Such an estimate has been shown to be asymptotically unbiased [5]. Since censoring has similar effects on the CWD and CWR, adjusting by such weights is anticipated to reduce the bias in these estimates when right censoring increases. In light of this, for our example with data from an oncology trial, we derive an IPCW-adjusted win ratio, CWR and CWD to provide estimates and associated CIs, which are generally expected to be wider than those for the unadjusted case. However, the IPCW-adjusted CWR and CWD have not been shown to be unbiased estimators under independent censoring and need further exploration to establish their properties.

In our derivation of the CWR and CWD, time advantages in the time to death and progression are not differentiated. This is similar to how the WR, net-benefit and win odds are derived; once the events are prioritized (say death), we prioritize the time advantage for death over the time advantage for progression.

Further details on the algorithm for deriving the CWR and CWD can be found in Appendix Table 2.

For the WR, CWR, CWD, and other measures, we recommend using bootstrapping to calculate the 95% Confidence Interval (CI) with a single data set, as done in the analysis of the oncology data set in the results section.

2.2 Net Benefit and Win Odds

The net benefit measures the difference in win proportions $(N_W - N_L)/N_p$, N_p is the number of pairs, and the win odds is defined as an odds of win proportions where a tie results in a half win being assigned to the treatment arm and a half win to the control arm ($[N_W+0.5*N_t]/[N_L+0.5*N_t]$, N_t is the number of ties) [7, 8]. The win ratio, net benefit and win odds test the same hypothesis of no difference in proportion of wins and result in similar p-values.

2. 3 RMST Difference, RMST Ratio, RMT-IF/Pairwise Win time

The RMST difference is the absolute gain or loss in the event-free survival time due to the treatment, where the event times are restricted to be less than or equal to time τ. The RMST values for each simulation are calculated up to a common time τ, which is the minimum of the largest observation times in the two arms i.e. control and treatment arm. Two more recent references [9, 10] introduce a concept called restricted mean time in favour of treatment (RMT-IF), which considers multiple events or states such as cancer progression, metastasis, and death in determining if the treatment arm performs better than the control arm. RMT-IF is defined as the net average time those in the study drug arm spend in a more favourable state than those in the control arm over a pre-specified time window. RMT-IF reduces to the RMST difference if there is only one event e.g. death. Here, we consider the RMST difference for PFS as well as another measure similar to RMT-IF called the pairwise win time ([11] and defined in Table 2) and compare them with the continuous win difference, since they all have the same units of time. The RMST ratio is the ratio of the area under the Kaplan-Meier curve for the treatment arm to that under the control arm, where the area is calculated up to a common time τ.

3. Simulations

A total of four scenarios are included and described in Table 4. For each scenario, we simulate 1000 trials with time-to-event data for 100 patients in the control arm and 100 patients in the treatment arm. Weibull or exponential distributions, with the exponential distribution being a special case of the Weibull distribution with a shape parameter value of 1, are used to independently generate the times to progression and death. The reference time for each patient is the time the patient is randomized to the study and the patient is followed until death, drop out or the end of study. In our simulations, we assume that all patients enter the study at the same time and there is no accrual time. We simulated censoring times in two ways:

1) Random Censoring: Censoring times are generated from an exponential or Weibull distribution with varying parameters to achieve different levels of censoring. An exponential distribution is used for censoring when the distribution used to generate the time to event is exponential and a Weibull distribution is used for censoring when the distribution used to generate the time to event is Weibull.

2) Administrative Censoring: Censoring is generated by a fixed upper limit on follow-up time.

With either type of censoring, the event of death is observed only if the time of death is earlier than the censoring time. The event of progression is observed only if the time to progression is earlier than either the censoring time or the time of death. We allow observation of follow-up time after a disease progression event, but not after either death or censoring. Patients who have neither progression nor death before the time of censoring are followed up to the censoring time.

In each of the 1000 simulations, we then create 10,000 pairs of control and treatment arm patients and determine in how many pairs the treatment arm wins compared to the control arm (and the time advantage for treatment) and in how many pairs the treatment arm loses compared to the control arm (and the time disadvantage for treatment), to calculate the WR, the CWR and the CWD. We first calculate the WR, the CWR and the CWD when progression is viewed as the primary event and death the secondary event, and then when death is viewed as the primary event and progression the secondary event. We estimate the mean WR, the mean CWR and the mean CWD as well as the mean win odds and the mean net benefit from the 1000 simulations and provide the 2.5 and 97.5 percentile of the distribution, the bias and the variance for each measure.

By using Weibull distributions for the time to progression and the time to death, we can consider both situations of proportional hazards (PH) and non-proportional hazards (NPH). When the shape parameters in the Weibull distributions for both arms are the same, the Weibull distributions fulfil the PH assumption. When the shape parameters in the Weibull distributions are

different for the treatment and control arms, the Weibull distributions do not fulfil the PH assumption, and we create NPH scenarios and estimate the mean values for the WR, CWR, CWD, win odds and net benefit.

In addition, we estimate the mean RMST ratio and mean RMST difference from the PFS curves for the two arms and the pairwise win time for all four scenarios. Both the 2.5 and 97.5 percentile of the distribution for the RMST ratio and difference and pairwise win time are provided.

Simulations for the WR, CWR, CWD, net benefit, win odds, RMST ratio and RMST difference and pairwise win time were carried out using SAS and replicated in R in most cases. Codes for these simulations are attached in the Appendix.

Table 4: Parameters associated with the Different Distributions/Scenarios

Scenarios	Distribution Parameters	Resulting Hazard Ratio
(1) time to PD, death and censoring from exponential distributions with the same HR for PD and Death (PH case)	Treatment PD ~ exp with hazard rate 0.06 (mTime=11.6) Death ~ exp with hazard rate 0.03 (mTime=23.1) Censoring ~ exp with hazard rate 0.0005	Death : 0.5 PD: 0.5
	Control PD ~ exp with hazard rate 0.12 (mTime =5.8) Death ~ exp with hazard rate 0.06 (mTime=11.6) Censoring ~ exp with hazard rate 0.0005	
(2) time to PD, death and censoring from exponential distributions with opposing HRs for PD and Death (PH case)	Treatment PD ~ exp with hazard rate 0.37 (mTime=1.9) Death ~ exp with hazard rate 0.03 (mTime=23.1) Censoring ~ exp with hazard rate 0.0005	Death : 0.5 PD: 2.64
	Control PD ~ exp with hazard rate 0.14 (mTime= 5) Death ~ exp with hazard rate 0.06 (mTime=11.6) Censoring ~ exp with hazard rate 0.0005	
(3) time to PD, death and censoring from Weibull distributions with the same shape but different scale parameters and the same HR for PD and Death (PH case)	Treatment PD ~ Weibull with shape parameter 2 andscale parameter 16.67 (mTime=13.9) Death ~ Weibull with shape parameter 2 andscale parameter 33.34 (mTime= 27.8) Censoring ~ Weibull with shape parameter 2 and scale parameter 2000	Death : 0.25 PD: 0.25
	Control PD ~ Weibull with shape parameter 2 and scale parameter 8.33 (mTime=6.9) Death ~ Weibull with shape parameter 2 and scale parameter 16.67 (mTime=13.9) Censoring ~ Weibull with shape parameter 2 and scale parameter 2000	

(4) time to PD, death and censoring from Weibull distributions with different shape and scale parameters (NPH case)	Treatment PD ~ Weibull with scale parameter 1.2 and shape parameter 16.67 (mTime=12.3) Death ~ Weibull with shape parameter 1.2 and scale parameter 33.34 (mTime=24.6) Censoring ~ Weibull with shape parameter 1.2 and scale parameter 2000	Not constant over time
	Control PD ~ Weibull with shape parameter 1 and scale parameter 8.33 (mTime=5.8) Death ~ Weibull with shape parameter 1 and scale parameter 16.67 (mTime=11.6) Censoring ~ Weibull with shape parameter 1 and scale parameter 2000	
~: distribution exp: exponential distribution mTime: median time in months and units of hazard rate is per month $exp(\lambda): f(t) = (1/\lambda)\, exp(t/\lambda)$ $Weibull(\lambda, k): f(t) = (k/\lambda)(t/\lambda)^{k-1} exp(-(t/\lambda)^{k})$		

4. Results

For the results presented below, we used the random censoring approach. In general, the results for the WR match closely for the two cases of *censoring* (random and fixed (administrative) censoring) when the percentage of pairs in the tied category is low and match closely for the CWR when the percentage of pairs in the right censored category is low (see Section 4.1.3 for details).

4.1 Scenario 1: Exponential distribution is used for time to PD, death and censoring and the HRs for PD and death are the same

4.1.1 Relationship between HR, win ratio, and RMST ratio under PH

The HR in our simulations is 0.5 for both PD and death and hence for PFS (see Note 1 in Appendix). This implies that patients in the treatment arm progress later and survive longer than those in the control arm. The exponential distributions used and the hazard rates assumed for PD, death and censoring are presented in Table 4, Scenario 1, and the results can be found in Table 5, Scenario 1.

In this case (Table 5, Scenario 1), we observe that the WR is close to the reciprocal of the HR for PFS both when PD is prioritized and when death is prioritized, since the HRs for PD and death are the same and PH is fulfilled. The win odds are close to the WR for both cases, since the percentage of ties is low. The values for net benefit when PD is prioritized and when death is prioritized are close to each other since the HRs for PD and death are the same. We observe that when the event rates in the exponential distributions used to model the time to PD and death are low, the RMST ratio for PFS is close in value to the reciprocal of the HR for PFS [12] and to the WR when PD is prioritized. The RMST difference from the PFS curves is smaller than the CWD when PD is prioritized. Since the time to PD is considered in the calculation of PFS (*PFS=min(PD, death)*) even for subjects who die later, we expect the RMST difference from the PFS curves to be close in value to the CWD when PD is prioritized. However, they are different measures of the treatment effect as one measures the time to the first event while the other is based on prioritized outcomes, and they need not produce the same value. The pairwise win time estimates the time in favor of treatment considering PD and death as the two states. In this scenario, the treatment prolongs the time to both PD and death.

4.1.2 Relation between win ratio and CWR under PH

For Scenario 1, where exponential distributions are used to model the time to PD and death with the same HR for PD and death, it is found that the CWR is equal to the square of the WR when PD is prioritized (Figure 1 and Appendix for derivation). The result also holds for a) when death is prioritized and b) when administrative censoring is applied, if the percentage of right censored cases is low (data not shown).

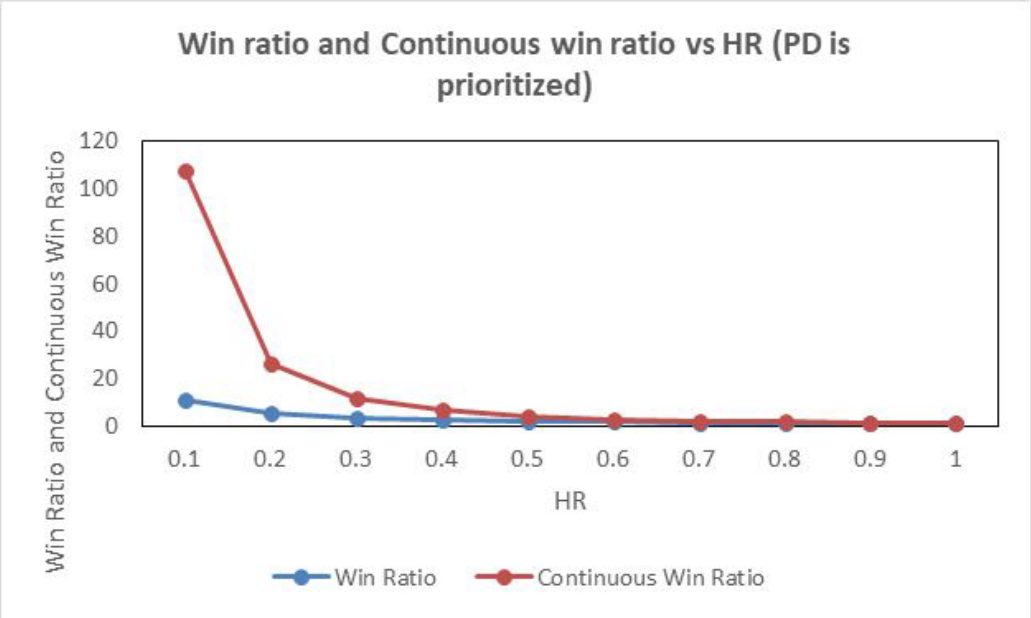
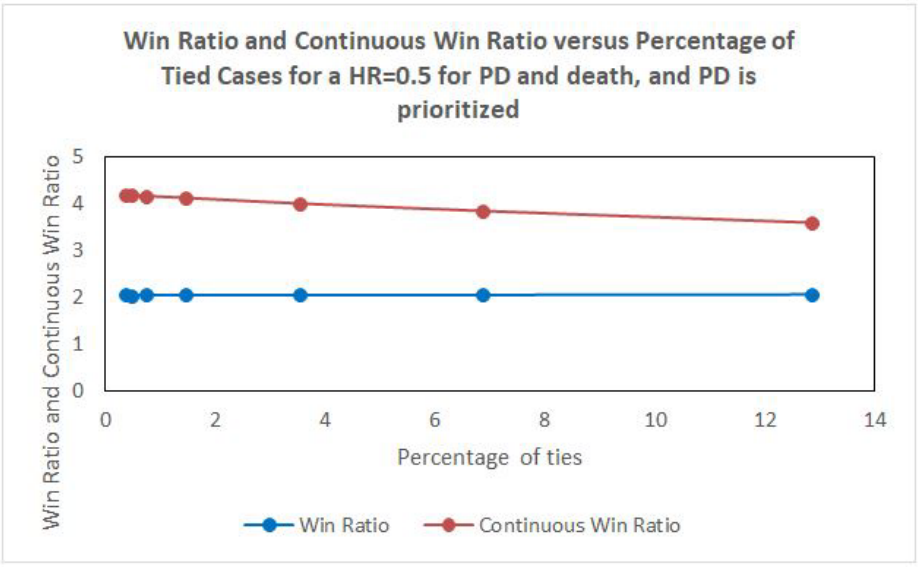


Figure 1: Relation between WR and CWR using exponential distributions to model the time to PD and death for different HRs (same HR for PD and death)

4.1.3. Under PH: impact of percentage of tied cases on the WR and impact of percentage of right censored cases on the CWR

Figure 2 shows how the WR and CWR are impacted when the percentage of tied cases and the percentage of right censored cases respectively increase. This is shown in the specific case when exponential distributions are used to model the time to PD and death for the treatment and control arm with the same HR for PD and death of 0.5. An exponential distribution is used to model censoring (the hazard rate for censoring is changed from 0.0005 to 0.01 per month), and PD is prioritized. The WR estimates are not impacted much with an increase in censoring for this PH scenario involving exponential distributions although the uncertainty of the estimate becomes larger with increasing censoring, as noted in [5] for their scenario a). For the CWR, there is a reduction in value as the percentage of right censoring increases. This is due to selective censoring of the treatment arm patients who tend to have longer time to event values than the control arm patients. In this scenario based on a HR of 0.5, patients in the treatment arm are winning twice as often as they lose and thus are more likely to be censored. As the HR gets closer to 1, the number of right censored cases decreases for a similar percentage of ties, and with it, its impact on the CWR decreases. We show the impact of right censoring on power for the CWR and CWD in Section 4.5.



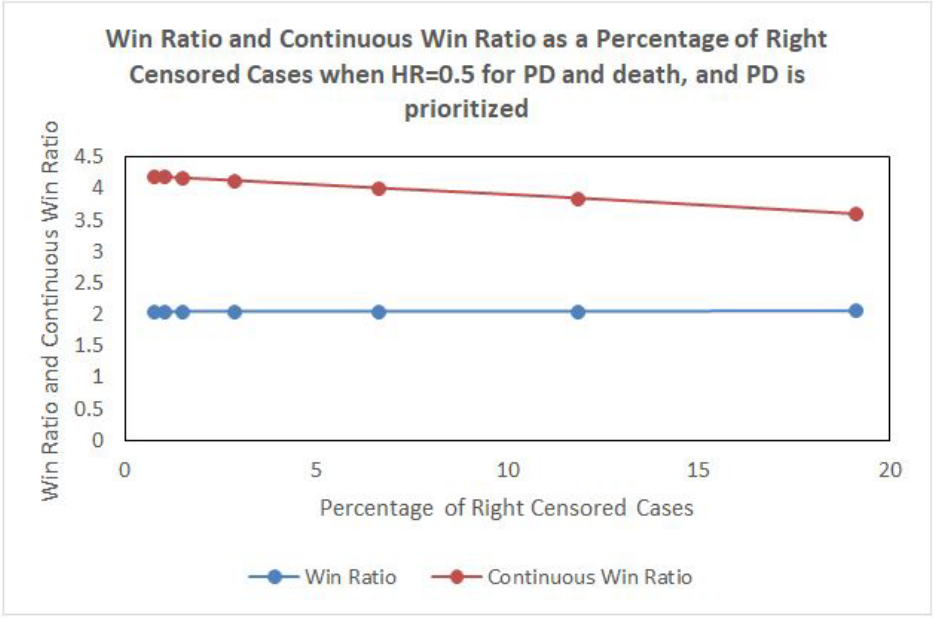


Figure 2: Relation between WR and percentage of tied cases and between CWR and percentage of right censored cases using exponential distributions to model the time to PD and death

4.2 Scenario 2: Exponential distribution is used for time to PD, death and censoring with HR for progression and death in opposing directions

Following Dong et al., 2020 [13], we set the HR for PD to 2.64 and the HR for death to 0.5 giving a HR for PFS of 2 for the hazard rates shown in the reference. This hypothetical scenario implies that patient progress faster but survive longer in the treatment arm than in the control arm. The exponential distributions used and the hazard rates assumed for PD, death and censoring are presented in Table 4, Scenario 2, and the results can be found in Table 5, Scenario 2.

The PFS curves for the control and treatment arm satisfy the PH assumption. In this case where the HRs for PD and death are in opposite directions, the WR is close in value to the reciprocal of the HR for PFS when PD is prioritized, and is close in value to the reciprocal of the HR for death when death is prioritized (Table 5, Scenario 2). The win odds are close in value to the WR

when PD is prioritized and when death is prioritized, since the percentage of ties is low. The net benefit when PD is prioritized and when death is prioritized are in opposite directions, as expected in this case of opposing HRs for PD and death. The CWR is close to 1 when PD is prioritized, which means that the treatment wins by as much time as it loses, and it is ~4.1 when death is prioritized, which means that the treatment wins by around four times as much time as it loses. In this case of opposing HRs for PD and death, it is difficult to obtain a relation for WR and CWR as was done when the HRs for PD and death are the same. For this case of PH using exponential distributions with low event rates used to model the time to PD and death, the RMST ratio for PFS is close in value to the win ratio when PD is prioritized. The RMST difference from the PFS curves is smaller than the CWD when PD is prioritized. In this scenario, the treatment prolongs death but hastens progression. Thus, the pairwise win time generally has a negative component for the time in favor of treatment for PD and a positive component for the time in favor of treatment for death. In contrast, for the CWD when death is prioritized, a judgement for win or loss for each pair is first made based on death and then on progression, if the decision cannot be made based on death. Hence, in this scenario, the pairwise win time is smaller than the CWD when death is prioritized.

4.3 Scenario 3: PH assumption holds for PD and death; time to PD, death and censoring modeled using Weibull distributions and the HRs for progression and death are the same

The HR in our simulations is 0.25 (see Note 2 in Appendix) for both PD and death and hence for PFS (see Note 3 in Appendix) (Table 4, Scenario 3). Thus, patients on the treatment arm progress much slower and survive much longer than those on the control arm. The Weibull distributions with the same shape parameter in both arms used to model the time to PD, death and censoring fulfil the PH assumption and are presented in Table 4, Scenario 3, and the results can be found in Table 5, Scenario 3.

We observe that the WR is close to the reciprocal of the HR for PFS both when PD is prioritized and death is prioritized, since the Weibull distributions used fulfil the PH assumption and since we assume the same HR for PD and death (Table 5, Scenario 3). However, the CWR is not the square of the win ratio in this case of Weibull distributions (Table 5, Scenario 3). The net benefit when PD is prioritized and when death is prioritized are similar in value since the HRs for PD and death are the same, which also holds true for the win odds. For this case of PH with Weibull distributions, the RMST ratio for PFS is not close in value to the win ratio when PD is prioritized. The RMST difference from the PFS curves is smaller than the CWD when PD is prioritized. In this scenario, the pairwise win time is greater than the CWD when PD is prioritized and when death is prioritized, since the treatment prolongs the time to both PD and death.

In Scenario 3, we investigated the relation between HR/WR and CWR/CWD under a specific hazard ratio of 0.25 for death and PD. We performed further simulations with Weibull distributions with various values of the same HR for PD and death and the same shape parameter in both arms to understand the relation between WR/HR and CWR/CWD. We find that when Weibull distributions with the same shape parameter in both arms of value greater than 1 are used to model the time to PD and death with the same HR as the corresponding exponential distributions, 1) the WR is the same, but the CWR and CWD are smaller in value than those for the corresponding exponential distributions (Figure 3 shows the relation between WR and CWR), and 2) the percentage of right censored cases is lower, and consequently its impact on the CWR and CWD is lower (data not shown). The WR is the same and the CWR and CWD are larger in value than those for the corresponding exponential distributions if Weibull distributions with the same shape parameter in both arms with a value less than 1 are used to model the time to PD and death with the same HR as the corresponding exponential distributions.

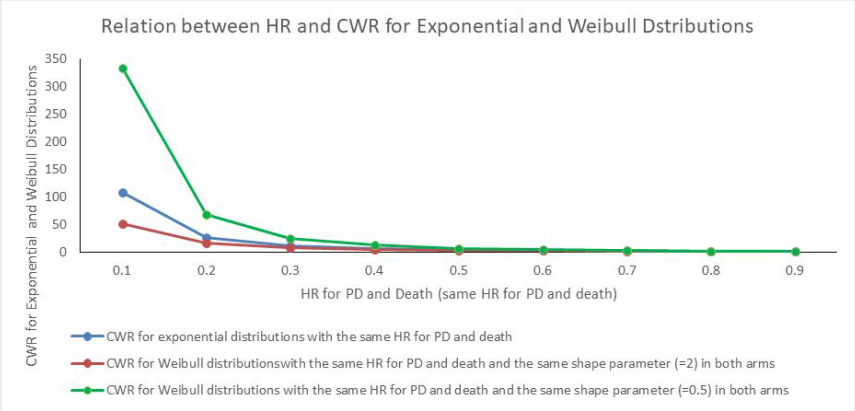


Figure 3: Relation between HR/WR and CWR for exponential and Weibull distributions (with shape parameter=2 for both arms in one case and =0.5 for both arms in the other case) with the same HR for PD and death and PD is prioritized. For the exponential distributions, CWR=WR^2, as in Figure 1, but CWR<WR^2 and CWR>WR^2 for the Weibull distributions with the same HR for PD and death as the corresponding exponential distributions and shape parameter in both arms =2 and the shape parameter in both arms =0.5 respectively.

The WR=1/HR in all cases since PH is fulfilled.

4.4 Scenario 4: Weibull distributions with different shape and scale parameters and thus PH assumption does not hold

The Weibull distributions with different shape parameters in the two arms used to model time to PD, death and censoring do not fulfil the PH assumption and are presented in Table 4, Scenario 4, and the results can be found in Table 5, Scenario 4.

Example PFS curves for this case from one simulation are shown in Figure 4. As shown in Table 5, Scenario 4, in this case of NPH, the WR when PD is prioritized and the RMST ratio from the PFS curves are not the same as the reciprocal of the HR from Cox regression for PFS, which is expected. The net benefit when PD is prioritized and when death is prioritized are not close in value, which is also expected in this case of NPH. The CWR is 5.2 when PD is prioritized and 4.1 when death is prioritized, and these values can be interpreted as before. The RMST difference from the PFS curves is smaller than the CWD when PD is prioritized. In this scenario, the pairwise win time is greater than the CWD when PD is prioritized and death is prioritized, since the treatment prolongs the time to both PD and death.

Table 5: Simulation Results for the Scenarios in Table 4

Statistics	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Mean HR for PFS (using cox regression)(2.5 percentile, 97.5 percentile)	0.502 (0.362, 0.670) 1/Mean HR=1.992	2.029 (1.488, 2.711) 1/Mean HR=0.493	0.250 (0.170, 0.339) 1/Mean HR=4	0.444 (0.316, 0.589) 1/Mean HR=2.252
Results when PD is prioritized				
Mean % of Ties	0.37%	0.17%	0%	0.27%
Mean % of Cases in Right Censored Category	0.74%	0.39%	0.01%	0.41%
Mean WR (2.5 percentile, 97.5 percentile) Bias, Variance	2.049 (1.429, 2.891) 0.049, 0.145	0.507 (0.357, 0.702) 0.007, 0.008	4.165 (2.809, 6.151) 0.165, 0.795	2.619 (1.799, 3.780) 0.0006, 0.256

Mean CWR (2.5 percentile, 97.5 percentile) Bias, Variance	4.195 (2.289, 7.297) 0.195, 1.672	1.024 (0.505, 1.827) -0.010, 0.118	10.723 (5.949,18.680) -0.002, 11.703	5.183 (2.793, 9.042) 0.012, 2.678
Mean CWD (2.5 percentile, 97.5 percentile) Bias, Variance	9.079 months (5.309, 13.228) -0.155, 4.220	-0.049 months (-2.299, 2.560) -0.037, 1.585	8.795 months (6.821,10.933) -0.001, 1.212	9.645 months (6.175,13.251) -0.063, 3.316
Mean Win Odds (2.5 percentile, 97.5 percentile) Bias, Variance	2.042 (1.429, 2.891) 0.042, 0.142	0.508 (0.357, 0.703) 0.008.0.008	4.164 (2.809, 6.151) 0.164, 0.795	2.610 (1.793, 3.771) -0.008, 0.252
Mean Net Benefit (2.5 percentile, 97.5 percentile) Bias, Variance	0.333 (0.177, 0.486) -0.0002, 0.006	-0.331 (-0.474, -0.174) 0.002, 0.006	0.602 (0.475, 0.720) 0.002, 0.004	0.436 (0.284, 0.581) -0.001, 0.006
Results when Death is prioritized				
Mean % of Ties	0.37%	0.17%	0%	0.27%
Mean % of Cases in Right Censored Category	1.49%	1.46%	0.03%	0.83%
Mean WR (2.5 percentile, 97.5 percentile) Bias, Variance	2.022 (1.425, 2.852) 0.022, 0.138	1.990 (1.416, 2.792) -0.010, 0.134	4.085 (2.820, 5.995) 0.085, 0.753	2.298 (1.622, 3.285) 0.004, 0.187
Mean CWR (2.5 percentile, 97.5 percentile) Bias, Variance	4.162 (2.199, 7.367) 0.162, 1.770	4.137 (2.192, 7.358) -0.055, 1.748	11.024 (6.243,19.339) -0.005, 12.529	4.104 (2.200, 7.365) 0.018, 1.783
Mean CWD (2.5 percentile, 97.5 percentile) Bias, Variance	16.112 months (8.802, 23.617) -0.431, 14.628	16.050 months (8.792, 23.580) -0.493, 14.631	14.685 months (11.305,18.085) -0.006, 3.262	14.435 months (8.173, 20.709) -0.144, 10.416
Mean Win Odds (2.5 percentile, 97.5 percentile) Bias, Variance	2.016 (1.422, 2.824) 0.016, 0.136	1.987 (1.415, 2.785) -0.013, 0.133	4.085 (2.820, 5.995) 0.085, 0.752	2.292 (1.617, 3.275) -0.002, 0.185
Mean Net Benefit (2.5 percentile, 97.5 percentile) Bias, Variance	0.327 (0.174, 0.477) -0.006, 0.006	0.321 (0.172, 0.472) -0.012, 0.006	0.596 (0.477, 0.714) -0.004, 0.004	0.383 (0.236, 0.532) -0.0005, 0.006
Other Measures				
Mean RMST Ratio and difference from the PFS curves (2.5 percentile, 97.5 percentile)	Ratio: 1.838 (1.410, 2.355) Difference: 4.558 months (2.545, 6.727)	Ratio: 0.554 (0.430, 0.710) Difference: -2.051 months (-3.0352, -1.142)	Ratio: 1.771 (1.535, 2.042) Difference: 5.066 months (3.675, 6.634)	Ratio: 1.985 (1.543, 2.519) Difference: 5.360 months (3.425, 7.405)
Mean Pairwise Win time (2.5 percentile, 97.5 percentile)	17.984 months (10.694, 25.556)	13.940 months (6.785, 21.509)	18.106 months (14.561, 21.611)	16.664 months (10.344, 23.019)

Scenario 1. Time to PD, death and censoring modeled using exponential distributions with the same HR for PD and death (the PH assumption holds: HR for PD and death is 0.5)
Scenario 2. Time to PD, death and censoring modeled using exponential distributions with opposing HRs for PD and death (the PH assumption holds: HR for PD is 2.64 and for death it is 0.5)
Scenario 3. Time to PD, death and censoring modeled Weibull distributions with the same shape parameters but different scale parameters for the control and treatment arms (the PH assumption holds: HR for PD and death is 0.25)
Scenario 4. Time to PD, death and censoring modeled Weibull distributions with different shape and scale parameters for the control and treatment arms (the PH assumption does not hold)
A hazard rate of censoring of 0.000000005 for exponential distributions or the same value for the scale parameter in the Weibull distribution for censoring was used to obtain the true value for CWR and CWD to estimate bias

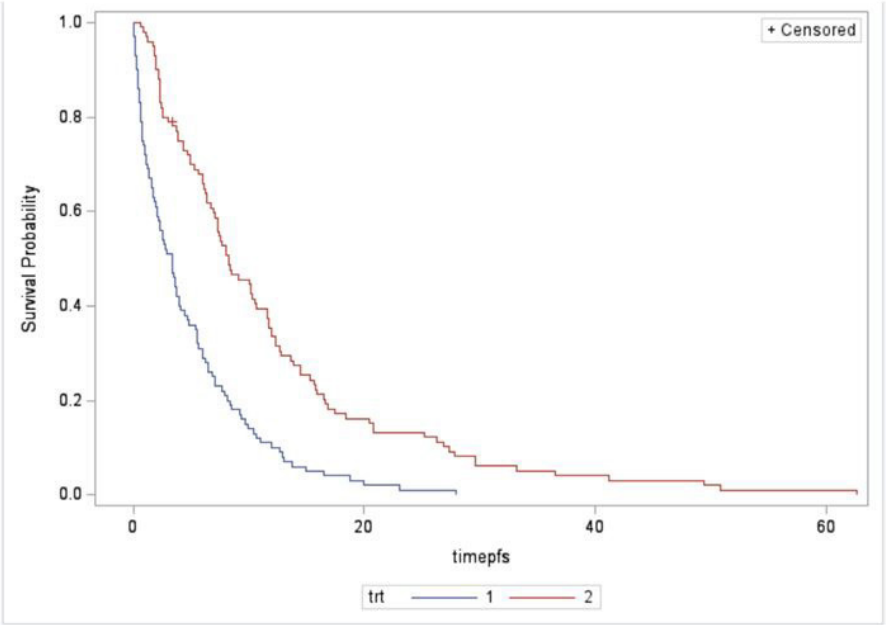


Figure 4: PFS curves for treatment and control arm using Weibull distributions to model the time to PD and death (NPH case) (example of 1 simulated curve).

4.5 Impact of Right Censoring on Power for the CWR and CWD

We studied the impact of right censoring on power for the CWR and CWD for the four scenarios in Sections 4.1 to 4.4 by changing the hazard rate or parameter for censoring. There are limitations in these simulations because the bootstrapping procedure used to obtain the confidence intervals for each simulation could be performed only a 100 times due to computational limits (with 100 simulations and 100 patients in each arm and 100 replicates for each simulation, there are 10^8 observations in the bootstrapping data set). However, the results (Table 6 and Appendix Table 3) show that for the four scenarios studied, the power remains high in general for the CWR even as the percentage of right censoring increases. Further, the CWD estimation is affected when the right censoring percentage reaches >20%. Note that our code also allows the estimation of power for the WR, net benefit and win odds for different parameters for censoring.

Table 6: Power for Detecting Differences in CRW and CWD when PD is prioritized and death is prioritized for Scenario 1 with time to PD, death and censoring from exponential distributions (the other 3 scenarios are shown in Appendix Table 3).

Scenario 1					
Hazard Rate for Censoring and percentage of Right Censoring when PD is prioritized	Power for CWD when PD is prioritized	Power for CWR when PD is prioritized	Hazard Rate for Censoring and percentage of Right Censoring when death is prioritized	Power for CWD when death is prioritized	Power for CWR when death is prioritized
0.0005, 0.8% 0.005, 7% 0.01, 12.4% 0.02, 19.1%	94% 83% 68% 13%	91% 93% 92% 90%	0.0005, 1.6% 0.005, 13% 0.01, 21.7%	92% 67% 34%	89% 90 % 87 %

The true values for CWR and CWD were obtained by conducting 1000 simulations with a very low hazard rate for censoring (unit of hazard rate is per month).

5. Application to Clinical Trial Data

We estimated the WR and the CWR using data from a Phase III trial of 172 patients with surgically resected non-small lung cancer randomized to receive radiotherapy with or without adjuvant platinum-based chemotherapy [14]. We analysed the data from 164 eligible patients, who had a mean follow-up time from randomization of 3.7 years [15]. We use the time to first relapse (PD) and the time to death as the events and calculate the WR,CWR, CWD (Table 7).

Note that the HR for PFS is 0.683 but the PFS curves do not satisfy the PH assumption. The results for the WR, CWR and CWD indicate that the arm with radiotherapy plus chemotherapy wins 1.8 times as often as it loses to the control arm with radiotherapy alone, wins by 1.8 times as many days as it loses, and wins by about 205 days compared to the control arm, when PD is prioritized. The radiotherapy plus chemotherapy arm wins 1.4 times as often as it loses to the control arm, it wins by 1.4 times as many days as it loses and wins by about 138 days compared to the control arm, when death is prioritized. Based on the low percentage of ties, our simulation results (see Section 4.1.3) suggest that the WR estimate is likely accurate. However, the CWD is likely not an accurate estimate of the duration of benefit due to the large number of right censored cases. The win odds when PD is prioritized and when death is prioritized are close in value to the WR in each case, which is expected given the small percentage of ties. The IPCW-adjusted WR when PD is prioritized and when death is prioritized are close in value to the unadjusted WR in each case, given the small percentage of ties. However, we still estimated the IPCW-adjusted WR to illustrate that this can done if the percentage of ties is higher. We also provided the IPCW-adjusted CWR and CWD when PD is prioritized and when death is prioritized. The net benefit when PD is prioritized and when death is prioritized are of similar positive magnitude. In this case of NPH, the RMST ratio from the PFS curves is not equal to the WR when PD is prioritized. The RMST difference from the PFS curves of ~223 days is slightly larger than the CWD of ~205 days and larger than the pairwise win time of ~182 days.

Table 7: WR, CWR, CWD, Win Odds, Net Benefit, RMST ratio and difference and Pairwise win time using data from an oncology trial

HR for PFS	**0.683 (0.491, 0.950)1/Mean HR=1.464**
Results when PD is prioritized	
% of Ties	2.68%
% of Cases in Right Censored Category	22.08%

Win Ratio	1.815 (1.221, 2.620)*
IPCW Adjusted Win Ratio	1.679 (1.157, 2.428)*
Continuous Win Ratio	1.804 (0.936, 3.519)*
Continuous Win Difference	204.64 (-22.591, 416.105) days*
IPCW Adjusted Continuous Win Ratio	1.707 (0.891, 3.410)*
IPCW Adjusted Continuous Win Difference	206.14 (-43.706, 493.593) days*
Win Odds	1.785 (1.217, 2.542)
Net Benefit	0.282 (0.098, 0.435)
Results when Death is prioritized	
% of Ties	2.68%
% of Cases in Right Censored Category	27.98%
Win Ratio	1.436 (0.980, 2.103)*
IPCW Adjusted Win Ratio	1.350 (0.922, 2.012)
Continuous Win Ratio	1.430 (0.790, 2.581)*
Continuous Win Difference	138.435 (-97.176, 354.218) days*
IPCW-adjusted Continuous Win Ratio	1.402 (0.769, 2.507)*
IPCW-adjusted Continuous Win Difference	135.421 (-112.507, 373.915) days*
Win Odds	1.422 (0.981, 2.048)*
Net Benefit	0.174 (-0.01, 0.344)*
Other Measures	
RMST Ratio and Difference from the PFS curves	Ratio (95% CI): 1.388 (0.949, 2.124) * Difference (95% CI): 222.96 (-53.406, 499.326) days
Pairwise Win time	182.5 (-70.363, 414.957) days*

* The 95% CI is calculated via bootstrapping

6. Discussion

The WR is an alternate measure of the treatment effect in the case of a composite endpoint. In this approach, the events that make up the composite endpoint can be ordered or prioritized. We have studied the WR and proposed the novel CWR and CWD, and considered the properties of these measures using simulations with exponential and Weibull distributions covering PH and NPH scenarios. We have also estimated the WR, CWR and CWD for an oncology data set. In addition, we have calculated the net benefit, win odds, RMST ratio and RMST difference and pairwise win time to put our results in context.

For the common oncology composite endpoint of progression-free-survival, defined as the earliest of progression or death, we have determined the WR when PD is prioritized and when death is prioritized. To simulate the endpoint, we have considered the case of PH and NPH for the PFS curves. The WR is the reciprocal of the HR for a single time to event endpoint or a composite time to event endpoint with the same HR for all events (e.g. PD and death) if the PH assumption holds. The WR when PD is prioritized is close in value to the RMST ratio from the PFS curves when exponential distributions are used to model the time to PD and death and when the event rates are low [12]. When exponential distributions are used to model the time to PD

and death for both the treatment and control arm with the same HR for PD and death, it is found that the CWR is equal to the square of the WR (see Appendix for Proof). When Weibull distributions are used to model the time to PD and death with the same HR for PD and death as the corresponding exponential distributions and with the same shape parameter in both arms of value greater than 1, the WR is the same, but the CWR and CWD are smaller in value than those for the corresponding exponential distributions. Thus, when specific distributions are used to model the time to PD and death, such as exponential and Weibull distributions, with the same HR for PD and death, it is possible to study the relation between HR/WR and CWR/CWD. However, when the HRs for PD and death are in opposing directions, no general relation between WR and CWR can be obtained, and it needs to be studied on a case-by-case basis.

We found that when exponential distributions describe the time to PD and death, which automatically fulfil the PH assumption, the WR is close to the reciprocal of the HR for PFS when PD is prioritized and is close to the reciprocal of the HR for death when death is prioritized. However, when PH is not fulfilled, as expected, the WR is different than the reciprocal of the HR for PFS, when PD is prioritized. In the case of NPH and in the case of opposing HRs for PD and death, the CWR, CWD, RMST ratio and difference and pairwise win time can be used as additional measures to the WR to understand the treatment effect. The RMST difference and the CWD both provide a measure of time gained (or lost) on the treatment arm compared to the control arm but are not the same measures of the treatment effect as one measures the time to the first event while the other is based on prioritized outcomes. In the PH scenarios studied, the RMST difference provided a more conservative estimate than the CWD of the time gained or lost on the study drug. Although these measures of the treatment effect are estimated in different ways, they were close in value in most cases we examined. We provided a third measure calculated in months or days, the pairwise win time, which estimates the time in favor of treatment considering PD and death as the two states. We also provided results for the net benefit and win odds for all the scenarios. The win odds were very close in value to the WR in the cases we examined since the percentage of ties was low. These measures could complement the WR in understanding the results.

We also simulated the censoring for PFS in two different ways to mimic random censoring and administrative censoring. Using these censoring methods, we obtain very similar results in general for the WR if the percentages of ties is low. We have shown that the WR is stable however even if the percentage of ties increases when exponential distributions are used to model the time to PD and death. Other work has also shown that in general, as the percentage of ties increases, the accuracy of the WR decreases and an IPCW-adjusted WR gives an unbiased estimate of the treatment effect (5). We obtain similar values for the CWR and CWD using the two censoring methods, if the percentage of right censored cases is low. We estimated the IPCW-adjusted WR, CWD and CWR to adjust for censoring for our example with data from an oncology trial. The CWR like the WR can be affected by censoring, leading to several ties. We have shown that right censoring impacts the CWR more than the WR. However, compared to the CWD, the CWR maintains the power even when the percentage of right censoring is high (example 20%) in the scenarios we examined. The utility of these measures should be further explored and validated in larger data sets from cancer trials.

Besides the WR, other prioritized comparisons include the Finkelstein-Schoenfeld method and the net benefit; note that the Finkelstein-Schoenfeld statistics differ from the net benefit estimate only by a constant (see Table 2 of [16] for a summary of the tests), and thus should result in the same p-values. The WR, net benefit, win odds, Finkelstein-Schoenfeld test all fall under the framework of the generalized pairwise comparisons methods. In general, in this framework, the difference in endpoints (Y-X) (intervention arm - control arm) can be compared to a threshold of clinical relevance Δ. For a single endpoint, if $(Y > X + \Delta)$, the pair is counted as a win and if $(X \geq Y + \Delta)$, the pair is counted as a loss[17]. Otherwise, the pair is a tie. In the case of censoring, the pair is counted as uninformative as it might not be possible to compare the difference to a threshold. When there are several endpoints, the endpoints are prioritized based on clinical relevance [17]. The endpoint with the highest priority is analysed first, and neutral and uninformative pairs are analysed based on the endpoint of lower priority. Our approach is a little different with the CWR and CWD. These new measures give us an idea of how much benefit in terms of time the study drug provides to the patient compared to the control, and may help doctors and payers better understand the effect of the study drug.

Appendix

Relation between win ratio (WR) and continuous win ratio (CWR) for exponential distributions:

The WR was introduced in Pocock et al. [6]. Consider the case where patients are to be compared by two prioritized outcomes i.e. patients are first compared with respect to the first priority outcome and are compared with respect to the second outcome only if the comparison of the highest priority outcome does not result in a win or loss. If a pair of patients is randomly picked, where one is from the treatment arm and the other is from the control arm, the WR is the ratio of the probability that a patient from the treatment arm wins to that a patient in the control arm wins. A WR >1 implies that the treatment is beneficial. Mathematically this can be defined as follows.

Let,

T = time to event for treatment arm

C = time to event for control arm

and

p(t) = win probability for treatment arm

p(c) = win probability for control arm

Then the WR is defined as

$$WR = \frac{p(t)}{p(c)}$$

In this paper we are introducing an extension to the win ratio measure which we are calling the continuous win ratio. We define this to be the ratio of the number of months the treatment arm wins by when it wins to the number of months it loses by when it loses. Mathematically, we define this as

$$CWR = \frac{p(t) * E(T - C|T > C)}{p(c) * E(C - T|C > T)}$$

Here $E(x)$ represents the expected value, or mean value of the variable x. There is a relationship between WR and CWR.

$$CWR = WR * \frac{E(T - C|T > C)}{E(C - T|C > T)}$$

It can be shown for the special case that T and C follow exponential distributions with hazard rates of λ_T and λc respectively, that

$$E(T - C|T > C) = \frac{\lambda_T \lambda_C}{\lambda_T (\lambda_T + \lambda_C)}, and\ \ E(C - T|C > T) = \frac{\lambda_T \lambda_C}{\lambda_C (\lambda_T + \lambda_C)}$$

So, with T and C exponential,

$$CWR = WR * \frac{E(T)}{E(C)} = WR * \frac{\lambda_C}{\lambda_T}$$

$$CWR = WR * \frac{1}{HR} = WR^2$$

The last equality following from the exponential distributions having proportional hazards.

Continuous Win Ratio Assuming Exponential Distributions for PD and Death and the Same HR for PD and Death and Assuming Administrative Censoring (PH Case):

We consider the case where the PH assumption is fulfilled (exponential distributions to model PD and death for both the treatment and control arm). The censoring is based on a fixed follow-up of 120 months. In this case, we provide the hazard ratio (HR) for PFS and the WR CWR, CWD and other measures when PD is prioritized and when death is prioritized. The HR in our simulations for the proportional hazards case is 0.5 for both PD and Death and hence for PFS. The hazard rates for PD are 0.12 (median time of 5.8 months) and 0.06 per month (median time of 11.6 months) in the control and treatment arm respectively and for death 0.06 (median time of 11.6 months) and 0.03 per month (median time of 23.1 months) in the control and treatment arm respectively.

Appendix Table 1: WR, CWR, CWD and other measures assuming exponential distributions for PD and death and the same HR for PD and death and assuming administrative censoring (PH case)

Mean HR for PFS (using cox regression)(2.5 percentile, 97.5 percentile)	0.508 (0.375, 0.661)1/ Mean HR=1.969
Results when PD is prioritized	
Mean % of Ties	0%
Mean % of Cases in Right Censored Category	0.69%
Mean Win Ratio (2.5 percentile, 97.5 percentile)	2.023 (1.453, 2.807)
Mean Continuous Win Ratio (2.5 percentile, 97.5 percentile)	4.045 (2.261, 6.679)
Mean Continuous Win Difference (2.5 percentile, 97.5 percentile)	8.938 (5.358, 12.786) months
Mean Win Odds (2.5 percentile, 97.5 percentile)	2.023 (1.453, 2.807)
Mean Net Benefit (2.5 percentile, 97.5 percentile)	0.330 (0.185, 0.475)
Results when Death is prioritized	
Mean % of Ties	0%
Mean % of Cases in Right Censored Category	2.75%
Mean Win Ratio (2.5 percentile, 97.5 percentile)	2.014 (1.474, 2.806)
Mean Continuous Win Ratio (2.5 percentile, 97.5 percentile)	3.987 (2.217, 6.734)
Mean Continuous Win Difference (2.5 percentile, 97.5 percentile)	15.590 (9.010, 22.651) months
Mean Win Odds (2.5 percentile, 97.5 percentile)	2.014 (1.474, 2.806)
Mean Net Benefit (2.5 percentile, 97.5 percentile)	0.328 (0.192, 0.475)
Other Measure	
Mean RMST Ratio and Difference from PFS Curves (2.5 percentile, 97.5 percentile)	Ratio: 1.830 (1.437, 2.336) Difference: 4.559 (2.736, 6.586) months

Appendix Table 2: Calculation of the win ratio and continuous win ratio assuming death is prioritized over progression. An algorithm following similar logic can be used when progression is prioritized over death.

	Win ratio	**Continuous win ratio**
If both the control and treatment arm patient in the pair die	The winner is the one who dies later (flag=1 if the treatment arm patient dies later than the control arm patient and 0 if the treatment arm patient dies earlier).	For the treatment arm, we calculate how many months the treatment arm patient wins or loses by (death date of winner- death date of loser)
If either the control arm patient or the treatment arm patient in the pair dies but not both	The winner is the one who doesn't die (flag=1 if the treatment arm patient does not die and 0 if the treatment arm patient dies), assuming that the patient is followed for a longer period than the patient who dies. If the patient who doesn't die is not followed for a longer period than the patient who dies, then it needs to be observed which of the two patients in the pair progresses first to decide the winner.	For the treatment arm, we calculate how many months the treatment arm patient wins or loses by (censoring date of winner- death date of loser), assuming that the winner is followed for a longer period than the loser.
If neither the control arm patient nor the treatment arm patient in the pair dies	It needs to be seen as to who progresses first to decide the winner.	
If both the control and treatment arm patient in the pair progress	After considering death, if it cannot be determined who wins and if both patients in the pair progress, then the patient who progresses later wins (flag=1 if the treatment arm patient progresses later than the control arm patient and 0 if the treatment arm patient progresses earlier).	For the treatment arm, we calculate how many months the treatment arm patient wins or loses by (progression date of winner- progression date of loser)
If either the control arm or the treatment arm patient in the pair progresses but not both	After considering death, if we cannot determine who wins, and only either the control arm or the treatment arm patient progresses, then the person who does not progress wins (flag=1 if the treatment arm patient does not progress and 0 if the treatment arm patient progresses), assuming that the patient is followed for a longer period than the patient who progresses. If the patient who doesn't progress is not followed for a longer period than the patient who progresses, we declare a tie (we assign flag=0.5).	For the treatment arm, we calculate how many months the treatment arm patient wins or loses by (max(PFS date or censoring date of winner)- progression date of loser), assuming the winner is followed for a longer period than the loser. If not, we declare a tie, and we assign the number of months to both arms to be 0.
If neither the control arm nor the treatment arm patient in the pair progresses	After considering death, if we cannot determine who wins, and neither the control arm nor the treatment arm patient progresses, then we declare a tie (we assign a flag=0.5).	We assign the number of months to both arms to be 0.

	The win ratio is the sum of flag=1 divided by sum of flag =0, considering all control arm and treatment arm pairs	The continuous win ratio is the sum of months where flag=1 divided by the sum of months where flag=0, considering all control arm and treatment arm pairs. The continuous win difference is the sum of months where flag=1 minus the sum of months where flag=0, considering all control arm and treatment arm pairs. This result must be divided by the number of pairs.

Appendix Table 3: Power (shown in bold) for Detecting Differences in CRW and CWD when PD is prioritized and death is prioritized

	Scenario 1	**Scenario 2**	**Scenario 3**	**Scenario 4**
CWD and CWR when PD is prioritized				
CWD (hazard rate or parameter for censoring, percent of right censoring, power)	0.0005, 0.8%, 94% 0.005, 7%, 83% 0.01, 12.4%, 68% 0.02, 19.1%, 13%	0.0005, 0.38%, 95% 0.01, 6.9%, 88% 0.02, 11.4%, 80% 0.05, 20%, 61%	0.0005, 0.01%, 96% 0.01, 3.2%, 89% 0.02, 10.8%, 68% 0.0333, 21.1%, 6%	0.0005, 0.44%, 92% 0.01, 10.6%, 76% 0.02, 18.4%, 22%
CWR (hazard rate or parameter for censoring, percent of right censoring, power)	0.0005, 0.8%, 91% 0.005, 7%, 93% 0.01, 12.4%, 92% 0.02, 19.1%, 90%	0.0005, 0.38%, 95% 0.01, 6.9%, 84% 0.02, 11.4%, 78% 0.05, 20%, 58%	0.0005, 0.01%, 94% 0.01, 3.2%, 94% 0.02, 10.8%, 90% 0.0333, 21.1%, 85%	0.0005, 0.44%, 91% 0.01, 10.6%, 94% 0.02, 18.4%, 92%
CWD and CWR when Death is Prioritized				
CWD (hazard rate or parameter for censoring, percent of right censoring, power)	0.0005, 1.6%, 92% 0.005, 13%, 67% 0.01, 21.7%, 34%	0.0005, 1.5%, 92% 0.005, 12.7%, 59% 0.01, 21.1%, 22%	0.0005, 0.04%, 93% 0.005 , 2.4%, 91% 0.01, 9.2%, 73% 0.02, 26.4%, 19%	0.0005, 0.87%, 92% 0.005, 10.4%, 85% 0.01, 19.2%, 56%
CWR (hazard rate or parameter for censoring, percent of right censoring, power)	0.0005, 1.6%, 89% 0.005, 13%, 90 % 0.01, 21.7%, 87%	0.0005, 1.5%, 89% 0.005, 12.7%, 88% 0.01, 21.1%, 82%	0.0005, 0.04%, 91% 0.005, 2.4%, 92% 0.01, 9.2%, 90% 0.02, 26.4%, 83%	0.0005, 0.87%, 91% 0.005, 10.4%, 91% 0.01, 19.2%, 91%

The true values for the CWR and CWD were obtained by conducting 1000 simulations with a very low hazard rate of censoring of 0.000005 per month for exponential distributions or a value of 0.000005 for the scale parameter for Weibull distributions.

Power is shown in bold in the above table.

Notes:

Note 1) : The HR for PFS is the ratio of the sum of the hazard rates for PD and death in the treatment arm to that in the control arm. This is because if X and Y are independent exponential random variables with hazard rate λ_1 and λ_2 respectively, then

$Z=min(X, Y)$ is an exponential random variable with hazard rate $\lambda_1+\lambda_2$. In reality, PD and death may not be independent and this is an approximation for PFS.

Note 2): For two Weibull distributions with the same shape parameter 'a' for the treatment and control arms, the hazard ratio HR is $(b0/b1)^a$, where b0 and b1 are the scale parameter (event rates) of the control and treatment arm respectively.

Note 3): It can be shown that the minimum of two independent Weibull random variables with the same shape parameter is a Weibull random variable with the same shape parameter and a modified scale parameter (modified event rate). Using this, it can be shown that the hazard ratio for PFS is λ, assuming that independent Weibull random variables all with distributions having the same shape parameter are used to model PD and death in the control and treatment arm respectively and assuming that the hazard ratio for both PD and death is λ.

Codes for Win Ratio

https://www.annexpublishers.com/articles/JBIA/10103-Code.pdf

Consent to Participate

No human clinical trial was conducted, but data from an old clinical trial have been analyzed. The reference for the data is below.

Piantadosi S. Clinical Trials: A Methodologic Perspective, 4th Edition, 2024. Code and data are publicly available at: https://www.wiley.com/en-us/Clinical+Trials%3A+A+Methodologic+Perspective%2C+4th+Edition-p-9781394195671#downloadstab-section

Funding Declaration

There was no funding for this manuscript

Ethics Declaration

Not applicable

Consent to Publish

No human trial was performed. See above in "Consent to Participate" section

Data Availability Declaration

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request

Competing Interest Declaration

The authors have no competing interests to declare

References

1. Oakes D (2016) On the win-ratio statistic in clinical trials with multiple types of event," Biometrika, Biometrika Trust, 103: 742-5.

2. Dong G, Hoaglin DC, Qiu J, Matsouaka RA, Chang Y, Wang J, Vandemeulebroecke M (2020) The Win Ratio: On Interpretation and Handling of Ties, Statistics in Biopharmaceutical Research, 12: 99-106.

3. Péron J, Buyse M, Ozenne B, Roche L, Roy P (2018) An extension of generalized pairwise comparisons for prioritized outcomes in the presence of censoring. Stat Methods Med Res, 27: 1230-9.

4. Royston P, Parmar MK (2013) Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. BMC Med Res Methodol, 13: 152.

5. Dong G, Mao L, Huang B, Gamalo-Siebers M, Wang J, Yu G, Hoaglin DC (2020) The inverse-probability-of-censoring weighting (IPCW) adjusted win ratio statistic: an unbiased estimator in the presence of independent censoring. J Biopharm Stat, 30: 882-99.

6. Pocock SJ, Ariti CA, Collier TJ, Wang D (2012) The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. Eur Heart J, 33: 176-82.

7. Dong G, Huang B, Verbeeck J, Cui Y, Song J et al. (2023) Win statistics (win ratio, win odds, and net benefit) can complement one another to show the strength of the treatment effect on time-to-event outcomes. Pharm Stat, 22: 20-33.

8. Dong G, Hoaglin, DC, Huang, B, et al. (2023) The stratified win statistics (win ratio, win odds, and net benefit). Pharmaceutical Statistics, 1- 9.

9. Mao L (2023) Study design for restricted mean time analysis of recurrent events and death. Biometrics.

10. Mao L, Wang T (2024) Dissecting the restricted mean time in favor of treatment. J Biopharm Stat, 34: 111-26.

11. Troendle JF, Leifer ES, Yang S, Jeffries N, Kim DY, Joo J, O'Connor CM (2024) Use of win time for ordered composite endpoints in clinical trials. Stat Med, 43: 1920-32.

12. Han K, Jung I (2022) Restricted Mean Survival Time for Survival Analysis: A Quick Guide for Clinical Researchers. Korean J Radiol, 23: 495-99.

13. Dong G, Huang B, Chang YW, Seifu Y, Song J, Hoaglin DC (2020) The win ratio: Impact of censoring and follow-up time and use with nonproportional hazards. Pharm Stat, 19: 168-77.

14. Piantadosi S. Clinical Trials: A Methodologic Perspective, 4th Edition (2024) Code and data are publicly available at https://www.wiley.com/en-us/Clinical+Trials%3A+A+Methodologic+Perspective%2C+4th+Edition-p-9781394195671#downloadstab-section.

15. Lad T, Rubintein L., Sadeghi, A (1998) The benefit of adjuvant treatment for resected locally advanced non-small-cell lung cancer. The Lung Cancer Study Group. J Clin Oncol, 6: 9-17.

16. Verbeeck J, Spitzer E, de Vries T, van Es GA, Anderson WN et al. (2019) Generalized pairwise comparison methods to ana-

lyze (non)prioritized composite endpoints. Stat Med, 38: 5641-56.

17. Buyse Test: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://cran.r-project.org/web/packages/BuyseTest/BuyseTest.pdf