Contemporary Clinical Trials 48 (2016) 119–124

Contents lists available at ScienceDirect

Contemporary Clinical Trials

journal homepage: www.elsevier.com/locate/conclintrial

The win ratio approach to analyzing composite outcomes: An application to the EVOLVE trial

Safa Abdalla [a,*], Maria E. Montez-Rath [a], Patrick S. Parfrey [b], Glenn M. Chertow [a]

[a] Stanford University School of Medicine, Palo Alto, CA, USA
[b] Health Sciences Center St. John's, Newfoundland, Canada

ARTICLE INFO

Article history:
Received 6 November 2015
Received in revised form 17 March 2016
Accepted 5 April 2016
Available online 11 April 2016

Keywords:
Win ratio
Composite endpoints
EVOLVE trial
Cinacalcet

ABSTRACT

Background: Unlike conventional time-to-event analysis of composite endpoints in clinical trials, the "win ratio" method allows for flexibility in prioritizing their components. Here, we compare the EVOLVE trial findings using the win ratio with those from time-to-event analysis.
Methods: Exposure: Randomization to cinacalcet or placebo.
Outcome: The primary composite endpoint combining all-cause mortality and non-fatal myocardial infarction, hospitalization for unstable angina, heart failure, and peripheral vascular events.
Analysis: In an unadjusted analysis, we paired each participant from the cinacalcet arm with every participant from the placebo arm within randomization strata. Pairs were classified as "winners" or "losers," according to which participant died first during the shared follow-up time, or experienced the next ranked event first. We ranked non-fatal events in two ways: 1) all ranked evenly; and 2) prioritized by their effect on health-related quality of life. The win ratio equaled the total winners divided by total losers. Further analyses were conducted where the win ratio was stratified by, or adjusted for, age.
Results: The unadjusted win ratio for the primary composite endpoint was 1.09 (95% CI 0.97 to 1.21), a statistically non-significant result which supports the primary trial result — unadjusted hazard ratio 0.93 (95% CI 0.85 to 1.02). Age-stratified analyses showed a nominally significant benefit of cinacalcet (win ratio 1.14, 95% CI 1.04 to 1.26). Ranking of non-fatal outcomes by their relative effects on quality of life did not materially alter the results.
Conclusion: The win ratio method corroborated the findings of EVOLVE based on conventional time-to-event analysis.
EVOLVE ClinicalTrials.gov number: NCT00345839.

1. Introduction

Clinical trials in nephrology are relatively scarce; in the United States, fewer than 3% of clinical trials registered between 2007 and 2010 were directly related to kidney disease. Of those, only 4% had a sample size of 500–1000 and only 1.7% included more than a thousand patients [1], underscoring the importance of statistical power for trials in this field. Composite endpoints, commonly used in clinical trials, offer the advantages of higher statistical power, a more comprehensive evaluation of treatment effect, and minimal issues with competing risks. However, conventional time-to-event analysis implicitly attaches the same level of importance to each of their components. This approach may not produce fully relevant results, particularly if one component of a composite endpoint carries more importance to patients than others or when there is heterogeneity in the treatment effect across the individual components [2,3]. For example, in a meta-analysis of 114 cardiovascular clinical trials, the pooled treatment effect on mortality and critical outcomes, which were considered to be more important, was smaller than that for less critical outcomes [4]. Effects of interventions on individual components of the composite endpoint are often included as secondary trial endpoints, but lower event rates reduce statistical power and subsequently the precision of effect estimates; in a meta-analysis of anti-platelet agent trials, the pooled effect of anti-platelet therapy on the primary composite endpoints indicated a treatment benefit that did not manifest when focusing on the all-cause mortality component only [3].

Several methods were proposed to cater for differences in relative importance or severity among composite endpoint components, while evaluating the net benefit of treatment. In a trial of thrombolytic regimens [5], a Delphi panel of experts developed *a priori* severity weights for clinical efficacy endpoints common in cardiovascular trials, using them in a Kaplan Meier analysis. O'Brien [6] proposed the global rank method where participants are ranked by their worst outcome — those with the same outcome are further ranked by the time to that outcome. The data can then be analyzed by conventional statistical methods for rank data. Buyse described the proportion in favor of treatment, based on pairwise comparisons of each participant from one

* Corresponding author.
E-mail addresses: sabdalla@stanford.edu, safa.abdalla@gmail.com (S. Abdalla).

http://dx.doi.org/10.1016/j.cct.2016.04.001
1551-7144/

randomization arm with every participant from the other arm [7,8]. The win ratio method is closely related but groups pairs into winners and losers based on which pair member had the event considered first, working from top-ranked events downwards [9]. The win ratio is also related to the global rank method [10].

Patients with end-stage kidney disease are at an exceptionally high risk of mortality and morbidity, of which cardiovascular disease is the major cause. The suite of highly correlated fatal and non-fatal clinical events that occur frequently in these patients justifies the use of composite endpoints when evaluating the effect of interventions, but raises the aforementioned concerns. The EValuation Of Cinacalcet Hydrochloride (HCl) Therapy to Lower CardioVascular Events (EVOLVE) trial was the largest trial in patients with end-stage kidney disease, with primary results based on conventional analysis. Herein, we apply the win ratio method to evaluate the unadjusted and age-adjusted effect of cinacalcet on mortality and major cardiovascular events, comparing results with conventional time-to-event analysis.

2. Methods

2.1. Overview of the EVOLVE trial

EVOLVE trial aimed to evaluate the effect of cinacalcet *versus* placebo on mortality and major cardiovascular events in patients with end-stage kidney disease with moderate to severe secondary hyperparathyroidism (sHPT) on hemodialysis. Trial design and baseline characteristics of participants have previously been published [11–13]. In brief, 3883 patients on hemodialysis with moderate to severe sHPT (intact parathyroid hormone (PTH) > 300 pg/ml) were randomized to receive either cinacalcet (N = 1948) or placebo (N = 1935), in addition to conventional therapy (typically phosphate binders and vitamin D sterols) at the discretion of the treating physician. Randomization was stratified by country and diabetes status of the participant. Pre-specified fatal and non-fatal events occurring during follow-up were adjudicated by an independent Clinical Events Classification committee. The study was approved by the Institutional Review Boards at each participating site and all participants gave written informed consent.

The primary composite endpoint was the first of any of the following: death or non-fatal myocardial infarction (MI), hospitalization with unstable angina, heart failure or peripheral vascular event, including revascularization or non-traumatic amputation. The unadjusted hazard ratio was 0.93 (95% confidence interval [95% CI] 0.85 to 1.02), while a pre-specified analysis adjusting for baseline covariates (age being the most important) showed a nominally statistically significant 12% reduction in the primary composite endpoint rate in patients randomized to cinacalcet (p = 0.008) [13].

2.2. Outcome

Our outcome for the win ratio approach was the primary composite endpoint, prioritizing mortality over non-fatal events and in our initial analysis, giving equal priority to each of the non-fatal components. We analyzed the sensitivity of the findings to the choice of ranking scheme by repeating the analyses, ranking non-fatal events in order of decreasing impact on health-related quality of life in the EVOLVE trial, as reported by Briggs et al. [14]: peripheral vascular event, heart failure, myocardial infarction and hospitalization for unstable angina.

2.3. Analysis

2.3.1. Time-to-event analysis

We first estimated unadjusted and age-adjusted hazard ratios (HR) (cinacalcet *versus* placebo) for the primary composite endpoint and for each of its components separately. We then compared them using the marginal Cox model of Wei, Lee and Weissfeld (WLW) [15] — a method for testing the null hypothesis that an intervention has the same effect on different outcomes. The analysis was stratified by country and diabetes status to account for trial design.

2.3.2. Win ratio analysis

Computing the win ratio required pairing of participants from each randomization arm. Applying the intention-to-treat principle, participants in each pair were compared during a shared follow-up time defined as the minimum of their follow-up times. Pairs were classified as winners if participants randomized to placebo died first during the shared follow-up time and losers if those randomized to cinacalcet died first. If both participants in a pair completed or exited the study before death, they were classified according to who experienced any of the other non-fatal events first. A pair was tied if a decision could not be made on whether it was a winner or a loser. This happened when it was unknown who had the event of interest first during the shared follow-up time or no event occurred in the follow-up time for either member of a pair (Fig. 1).

The win ratio is computed as the total number of winner pairs divided by the total number of loser pairs or, in other words, the proportion of winner pairs divided by the proportion of loser pairs. It can vary from 0 (no winner pairs) to infinity (no loser pairs). A win ratio larger than one indicates benefit of the treatment being evaluated and the inverse of the win ratio can be compared with the HR (although they would still be different measures that are not directly comparable in magnitude).

2.3.2.1. Unadjusted win ratio. We paired each patient randomized to cinacalcet with every patient randomized to placebo within randomization strata (total = K strata), yielding a total of 343,184 pairs. We applied the rules described above and shown in Fig. 1 to find the number of winner pairs and the number of loser pairs in each stratum. Those were summed over the strata and the overall win ratio was computed as the total number of winner pairs divided by the total number of loser pairs. To get the overall variance, we first calculated the variance for the difference between the number of winner pairs and the number of loser pairs (win difference) in each stratum and summed it over strata using Finkelstein and Schoenfeld' method [16] after confirming the assumption of the method that the censoring distribution did not vary between the cinacalcet and placebo groups. Applying the delta method to the variance of the win difference yields a reasonable approximation for the variance of the win ratio as the former divided by the squared sum of loser pairs. Formulas 1–10 below detail these calculations:

For *stratum k* of size $N_k = N_{ck} + N_{pk}$, where N_{ck} is the total number of patients in the cinacalcet group and N_{pk} is the total number of patients in the placebo group within the stratum, comparing each patient (i) with every other patient (j) defines:

$$U_{ki} = \sum_{i \neq j} u_{kij} \tag{1}$$

where $u_{kij} = -1, +1$ or 0 depending on whether patient i had the event under consideration first, did not have the event first or the comparison was tied, respectively. The variance in each stratum k was then computed as

$$v_k = \frac{N_{ck}N_{pk}}{N_k(N_k-1)} \sum_{i=1}^{N_k} U_{ki}^2 \ . \tag{2}$$

Let n_{wk} be the total number of winner pairs and n_{lk} be the total number of loser pairs in the k-th stratum. We can define the overall win ratio as

$$W = \frac{\sum_{k=1}^{K} n_{wk}}{\sum_{k=1}^{K} n_{lk}} \tag{3}$$