The win ratio approach for composite endpoints

Nantharat Apiwantanakul
Commentator
16 Aug 2024

European Heart Journal (2012) **33**, 176–182
doi:10.1093/eurheartj/ehr352

SPECIAL ARTICLE

The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities

Stuart J. Pocock*, Cono A. Ariti, Timothy J. Collier, and Duolao Wang

Department of Medical Statistics, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK

Received 13 June 2011; revised 16 July 2011; accepted 15 August 2011; online publish-ahead-of-print 6 September 2011

European Heart Journal (2020) **41**, 4391–4399
doi:10.1093/eurheartj/ehaa665

CLINICAL REVIEW
Ischaemic Heart Disease

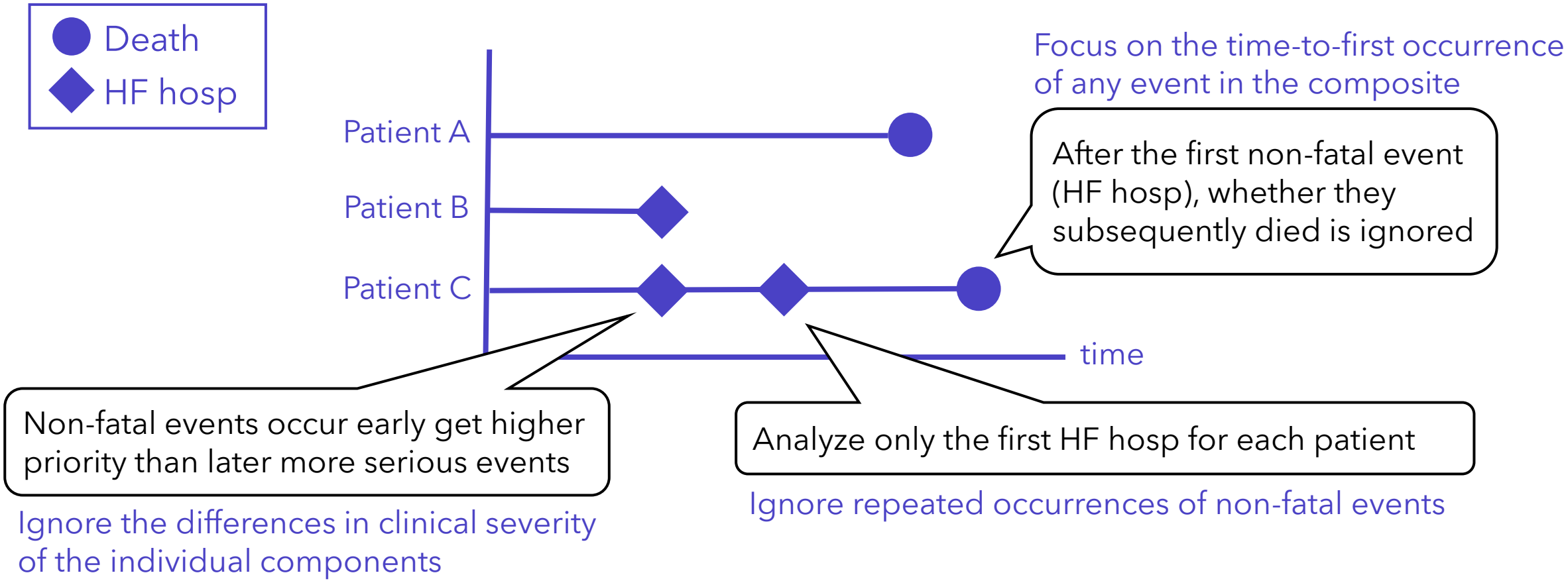
The win ratio approach for composite endpoints: practical guidance based on previous experience

Björn Redfors [1,2,3], John Gregson [4], Aaron Crowley [1], Thomas McAndrew [1], Ori Ben-Yehuda[1,2], Gregg W. Stone [1,5], and Stuart J. Pocock[4]*

[1]Clinical Trials Center, Cardiovascular Research Foundation, New York, NY, USA; [2]Division of Cardiology, NewYork-Presbyterian Hospital/Columbia University Irving Medical Center, New York, NY, USA; [3]Department of Cardiology, Sahlgrenska University Hospital, Gothenburg, Sweden; [4]Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London WC1E7HT, UK; and [5]The Zena and Michael A. Wiener Cardiovascular Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Received 14 May 2020; revised 1 July 2020; editorial decision 25 July 2020; accepted 29 July 2020; online publish-ahead-of-print 9 September 2020

Analysis of the composite outcome



Analysis of the composite outcome

Methods	Description
Conventional time to first event	Focuses on the time until the first event occurs. Ignores subsequent events Ignore the differences in clinical severity
Differentially Weighting	Assigns different weights to event types based on clinical importance
Win ratio method	Rank-based approach pairing patients based on the superiority of endpoint composites

Armstrong PW, Westerhout CM. Composite End Points in Clinical Research: A Time for Reappraisal. Circulation. 2017 Jun 6;135(23):2299-2307.

Win ratio

- was introduced in 2012 for examining composite endpoints
- **Step1** Form patient-to-patient pair
- **Step 2** Evaluate each pair
    - 'win'; if the patient on the new treatment has the better outcome
    - 'loss'; if the control patient does better
    - 'tie'; otherwise
  - Evaluate first based on the most important outcome (e.g. death) and secondly on the lesser event (e.g. HF hosp)
- **Step 3** Calculate win ratio (WR)
  - The win ratio is the number of pairs of treated-patient "wins" divided by number of pairs of treated-patient "loses"

$$WR = \frac{N_{win}}{N_{loss}}$$

Step 1 - Form patient-to-patient pair

- Unmatched approach
- Matched pairs approach

Unmatched approach

- Every patient in the Treatment group is compared with every patient in the Control group.
  - N_T = the number of patients in the Treatment group
  - N_C = the number of patients in Control group
  - N_T x N_C = all paired comparisons

Treatment arm	Control arm
302 patients	312 patients

302 x 312 = 94224 patient pairs

Matched pairs approach

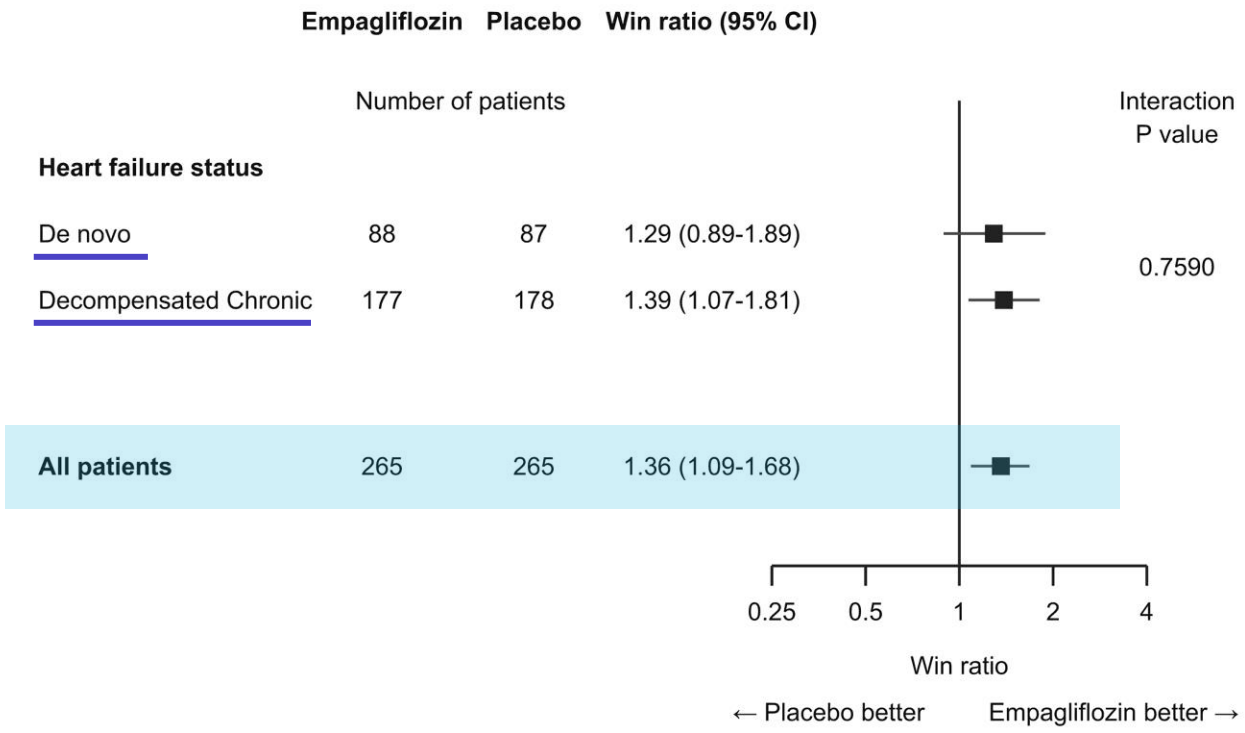
- Patients are formed into matched pairs **based on their risk profiles**
  - Avoid pairwise comparisons between patients with different baseline risk
  - Try to ensure that each matched pair of patients have a similar prognosis, thereby making all paired comparisons intrinsically fair (attempt to account for each patient's underlying risk)
- Using a **risk score**
  - Based on a pre-existing risk score obtained from an earlier study, not necessarily on the same composite outcome
  - Obtained from the trial data themselves with appropriate modelling of pre-defined predictors of the composite outcome

Matched pairs approach

- There will usually be slightly unequal patient numbers in the two groups, leaving a small number of patients unmatched in the larger group
- Example
  - 2737 patients in total
    - 1364 patients on eplerenone
    - 1373 patients on placebo
  - Get equal-sized groups by randomly removing nine patients from the placebo group
  - Risk-matches each eplerenone patient with each placebo patient using their risk scores.
- Unfortunately, experience has shown that it is difficult to objectively define the matching process in advance, and it is often not possible to match all patients.

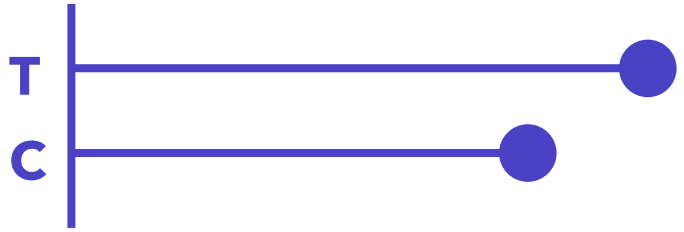
Unmatched with stratification

- Attempts to control for known prognostic variables
  - Dividing patients into strata based on prognostically meaningful variables
  - Perform pairwise comparison within the same strata
  - Count the wins and losses within each stratum
- Stratified win ratio
- The influence of the stratification variables on prognosis can be reduced



Step 2 – pairwise comparison

- Composite outcome
  - Time to death
  - HF hospitalization

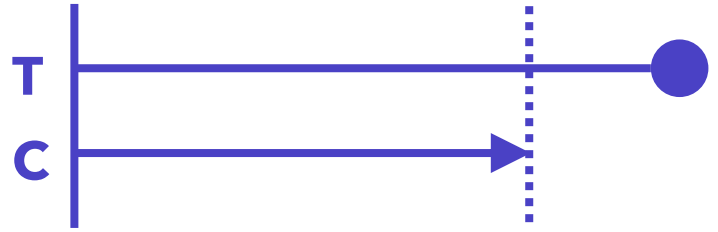


Assess who dies first?

(C died first)
T wins

Step 2 – pairwise comparison

- Composite outcome
  - Time to death
  - HF hospitalization



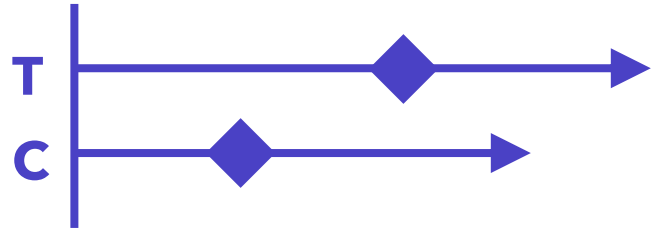
Assess who dies first?

- Treated patient had death only after the control patient was lost to follow-up
- That death should not be considered
- a **'tie'** would be declared for this comparison

Each patient pair can only be compared for the **shared follow-up duration** they both achieved

Step 2 – pairwise comparison

- Composite outcome
  - Time to death
  - HF hospitalization



Assess who dies first? → Assess who had HF hosp first?

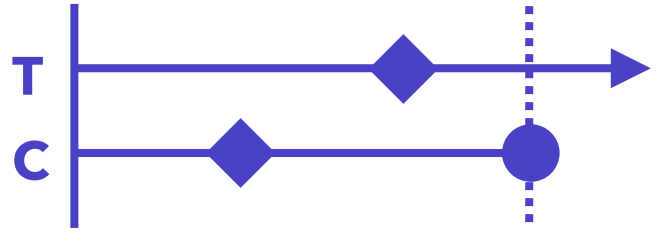
Neither of the patients dies

(C had HF hosp first)
T wins

Accounts for clinical priorities
Evaluates the component outcomes in descending order of importance until one of the pair shows a better outcome compared with the other.

Step 2 – pairwise comparison

- Composite outcome
  - Time to death
  - HF hospitalization



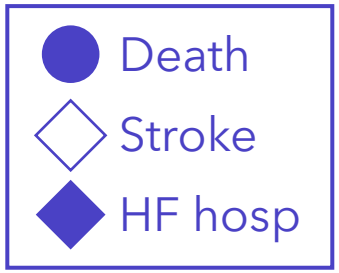
Assess who dies first? → Assess who had HF hosp first?

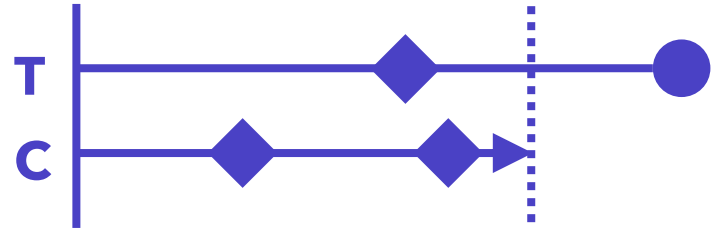
(C died first)
T wins

- Evaluates the component outcomes in descending order of importance
- Uses information from all deaths
- Ignore death that occurred after HF hosp (if in conventional time-to-first event)

Step 2 – pairwise comparison

- Composite outcome
  - Time to death
  - Stroke
  - Number of HF hospitalization





can be extended to a composite with three or more components with sensible ranking of the components

Assess who dies first? → Assess who had stroke first? → Assess who had higher numbers of HF hosp?

Neither of the patients dies

Neither of the patients had stroke

(C had more)
T wins

Can combine the time to death with the number of occurrences of HF hosp in a single hierarchical composite endpoint

Guidance on how to compare individual components

- When conducting the win ratio analysis, we **must decide whether to compare patients** in regard to
- (i) simply whether they experienced the event
  - We discourage simply comparing patients in regard to whether they had the event (option i), since ignoring information on timing or frequency of events omits important information.
- (ii) how soon they experienced the event
- (iii) how many events they experienced
- (iv) how severe the events were

Guided by clinical reason

Trial	Compare two coronary stent types	Examining the effect of two therapies for HF	Examining the effect of cerebral protection devices
Outcome	Target vessel revascularization (TVR)	Recurrent HF hospitalization	Ischemic stroke
Prioritize	(ii) **how soon** they experienced the event • Who had TVR first	(iii) **how many** events they experienced • The number of HF Hosp is strongly associated with prognosis in patients with HF	(iv) **how severe** the events were • More meaningful to compare patients with regard to the severity of the stroke
Rather than	(iii) **how many** events they experienced • Any subsequent TVR of the same vessel may be related to the second procedure rather than the study stent	(ii) **how soon** they experienced the event	(ii) **how soon** they experienced the event (iii) **how many** events they experienced

Step 2 – pairwise comparison

- Composite outcome
  - Time to death
  - Stroke
  - Number of HF hospitalization
  - Kansas City Cardiomyopathy Questionnaire (KCCQ)
  - Index of cardiac function (E.g. Left ventricular ejection fraction)

Patient reported outcomes

Pathophysiological measures

Can incorporate patient-reported outcomes or pathophysiological measures which is **quantitative variable**

- Including quantitative variables can be useful because their diversity of values means that <u>most pairwise comparisons identify a winner.</u>

Step 2 – pairwise comparison

- Composite outcome
  - Time to death
  - Stroke
  - Number of HF hospitalization
  - Kansas City Cardiomyopathy Questionnaire (KCCQ)
  - Index of cardiac function (E.g. Left ventricular ejection fraction)

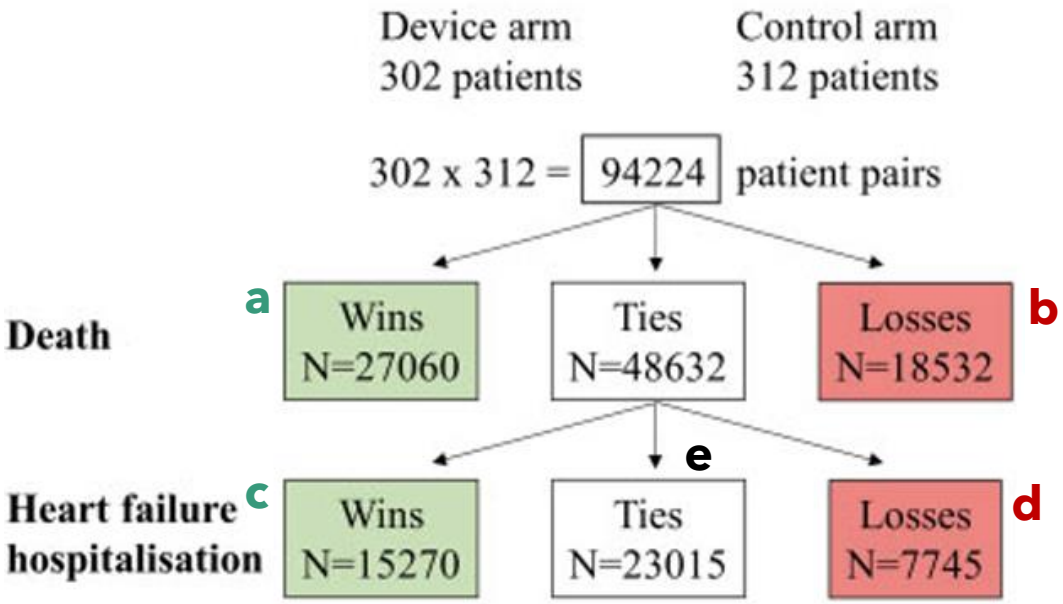
Patient reported outcomes

Pathophysiological measures

To declare a winner or loser	
Any difference	No matter how small to discriminate
Use margin of clinically meaningful quantity	Clinically meaningful win or loss Lower the number of decisions made at that level
Categorize	Patients can be classified into those that responded or worsened to therapy

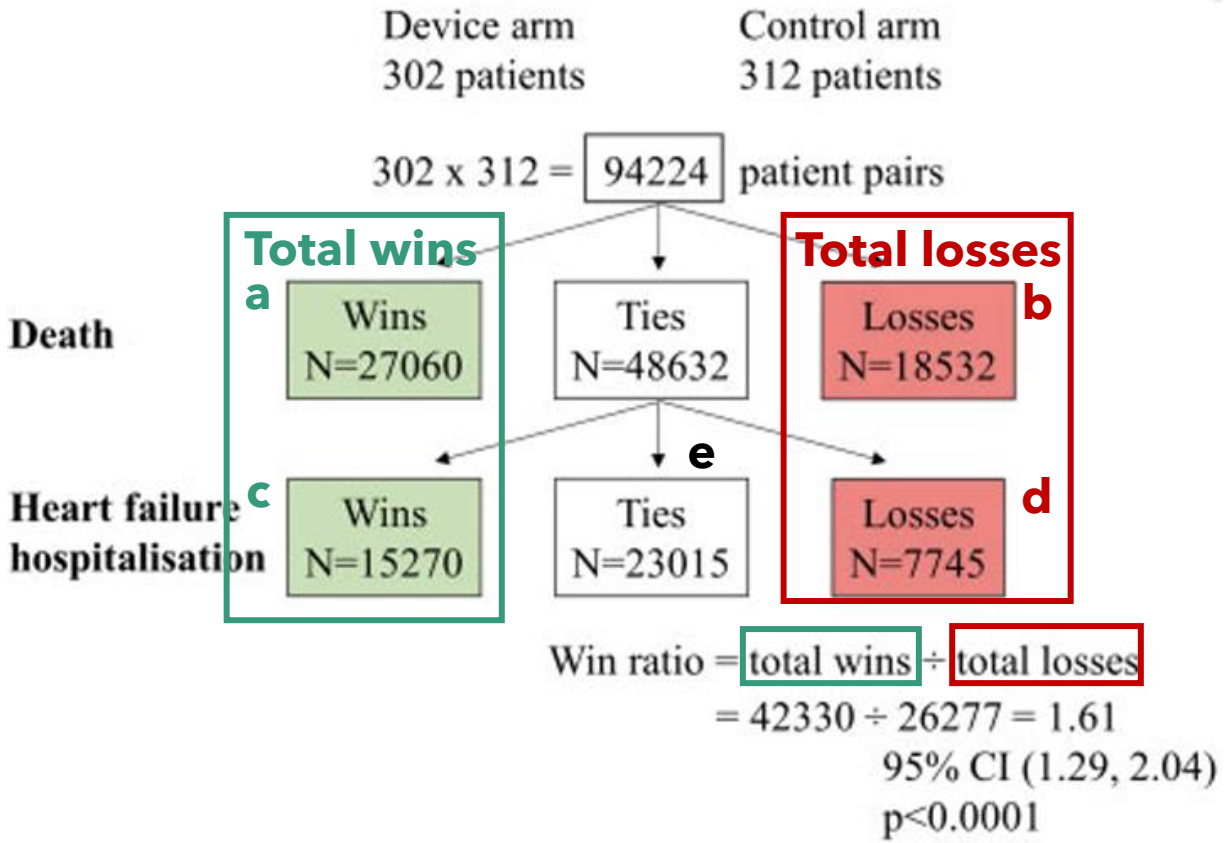
Step 2 – pairwise comparison

- Each pair is classified into one of categories
  - (a) Control patient had death first **(T wins)**
  - (b) Treated patient had death first **(T losses)**
  - (c) Control patient had HF hosp first **(T wins)**
  - (d) Treated patient had HF hosp first **(T losses)**
  - (e) None of the above **(Tie)**
    - Neither patient had an event (event-free)
    - One patient had an event, but the others' follow up time was shorter
      - 'unused' death
      - 'unused' HF hosp



Step 3 – calculate win ratio

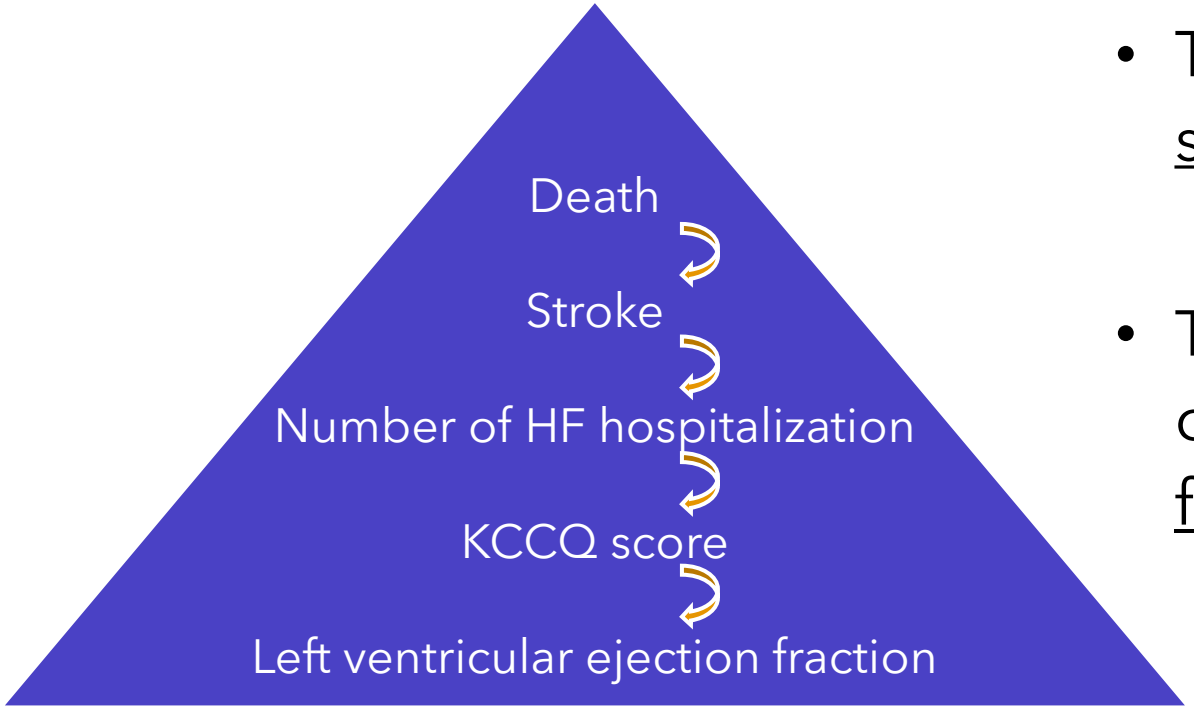
- $N_{win} = N_a + N_c$; the number of 'winners' for the new treatment
- $N_{Loss} = N_b + N_d$; the number of 'losers' for the new treatment
- Win ratio = $\frac{N_{win}}{N_{Loss}}$; overall
- WR >1 reflects a better outcome in the therapeutic group
- Patients with ties do not contribute to win ratio
- $P_{win} = \frac{N_{win}}{N_{win}+N_{Loss}}$; proportion of wins
- $P_{Tie} = \frac{N_e}{N}$; proportion of ties



Step 3 – calculate win ratio

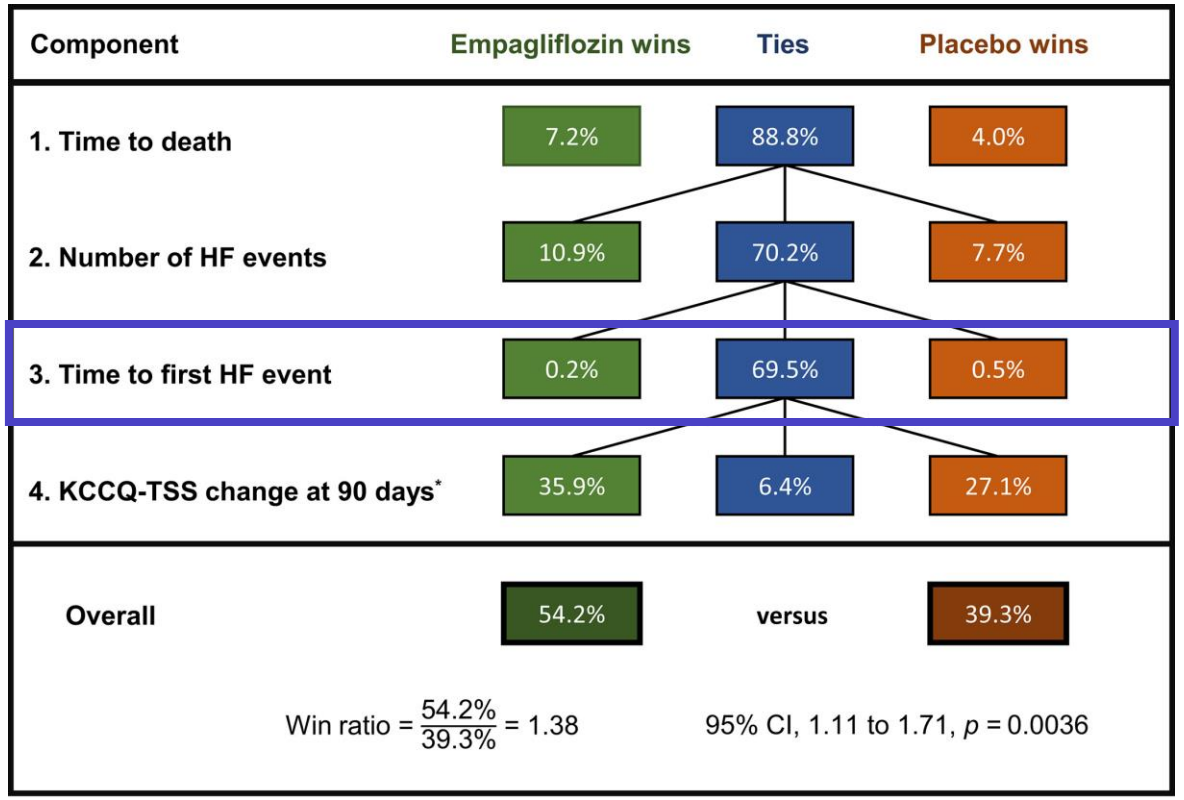
How do we interpret the value of the win ratio?

- What does a win ratio of 1.61 actually mean?
  - Win ratio $= \frac{N_{win}}{N_{Lose}}$ → Odds
    - If any two patients are compared, the odds that the treated patient is the winner is 1.61
    - Treated patient had 61% increased odds of winning
    - There are 61% more wins on the new treatment
    - Treated patient has odds '1.61 to 1' of doing better than control
  - $Prob = \frac{Odds}{1+Odds} = \frac{1.61}{1+1.61} = 0.62$
    - The probability of the treated patient wins is 0.62 (62%)



Win ratio
Composite outcome

- The estimate of the win ratio may be sensitive to the chosen order of outcomes.
  - Rank based on their clinical priorities
- The relative contribution of each component is sensitive to the duration of follow-up
  - More patients will die in a longer trial, and so the relative contribution of death will be greater.

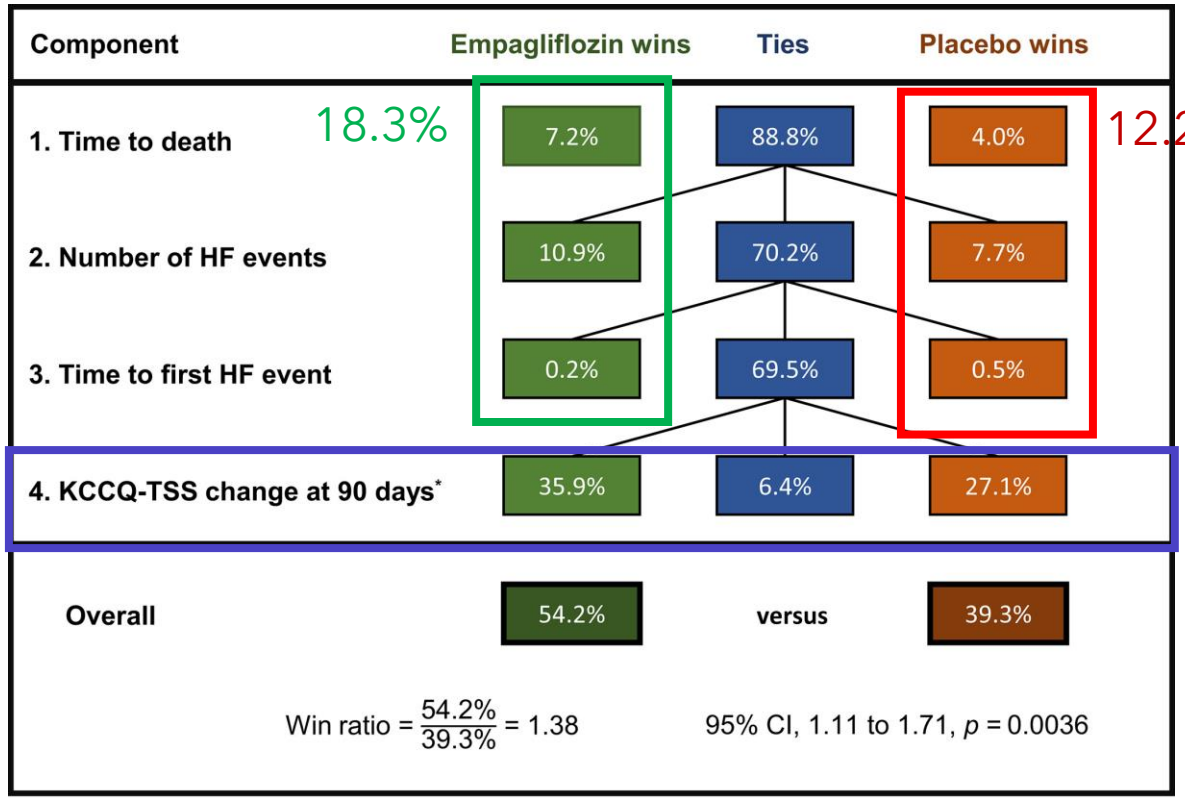


⁺ note that the predefined primary analysis is stratified (see Figure 2 and Figure 3)

* a win requires at least 5 points difference between patients

- Selecting appropriate components for the composite endpoint
  - The third level had only 0.2% and 0.5% wins on empagliflozin and placebo, with 69.5% tied.
  - This brings into question whether in hindsight this third step was worth including.
  - If a component were no treatment effect or an inverse effect, power would be adversely affected

Pocock SJ, Ferreira JP, Collier TJ, Angermann CE, Biegus J, Collins SP, Kosiborod M, Nassif ME, Ponikowski P, Psotka MA, Teerlink JR, Tromp J, Gregson J, Blatchford JP, Zeller C, Voors AA. The win ratio method in heart failure trials: lessons learnt from EMPULSE. Eur J Heart Fail. 2023 May;25(5):632-641.



Biggest contribution

+ note that the predefined primary analysis is stratified (see Figure 2 and Figure 3)

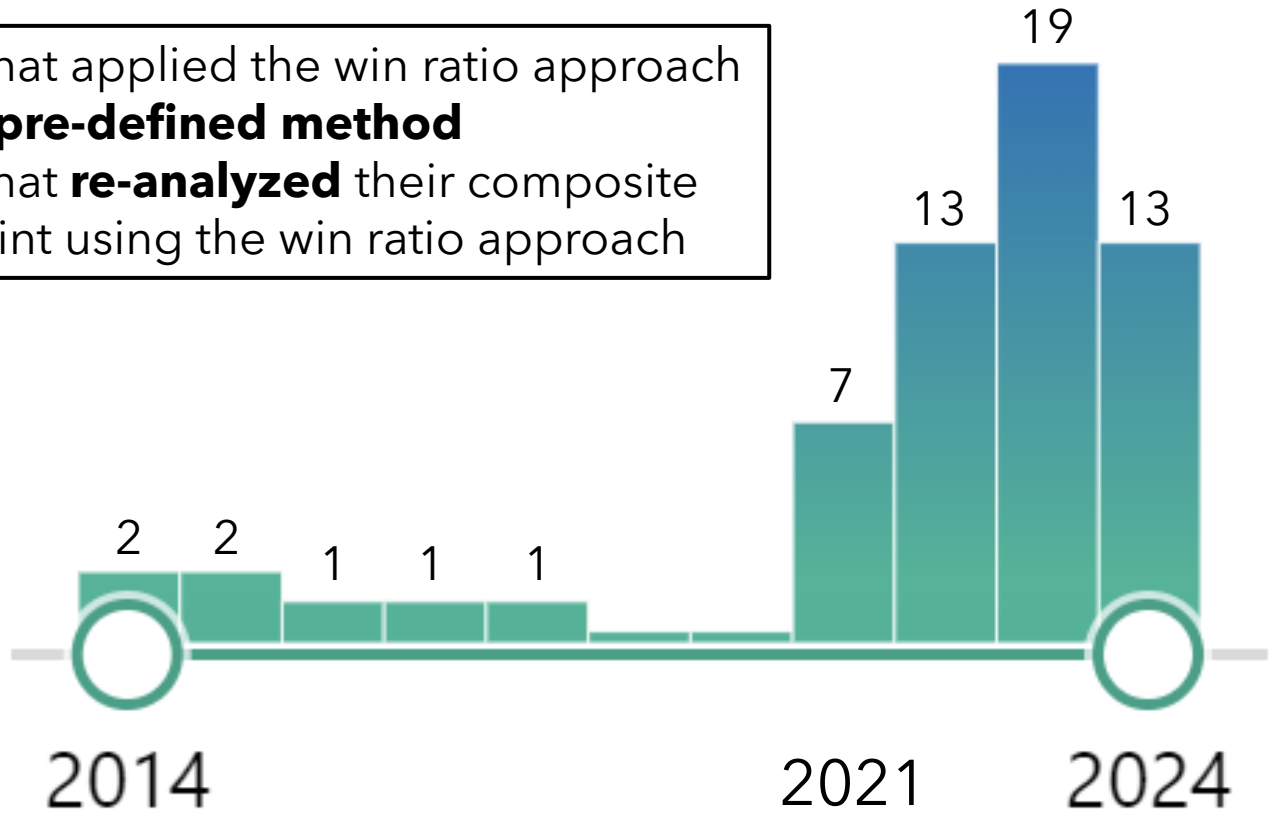
* a win requires at least 5 points difference between patients

- Step 1-4
  - Win ratio = 1.38 (**1.11**, 1.71)
- Step 1-3
  - Win ratio = 1.50 (**0.99**, 2.26)
  - underpowered to reach a definitive conclusion

Pocock SJ, Ferreira JP, Collier TJ, Angermann CE, Biegus J, Collins SP, Kosiborod M, Nassif ME, Ponikowski P, Psotka MA, Teerlink JR, Tromp J, Gregson J, Blatchford JP, Zeller C, Voors AA. The win ratio method in heart failure trials: lessons learnt from EMPULSE. Eur J Heart Fail. 2023 May;25(5):632-641.

Query	Results
Search: **"win ratio"** Filters: **Clinical Trial, Randomized Controlled Trial** ("win ratio"[All Fields]) AND (clinicaltrial[Filter] OR randomizedcontrolledtrial[Filter])	52

- Trials that applied the win ratio approach as the **pre-defined method**
- Trials that **re-analyzed** their composite endpoint using the win ratio approach



Trials that applied the win ratio approach as the **pre-defined method**

Year	Journal	Trial	Population	Treatment	Field
2021	Lancet	ACTION	Hospitalized COVID-19 patients	Therapeutic or prophylactic anticoagulation	Covid
2021	EuroIntervention	SPYRAL HTN-ON MED	patients with uncontrolled hypertension	renal denervation (RDN) vs sham control	Med
2021	Lancet Diabetes Endocrinol	DARE-19	Hospitalized COVID-19 patients	Dapagliflozin vs placebo	Covid
2022	Nat Med	EMPULSE	Patient hospitalized for acute heart failure	empagliflozin vs placebo	Cardio
2022	Eur Heart J	DIAMOND	Patients with HF with reduced ejection fraction	Patiromer vs placebo	Cardio
2022	N Engl J Med	-	Infants undergoing heart surgery	Methylprednisolone vs placebo	Cardio
2022	Circulation	COVID-PACT	Critically ill patients with COVID-19	Anticoagulation and antiplatelet therapy	Covid
2022	Eur Heart J	VIP-ACS	Acute coronary syndrome	standard-dose vs double dose influenza vaccination	Cardio
2023	N Eng J Med	STEP-HFpEF	Patients with heart failure with preserved ejection fraction	Semaglutide vs placebo	Cardio
2023	N Eng J Med	PARTNER 3	low-risk patients with severe, symptomatic aortic stenosis	Transcatheter aortic-valve replacement (TAVR) vs surgical aortic-valve replacement	Cardio
2023	N Eng J Med	TRILUMINATE Pivotal	Patients with tricuspid regurgitation	transcatheter edge-to-edge repair (TEER) vs medical therapy	Cardio
2023	J Am Coll Cardiol	PARAGLIDE-HF	Patients with mildly reduced or preserved ejection fraction and worsening heart failure	sacubitril/valsartan vs placebo	Cardio
2023	N Eng J Med	HEART-FID	Patients with hearth failure and iron deficiency	ferric carboxymaltose vs placebo	Cardio
2023	Lancet Respir Med	VT4COVID	COVID-19-related acute respiratory distress syndrome	ultra-low tidal volume (ULTV) vs low tidal volume (LTV)	ICU
2023	JACC Heart Fail	REDUCE LAP-HF II	Heart failure with preserved ejection fraction	Atrial shunt or sham procedure	Cardio
2024	NEJM Evid	DAPA-MI	patients with acute MI	dapagliflozin or placebo	Cardio
2024	N Engl J Med	ATTRibute-CM	patients with transthyretin amyloid cardiomyopathy	acoramidis hydrochloride vs placebo	Cardio
2024	N Engl J Med	STEP-HFpEF DM	Patients with obesity-related heart failure and type 2 DM	Semaglutide vs placebo	Cardio
2024	Lancet	STEP-HFpEF and STEP-HFpEF DM	Patients with obesity-related heart failure with preserved ejection fraction	Semaglutide vs placebo	Cardio
2024	JAMA	DEFENDER	Critically ill patients with acute organ dysfunction	Dapagliflozin vs standard ICU	ICU
2024	Eur J Heart Fail	BeAT-HF	Patients with heart failure and a reduced ejection fraction	Baroreflex activation therapy vs medical management	Cardio
2024	JAMA Cardiol	MESSAGE-HF	Patients after HF hospitalization	telemonitoring strategy vs standard care	Cardio

Trials that **re-analyzed** their composite endpoint using the win ratio approach

Year	Journal	Trial	Population	Treatment	Field
2017	Eurointervention	SYNTAX	De-novo 3-vessel or left main coronary artery disease	CABG vs PCI for multivessel CAD	Cardio
2018	J Clin Epidemiol	Kidney transplant trial	Kidney transplant	ramipril treatment vs. placebo	Nephro
2021	J Am Heart Assoc	EMPA-REG OUTCOME	patients with type 2 diabetes mellitus and established cardiovascular disease	Empagliflozin vs placebo	Cardio
2022	Eur J Heart Fail	TOPCAT	Patients with HF and a preserved ejection fraction	Spironolactone vs placebo	Cardio
2022	Eur J Heart Fail	PARADISE-MI	Patients with acute MI	sacubitril/valsartan vs ramipril	Cardio
2023	Nutrients	VITAL	healthy older adults	Marine n-3 fatty acid supplementation vs vitamin D3	Med
2023	Thromb Haemost	mAFA-II	Atrial fibrillation patients	mAFA intervention vs usual care	Cardio
2024	Clin Infect Dis	MERINO	Ceftriaxone-Nonsusceptible Escherichia coli or Klebsiella pneumoniae Bloodstream Infections	piperacillin-tazobactam vs meropenem	Infection
2024	Circ Cardiovasc Qual Outcomes	ENGAGE AF-TIMI	Patients with atrial fibrillation	Edoxaban vs warfarin	Cardio
2024	Nat Med	DAPA-HF and DELIVER	patients with heart failure and reduced ejection fraction and mildly reduced or preserved ejection fraction	dapagliflozin vs placebo	Cardio
2024	J Crit Care	STARRT-AKI	Acute kidney injury	Accelerated vs standard renal replacement therapy (RRT)	Nephro
2024	J Crit Care	RENAL	Critically ill patients	Two different RRT doses	Nephro

Table 2 **Trials that have applied the win ratio approach as the pre-defined method to analyse their primary composite endpoint**

Trial	Population	Randomized treatment	Hierarchical level	Primary composite endpoint	Win ratio (95% CI)
ATTR-ACT[14]	Transthyretin amyloid cardiomyopathy	Tafamidis vs. placebo	2	All-cause mortality > number of heart failure hospitalizations	1.70 (1.26–2.29)
CHART-1[16]	LVEF ≤35%	Cardiopoietic stem cells vs. placebo	6	Time to death > *N* of HF events > MLHFQ score ≥10-point improvement > 6MWT improvement ≥40 m > LVESV change ≥15 mL > LVEF change ≥4%.	1.17 (0.89–1.55)
TAVR-UNLOAD[18]	Moderate AS and reduced LVEF	TAVR vs. medical therapy	4	Time to death > disabling stroke > hospitalizations due to HF, aortic valve disease, or non-disabling stroke > change in KCCQ relative to baseline	Ongoing
RELIEVE-HF (NCT03499236)	NYHA class III and IV heart failure	Inter-atrial shunt vs. medical therapy	4	Time to death > time to heart transplant or LVAD > number and time of hospitalizations due to HF > improvement in 6MWT	Ongoing
CARILLION (NCT03142152)	Functional MR associated with HF	Carillion implant vs. medical therapy	5	Death > cardiac transplantation or LVAD > percutaneous or surgical mitral valve intervention > time to first HF hospitalization > improvement in 6MWT	Ongoing
ACTIVE (NCT03016975)	Functional MR associated with HF	Cardioband implant vs. medical therapy	4	Death > number of HF hospitalizations > improvement in 6MWT > improvement in KCCQ	Ongoing
PARACHUTE-HF (NCT04023227)	HF with reduced LVEF caused by chronic Chagas disease	Sacubitril/valsartan vs. enalapril	3	CV death > HF hospitalization > relative change in NT-proBNP from baseline to week 12	Ongoing

6MWT, 6-min walk test; AS, aortic stenosis; HF, heart failure; KCCQ, Kansas City Cardiomyopathy Questionnaire; LVAD, left ventricular assist device; LVEF, left ventricular ejection fraction; LVESV, left ventricular end-systolic volume; MLHFQ, Minnesota Living with Heart Failure Questionnaire; MR, mitral regurgitation; TAVR, transcatheter aortic valve replacement. > Designates the order of the win ratio hierarchy, which decreases from left to right.

Table 3 Trials that have re-analysed their primary composite endpoint using the win ratio approach

Trial	Population	Randomized treatment	Primary composite endpoint	Outcome	
				Primary analysis	Win ratio
PARTNER B*[6]	Severe symptomatic aortic stenosis	TAVR vs. OMT	Death, or hospitalization due to valve- or procedure-related clinical deterioration	1/HR 2.17 (1.69, 2.86)	1.87 (1.35, 2.54)

$\frac{1}{HR}$ **= 2.17 → HR = 0.46**

Patients in the new treatment group were 54% less likely to die or hosp than patients in the control group

WR = 1.87

The odds that the treated patient is the winner is 1.87

Patient on new treatment has 61% increased odds of winning

Treatment group is better	
HR < 1 $\frac{1}{HR}$ > 1	WR >1

Sample size for win ratio

Table 4 Examples of sample size estimation for the win ratio

Follow-up	First tier: death			Second tier: number of CVH			Third tier: reduction in KCCQ			Power (%)	Sample size
	Control rate (per year) (%)	Hazard ratio	Decisions[a] (%)	Control rate (per year)	Rate ratio	Decisions[a] (%)	Control, mean ± SD	Treatment, mean ± SD	Decisions[a] (%)		
3 years[b]	21.5	0.75	67.8	0.70	0.75	32.1		No third tier		80	1050
1 year[c]	21.5	0.80	47.1	0.70	0.80	52.9		No third tier		80	1948
1 year[c]	21.5	0.80	31.0	0.70	0.80	36.6	5 ± 12.0	0 ± 12.0	32.4%	80	590

[a]The decisions (%) columns refer to the percentage of all non-tied patient pairs that have a winner/loser at that level in the hierarchy.
[b]Median follow-up of 3 years; variable between 2.5 and 3.5 years.
[c]Fixed follow-up of 1 year.
CVH, cardiovascular hospitalization; KCCQ, Kansas City Cardiomyopathy Questionnaire; SD, standard deviation.

1284 **Cox PH**

▲ increase

▼ reduce

- Examine the treatment effect at 1 year
- Expected treatment effect is smaller (25%→20%)

Increase sample size

Include an additional endpoint in the hierarchical composite

Reduce sample size

Win ratio – Take home messages

- Hierarchical structure
  - accounts for clinical priorities
- Flexibility
  - can analyze composites composed of time-to-event, recurrent event, continuous, and/or categorical outcomes
- Statistical power
  - using all available information contained in the component outcomes

Advantages	Challenges
1. All Key Elements Included The win ratio recognizes all events, not just the first one, e.g. a death after a non-fatal event gets included in the analysis	**1. Lack of Familiarity** The win ratio is a relatively new statistical method: *This article should facilitate a better understanding of the concept and its potential value.*
2. Clinical Priorities Recognized The win ratio forms the component outcomes into a hierarchy based on their relative clinical importance, e.g. death gets top priority.	**2. Statistical Software** Calculation of the win ratio (and its CI and p-value) requires statistical programs being readily available: *We provide links to such software.*
3. Repeat Events Easily Incorporated The win ratio can be readily extended to account for recurrent events (e.g. hospitalizations) without statistical complexity.	
4. Non-Event Outcomes can be Included The win ratio can be extended to include visit-related items, e.g. quality of life scores and physiological measures.	**3. Determining Sample Size** Power calculations for the win ratio entail simulations: *We have created new software to facilitate this task*
5. Conceptually Straightforward Counting up the "winners" and "losers" across all pairwise comparisons is a simple concept, compared to explaining what a hazard ratio means.	

Take home figure The win ratio method's key advantages and challenges compared with conventional methods for composite outcomes in randomized controlled trials. CI, confidence interval.