Journal of Clinical Epidemiology 196 (2026) 112315

ORIGINAL RESEARCH

The win ratio in contemporary clinical trials: growing adoption and interpretation challenges—a meta-epidemiological study

Qize Li[1], Yingxuan Zhu[1], Yanyan Zhao, Chilie Danzeng, Duoer Wang, Ru Jiang, FeiLong Chen, Xiaoying Lin, Lingqian Hu, Yang Wang*

Medical Research and Biometrics Center, Chinese Academy of Medical Sciences & Peking Union Medical College/National Center for Cardiovascular Diseases, Fuwai Hospital, Beijing 100037, China

Accepted 1 May 2026; Published online 12 May 2026

Abstract

Objectives: The win ratio (WR) has been used with increasing frequency in clinical trials with hierarchical composite end points. This meta-epidemiological study aims to systematically evaluate the contemporary use of the WR in clinical trials. Specifically, we sought to (1) describe temporal trends in WR application, (2) quantify its use across therapeutic areas, and (3) examine component directional contributions and conclusion consistency between analyses summarized by the WR and the original analyses.

Methods: We conducted a systematic review (PRISMA 2020) of PubMed, Embase, and Web of Science (search date: 26 March 2025). Eligible studies were published clinical trials or reanalyses that summarized treatment effects using the WR based on individual data. Studies were classified by application type (prespecified vs reanalysis) and mapped to International Classification of Diseases (ICD-10) categories. After inclusion, we further screened for studies with comparable data, then performed two comparisons: first, the component directional contribution between end point component and the overall WR result; and second, the conclusion-level consistency between each WR-reported analyses and the conclusions of the original trials.

Results: A total of 90 studies were included: 35 prespecified end point analyses and 55 reanalyses. Circulatory system trials accounted for about 80% of the applications. We included 57 studies for comparison, and discrepancies between component-level results and overall conclusions were common. In prespecified end point analyses, 36 of 139 components (25.9%) had directional discordance or neutrality with the overall WR point estimate, compared with 17.0% in reanalyses, although the difference was not statistically significant (generalized linear mixed-effects model with study as a random effect, $P = .114$). Among the 35 studies included for conclusion-level comparison, 22 reached WR-reported conclusions that were consistent with those of the original trials. Nearly all observed discrepancies in conclusions were associated with changes in the selection or prioritization of end point components.

Conclusion: The use of the WR in clinical research has steadily increased and is expanding beyond cardiovascular trials. Hierarchical composite end points constructed using the WR may differ from conventional composites in component selection, and directional discordance among components is not uncommon. In WR-reported reanalyses, changes to end point components may influence clinical interpretation. Careful attention to the appropriate use of the analytical framework, end point definition, and result interpretation is essential to ensure the interpretability and validity of WR-reported analyses. © 2026 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Keywords: Win ratio; Hierarchical composite end points; Clinical trial methodology; Generalized pairwise comparisons; Meta-epidemiology; End point selection

1. Introduction

The win ratio (WR) is being used more frequently in clinical trials, since its introduction in 2012 as an alternative to traditional composite outcomes [1]. Traditional composite end points are limited to components of the same type and often present interpretative challenges, particularly when weighing the relative importance of individual components. WR offers a statistical framework for

Funding: National Science and Technology Innovation 2030 Fund [2023ZD0505606].

[1] Qize Li and Yingxuan Zhu contributed equally to this work and shared first authorship.

* Corresponding author. Medical Research and Biometrics Center, Fuwai Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College/National Center for Cardiovascular Diseases, No. 167 Beilishi Road, Xicheng District, Beijing 100037, China.

E-mail address: wangyang@mrbc-nccd.com (Y. Wang).

https://doi.org/10.1016/j.jclinepi.2026.112315

0895-4356/© 2026 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

What is new?

Key Findings

- The use of the win ratio (WR) in clinical trials has increased, with most applications found in cardiovascular research. Directional discordance or neutrality between individual components and the overall WR result was common. Most conclusion-level discrepancies were associated with changes in end point selection or prioritization.

What this adds to what is known?

- This meta-epidemiological study systematically maps contemporary WR applications by application type and therapeutic area. It further links component-level directional discordance and conclusion-level consistency to changes in hierarchical end point structure.

What is the implication and what should change now?

- WR-reported analyses should clearly justify end point component selection and prioritization. Trial reports should transparently present level-specific win-loss results and inferential methods to support clinical interpretation.

analyzing hierarchical composite end points composed of mixed outcome types (eg, time-to-event, binary, continuous), and can also accommodate recurrent events [2,3], while explicitly incorporating prespecified clinical prioritization of components, thereby allowing the inclusion of patient-reported outcomes (PROs) and laboratory measures—an essential advancement for patient-centered trial design [4].

The flexibility in end point types, intuitive interpretability of results, suitability for analyses that violate conventional parametric assumptions, and improved information efficiency and statistical power have all contributed to the growing interest in the WR [5,6]. WR has often been used alongside the original Finkelstein–Schoenfeld (FS) test, and is now a widely used effect measure within hierarchical end point analyses under the generalized pairwise comparison framework [7,8].

Although the WR has gained considerable attention for analyzing prioritized outcomes, its real-world application has not been systematically quantified. In particular, its application may potentially influence clinical decision-making by altering the interpretation of trial results—especially when the incorporation of additional outcome components changes the clinical meaning of composite end points. In some cases, the WR has been used to introduce new components without adequately addressing their impact on the overall clinical conclusions.

This meta-epidemiology research aims to address these gaps by providing a comprehensive synthesis of how and where the WR has been applied in contemporary clinical trials. Specifically, this research aims to (1) describe the time trends of WR application in clinical trials; (2) quantify its application landscape across therapeutic domains; and (3) focus on component-level contributions and, where applicable, assess the consistency between conclusions from analyses summarized by the WR and those from the original statistical analyses, thereby evaluating the potential impact of methodological adaptations on clinical interpretation.

Collectively, this work seeks to offer an overview of the contemporary applications, methodological considerations, and practical challenges associated with using the WR in clinical trials.

2. Method

2.1. Systematic review of WR applications

This review was conducted in accordance with the PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [9]. A systematic literature search was conducted in three major electronic databases—PubMed, Embase, and Web of Science—on March 26, 2025. The search term was restricted to "win ratio," with no limitations on language or publication date. After deduplication, titles and abstracts were screened independently by two reviewers (Q.L. and Y.Z.), followed by full-text assessment. Disagreements about inclusion or exclusion were resolved by discussion; if unresolved, a third reviewer adjudicated. The following eligibility criteria were applied:

Inclusion criteria:

1. Clinical trials in which the WR method was applied to the analysis of trial results, including use in prespecified primary/secondary end point analyses and reanalyses (including subgroup analyses, pooled analyses, and long-term outcome assessments).
2. Analyses must be based on individual-level participant data, derived from either single trials or pooled datasets.

Exclusion criteria:

1. Methodological or theoretical studies and reviews;
2. Clinical trial protocols;
3. Conference abstracts and commentaries;
4. Observational studies, including cohort studies, registry studies, and cross-sectional studies.

Following systematic literature screening and selection, records were classified into two categories based on study

design and the specific application of the WR: (1) use of the WR to analyze prespecified end points (Published in major publications; excluding sensitivity and exploratory analyses) (2) reanalyses of completed trials using the WR, including subgroup analyses, pooled analyses, and long-term outcome assessments, defined as studies analyzing long-term follow-up outcomes not included in the original primary trial publication.

To ensure accurate identification and avoid duplication, the clinical trials referenced in each study were further tracked using unique trial identifiers, such as ClinicalTrials.gov registration numbers (NCT numbers) or other relevant registry codes.

In addition to methodological categorization, studies were also classified according to their therapeutic areas, aligned with the major chapters of the International Classification of Diseases (ICD-10) classification system. We also examined the application of the WR in long-term analyses, with a focus on whether the WR-reported hierarchical end point analysis was newly introduced compared with the original trial publications.

2.2. Data extraction and end point characterization

After completing the review and classification of the included studies, we further extracted detailed information regarding the use of the WR, focusing primarily on efficacy end point analyses, including trial characteristics and sample size (number randomized), end point components, win−loss counts by hierarchical level (when reported), and study conclusions. The characterization of end points was based on two aspects: changes in their composition (eg, the number of components and the types of newly added components), and the presence of a hierarchical structure. We regarded the end points as equivalent when the original study employed a composite end point and the subsequent hierarchical end point analysis summarized by the WR retained an identical set and number of underlying clinical components, even if their operational definitions or expressions were modified. This scenario was defined as ‘‘Composite Refined to Hierarchical (no changed in components)’’. We did not differentiate between different expressions of the same component (ie, (i) whether the event occurred, (ii) how soon the event occurred, (iii) how many events occurred, or (iv) how severe the events were) [4], and treated them as equivalent. For example, ‘‘any hospitalizations’’ was considered equivalent to ‘‘time to hospitalization’’. Component outcomes in traditional composite end points are typically expressed in a uniform format, such as binary or time-to-event outcomes, thereby restricting end point structures, whereas WR allows greater flexibility in the types of outcomes included. Given the common use of such transformations when moving from traditional composite end points to hierarchical composite end points, we assumed equivalence for the purpose of our assessment.

2.3. Component directional contribution

To evaluate study end points, we considered only the first WR-based composite end point reported in each study, typically corresponding to the primary analysis, without considering sensitivity analyses or alternative end point definitions. We defined directional contributions as concordant, discordant, or neutral [8]. We compared the numbers of wins and losses (absolute counts were used as the first criterion when available, followed by percentages) for each hierarchical level to determine whether the direction of each component was aligned with the direction of the overall WR point estimate. If aligned, the component was classified as exhibiting directional concordance; if opposite, as exhibiting directional discordance; and if wins and losses were equal, as directionally neutral. During data extraction, we found that not all studies reported win−loss counts for each level. When such information was unavailable, our principle was to infer which group the individual component favored.

After completing agreement assessments, we quantified the win contribution of hierarchical components by calculating the proportion of wins attributable to each level among all wins:

$$C_k = \frac{W_k}{\sum_{m=1}^{L} W_m}$$

where W_k denotes the number of wins determined at the k-th hierarchical level and L represents the total number of hierarchical levels. In the present analysis, we specifically compared the contribution of the highest-priority component (C_1) with that of the lowest-priority component (C_L).

2.4. Conclusion-level consistency

Conclusion-level consistency (within the same study) was defined as whether the conclusions derived from WR-reported hierarchical end point analyses were consistent with those from the original studies using their original analytic methods. To assess the consistency of conclusions, we adopted the concept of ‘‘regulatory agreement’’ [10].

In assessing conclusion-level consistency, considering the possibility that reanalyses may involve different approaches, such as matched vs. unmatched comparisons, or include various types of sensitivity analyses, we categorized the agreement between conclusions into three groups: agree, disagree, and possibly disagree. WR-based results were classified as agree if they supported a consistent conclusion, and as disagree if they did not. The ‘‘possibly disagree’’ category was assigned when different analytic strategies or component selections (eg, matched vs unmatched comparisons, or a 3-component vs 4-component end point) yielded divergent conclusions. Rather than testing methodological robustness, we aimed to assess whether the use of the WR resulted in meaningful shifts

in the direction or strength of clinical conclusions, particularly from a regulatory or decision-making perspective.

The applicable scope of the two types of consistency evaluation was as follows:

For WR-reported analyses of prespecified secondary end points (Published with major publications), we first assessed component directional contribution and then compared these results with the original primary end point conclusions.

For WR-reported reanalyses of the original trial population (excluding pooled analyses, subgroup analyses, and long-term follow-up), we again first examined component directional contribution and then reviewed each original trial to determine whether the WR-based findings remained consistent with the original trial conclusions.

2.5. Statistical analysis

For descriptive analyses, categorical variables were summarized as frequencies and percentages, and continuous variables were summarized as medians (IQR), because normality testing indicated that none of the variables were normally distributed. Publication year was summarized as a range (minimum—maximum). In addition, when comparing the proportion of different types of contributions across article types, because each study could contribute multiple

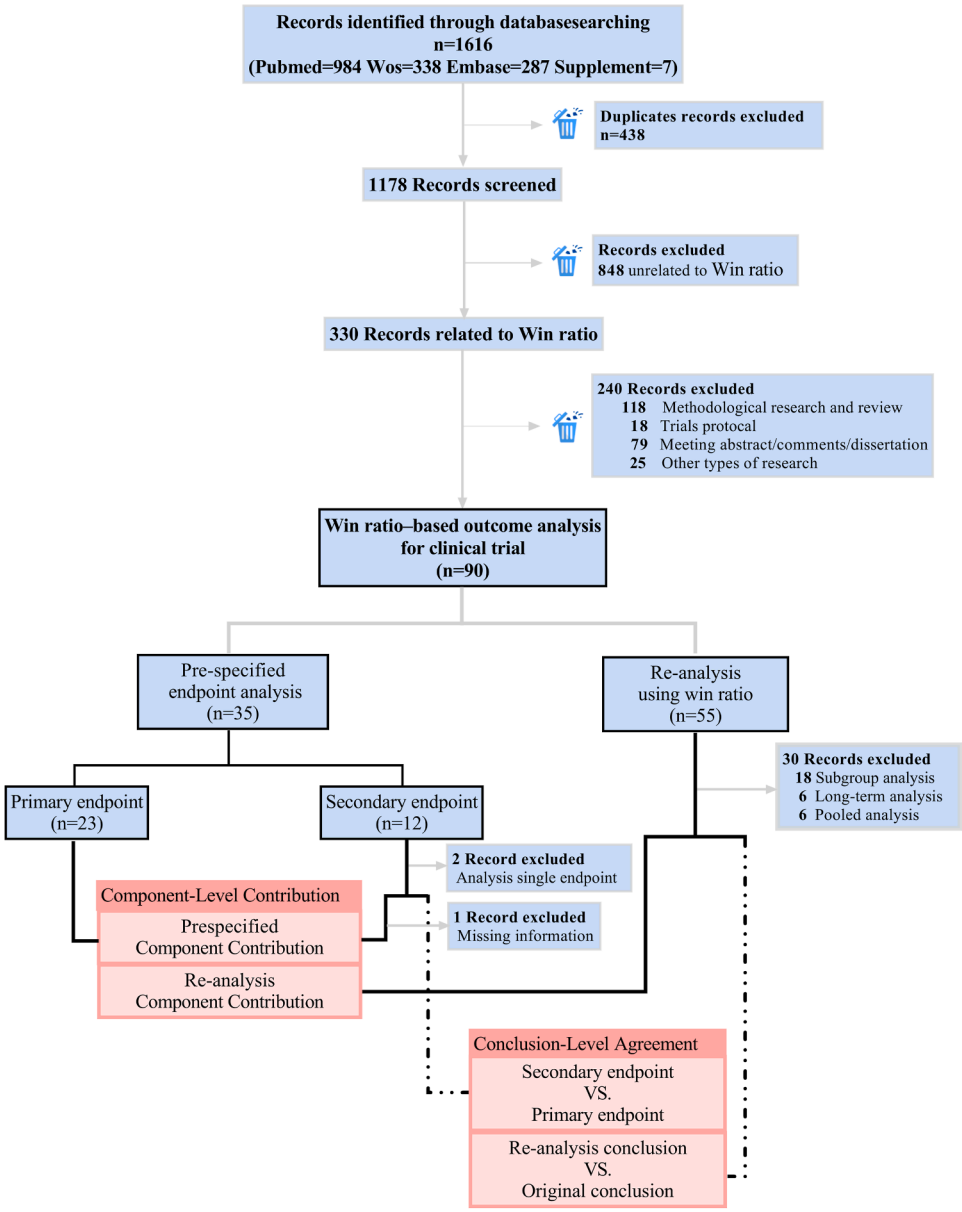


Figure 1. Flowchart of study selection and scenario establisher.

components, we used generalized linear mixed-effects models with study as the clustering unit. For the comparison between the highest-priority and lowest-priority components, we used the exact McNemar's test. The two-sided significance level was set at 0.05. Data collection and organization were performed using Excel 2021, with SAS 9.4 (SAS Institute Inc, Cary, North Carolina, USA) used for supplementary analyses and R 4.2.1 software employed for data visualization.

3. Results

3.1. Summary of WR applications

As shown in Figure 1, a total of 1609 records were identified through systematic searches of PubMed, Embase, and Web of Science. In addition, one study published after the search date was manually added, and six further studies were added through cross-checking during data extraction, bringing the total to 1616 records.

Following full-text review and application of inclusion and exclusion criteria, 90 studies were ultimately included and categorized based on their analytical focus. Among them, 35 studies analyzed prespecified end points, each corresponding to a unique clinical trial. Fifty-five studies were reanalyses of clinical trials, covering overall trial populations, prespecified secondary analyses that were not reported with the primary results, specific subgroups, as well as pooled and long-term analyses. In addition, one sensitivity analysis and one exploratory analysis of a prespecified end point were included.

Figure 2 illustrates the temporal trends in clinical studies applying the WR, categorized by type of application, from 2014 to the date of literature retrieval (March 26, 2025). The results reveal an overall upward trend in the number of studies using WR for outcome analysis over time, with a noticeable acceleration after 2020, although the formal test for trend was not statistically significant. All categories demonstrate increasing uptake, reflecting the growing recognition and acceptance of the WR across diverse clinical research settings. It is also worth noting that the data for 2025 only cover the first quarter; therefore, the total number of studies for the full year is likely to be higher, suggesting that the application of the WR will continue to rise. Among the 23 studies (each corresponding to a distinct clinical trial) using WR to summarize the primary end point, the trial information is shown. One study was a secondary analysis conducted after the primary end point did not meet its prespecified hypothesis; we categorized this study as using the WR to summarize secondary end points [11].

Figure 3 presents the distribution of disease categories, classified according to the major chapters of ICD-10, based on the primary objectives and primary end points of the clinical trials that applied the WR method.

Across all categories, diseases of the circulatory system (ICD-10 Chapter IX) overwhelmingly dominate the use of the WR, accounting for 78.3% (18/23) in studies analyzing prespecified primary end points, 75.0% (9/12) in those analyzing secondary end points, and 85.5% in reanalysis. Although less common, applications of the WR method were also observed in other ICD chapters, such as neoplasms, respiratory diseases, nervous system disorders,

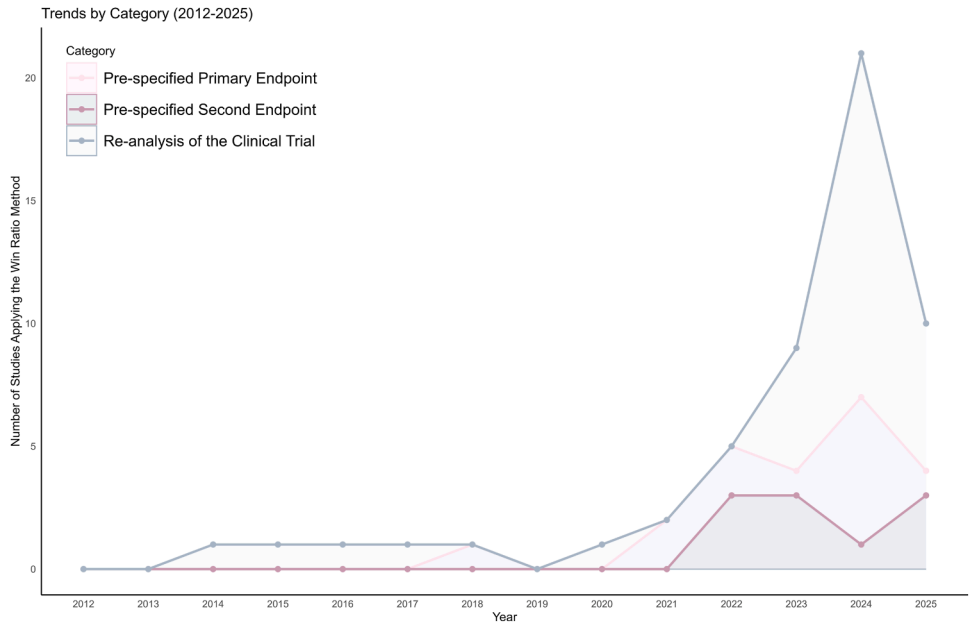


Figure 2. Temporal trends in the application of the win ratio in clinical trial outcome analyses.

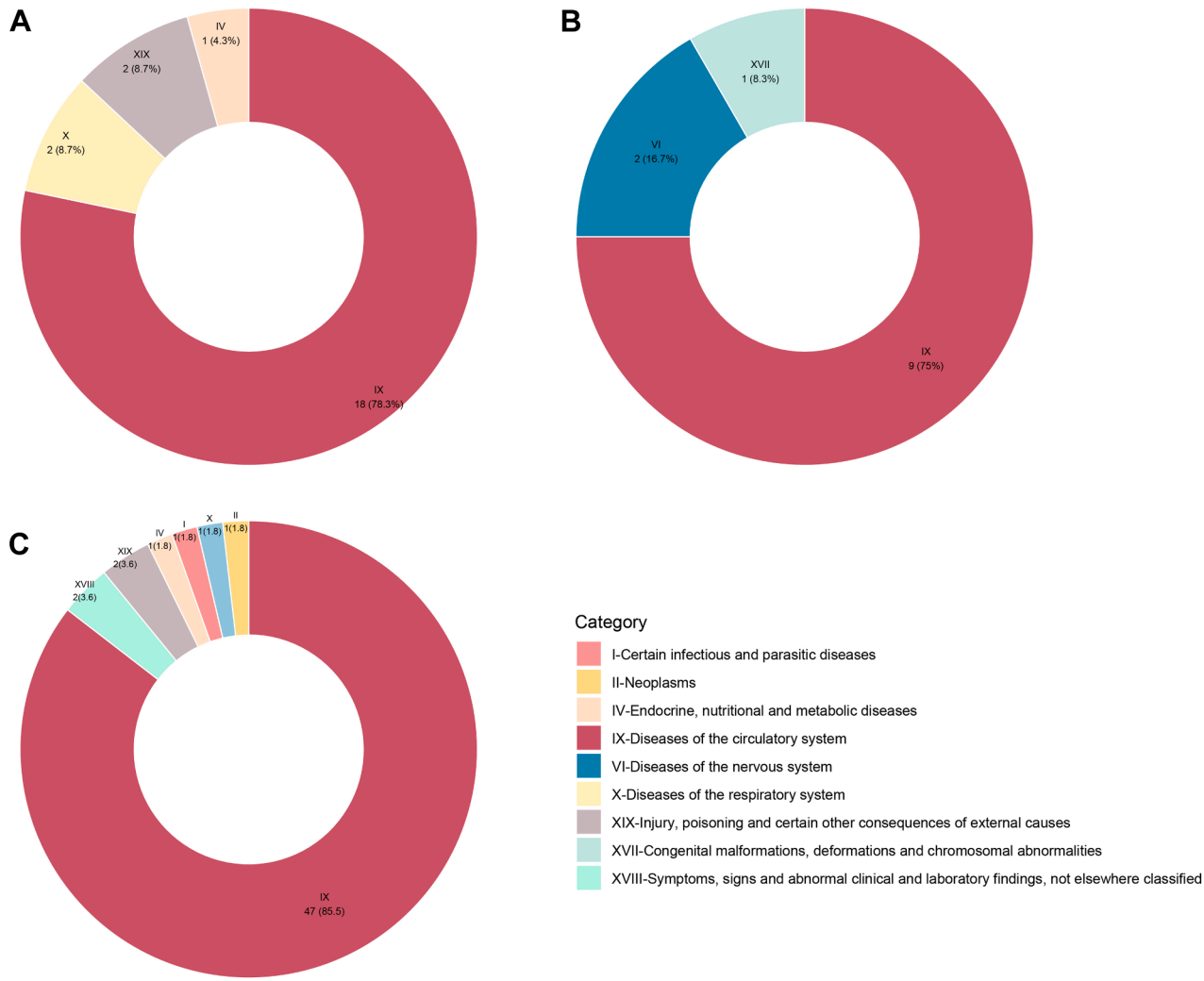


Figure 3. Distribution of disease categories applying the win ratio based on ICD-10 Classification. A, Prespecified analyses—primary end points. B, Prespecified analyses—secondary end points. C, Reanalyses. ICD-10, International Classification of Diseases.

endocrine and metabolic diseases, and infectious diseases—primarily within the context of reanalyses, each representing a smaller proportion of the overall studies.

For studies analyzing long-term outcomes using the WR (6 studies involving seven clinical trials), WR was newly introduced in four cases. The remaining two studies used the WR consistently across both initial and follow-up publications.

3.2. Characteristics of studies included in the assessment

After reviewing the application trends of the WR method, we examined the study end points and assessed component directional contribution and conclusion-level agreement (including all studies classified as prespecified analyses and reanalyses, excluding subgroup, pooled, and long-term follow-up analyses because these change the study population or follow-up status). In those trials applying the WR, the majority of analyses were performed on hierarchical composite end points. Two studies applied the WR to a nonhierarchical composite end point. These two studies were excluded from subsequent comparisons. The characteristics of all studies included in the comparisons are presented in Table 1.

Among the 33 prespecified studies (each an independent trial) published between 2018 and 2025, almost all were randomized (97.0%) and fewer than half were blinded (48.5%). The 25 reanalysis studies, conducted between 2015 and 2025 with no duplication of trial populations, were 96.0% randomized and 64.0% blinded. The median number of composite end point components was 4.0 (3.0−6.0) in prespecified studies and 3.0 (3.0−4.0) in reanalyses. The Finkelstein−Schoenfeld test was the most frequently reported method for *P* value calculation, whereas confidence intervals were mainly derived using bootstrap or asymptotic normal approximation (including analytic/asymptotic normal or Wald-type approximations, large-sample approximations from regression or rank-based tests, and the Gehan−Wilcoxon method); however, many studies did not report the specific inferential methods.

Table 1. Baseline and methodological characteristics of included studies for comparative analysis

Characteristic	Study type	
	Prespecified study (n = 33)	**Reanalysis study (n = 25)**
Publish-year, min-max	2018–2025	2015–2025
Randomization	97.0% (32 33)	96.0% (24/25)
Blinding	48.5% (16/33)	64.0% (16/25)
Unmatched	100% (33/33)	92.0% (23/25)
Sample size, median (IQR)	507 (337–632)	3883 (337–11052)
Primary, median (IQR)	507 (215–632)	NA
Secondary, median (IQR)	507.5 (360–878)	NA
Number of components, median (IQR)	4.0 (3.0–6.0)	3.0 (3.0–4.0)
Primary	4.0 (3.0–5.0)	NA
Secondary	5.5 (4.0–7.0)	NA
Patient-reported outcomes	42.4% (14/33)	20.0% (5/25)
Pathophysiological measures	39.4% (13/33)	36.0% (9/25)
Method for P value		
Finkelstein–Schoenfeld test	30.3% (10/33)	20.0% (5/25)
Bootstrap	3.0% (1/33)	4.0% (1/25)
Asymptotic normal inference	27.3% (9/33)	4.0% (1/25)
NR	33.3% (11/33)	44.0% (11/25)
NA	6.1% (2/33)	28.0% (7/25)
Method for CI		
Bootstrap	15.2% (5/33)	16.0% (4/25)
Asymptotic normal inference	18.2% (6/33)	32.0% (8/25)
NR	63.6% (21/33)	52.0% (13/25)
NA	3.0% (1/33)	0%
Conclusion (positive)	48.5% (16/33)	68.0% (17/25)
Primary	39.1% (9/23)	NA
Secondary	70% (7/10)	NA
Win–loss counts by level reported	69.7% (23/33)	84.0% (21/25)
Contribution of the C_1, median (IQR)	13.3% (4.9–24.0) (n = 21)	35.3% (19.4–45.9) (n = 20)
Contribution of the C_L, median (IQR)	37.3% (14.1–57.4) (n = 21)	30.5% (18.9–42.3) (n = 20)

CI, confidence interval; NA, not applicable (no distinction between primary and secondary end points; no P values; no confidence intervals [CIs]); NR, not reported; C_1, highest-priority component; C_L, lowest-priority component.

Table 2 shows more details on prespecified end point. In almost all studies, death was placed as the top-priority component in the hierarchical composite end point, and in many cases, it was defined as all-cause mortality. Only one study in the field of orthopedics used radiographic healing as the top-priority component, as adverse events such as death were extremely rare in that patient population [12]. Furthermore, PROs and pathophysiological components were also frequently included in these hierarchical composite end points. The WR was used to summarize secondary hierarchical end points consisting of 2–7 components. In nearly all cases, these hierarchical composite secondary end points encompassed the original primary end point. In most studies, intercomponent correlation was not explicitly considered in sample size planning at the trial design stage; only four studies incorporated assumptions about correlations between components.

Furthermore, we observed that the application of the WR method is not limited to extending adverse clinical events to include PROs or pathophysiological measures. In some cases, trials originally designed with pathophysiological measures as primary end points incorporated clinical adverse events, such as death, when using the WR for analysis. For the highest-priority component—death—both all-cause mortality and cause-specific mortality (eg, cardiovascular death) were used, whereas two studies conducting WR analyses excluded death from their hierarchical end points [12,13].

As shown in Table S1, from the perspective of end point definition when analyzing the same study population with different analytical approaches, 25 reanalyses (excluding

Table 2. Trial characteristics and end point definitions for prespecified end point analyses

Analysis ID	Author	Year	Trial	Trial identifiers	Sample size	End point
WR-based prespecified primary end point analysis						
PE-1	Mathew S. M.	2018	ATTR-ACT	NCT01994889	441	(1) Mortality (2) Hospitalization
PE-2	Kosiborod, M. N.	2021	DARE-19 [33]	NCT04350593	1250	(1) Death from any cause through day 30 (2) New/worsened organ dysfunction (3) Clinical status at day 30 for patients still hospitalized and without any worsening organ dysfunction (4) Hospital discharge before day 30 and alive at day 30
PE-3	Lopes, R. D.	2021	ACTION [34]	NCT04394377	615	(1) Time to death (2) Duration of hospitalization (3) Duration of supplemental oxygen use through 30 d.
PE-4	Bohula, E. A.	2022	COVID-PACT [35]	NCT04409834	390	(1) Death attributable to venous or arterial thrombosis (2) Pulmonary embolism (3) Clinically evident DVT (4) Type 1 myocardial infarction (5) Ischemic stroke (6) Systemic embolic or acute limb ischemia (7) Clinically silent DVT
PE-5	Fonseca, H. A. R.	2022	VIP-ACS [36]	NCT04001504	1801	(1) All-cause death (2) Myocardial infarction (3) Stroke (4) Hospitalization for unstable angina (5) Hospitalization for heart failure (6) Urgent coronary revascularization (7) Hospitalization for respiratory infections
PE-6	Hepprich, M.	2022	/[37]	NCT04510493	116	(1) Longer survival time (2) Longer ventilation-free time (3) Longer ICU-free time (4) Shorter hospitalization time within 29 days
PE-7	Shah, S. J.	2022	REDUCE LAP-HF II [38]	NCT03088033	626	(1) Cardiovascular death or nonfatal ischemic stroke (2) Recurrent HF events (3) Change in KCCQ
PE-8	Voors, A. A.	2022	EMPULSE [39]	NCT04157751	530	(1) Time to all-cause death (2) Number of HF events (3) Time to first HF event (4) A $\geq$5 point difference in change from baseline in the KCCQ-TSS after 90 d of treatment
PE-9	Freund, Y.	2023	/[40]	NCT04653597	237	(1) Death (2) ICU length of stay (3) Hospital length of stay
PE-10	Mentz, R. J.	2023	FAIR-HF [41]	NCT03037931	3065	(1) Death within 12 months (2) Hospitalizations for heart failure within 12 mo (3) Change in the 6-minute walk distance from baseline to 6 mo
PE-11	Richard, J. C.	2023	VT4COVID [42]	NCT04349618	215	(1) All-cause mortality at day 90 (2) Ventilator-free days at day 60
PE-12	Sorajja, P.	2023	TRILUMINATE [43]	NCT03904147	350	(1) Death from any cause or tricuspid-valve surgery (2) Hospitalization for heart failure (3) An improvement in quality of life as measured with the KCCQ, with an improvement defined as an increase of at least 15 points assessed at the 1-year follow-up

(*Continued*)

Table 2. Continued

Analysis ID	Author	Year	Trial	Trial identifiers	Sample size	End point
PE-13	Gillmore, J. D.	2024	ATTRibute-CM [44]	NCT03860935	632	(1) Death from any cause (2) Cumulative frequency of cardiovascular-related hospitalization (3) The change from baseline in the NT-proBNP level (4) The change from baseline in the 6-minute walk distance
PE-14	James, S.	2024	DAPA-MI [45]	NCT04564742	4017	(1) Death (2) Hospitalization for HF, (3) Nonfatal MI (4) Trialibrillation/flutter event (5) New diagnosis of type 2 diabetes (6) NYHA Functional Classification at the last trial visit (7) Body weight decrease of 5% or more from baseline to the last trial visit
PE-15	Lefaivre, K. A.	2024	COTS [12]	NCT01766648	193	(1) Radiographic healing (2) Clinical fracture-healing.
PE-16	Rohde, L. E.	2024	MESSAGE-HF [46]	NCT04062461	699	(1) Time to cardiovascular death (2) Time to first HF hospitalization (3) Relative change of more than 20% in NT-proBNP level from study inclusion to 180 days of follow-up
PE-17	Stone, G. W.	2024	RELIEVE-HF [47]	NCT03499236	508	(1) All-cause death (2) Cardiac transplantation or LVAD implantation (3) HFHs (4) Outpatient worsening HF events (5) Change in KCCQ from baseline by at least five points
PE-18	Tavares, C. A. M.	2024	DEFENDER [48]	NCT05558098	507	(1) Hospital mortality (2) Initiation of kidney replacement therapy (3) ICU length of stay through 28 d
PE-19	Estep, J. D.	2025	ALIVE [49]	NCT02931240	126	(1) Cardiovascular mortality (2) Hospital readmission for new or worsening HF (3) Improvement in the MLHFQ (4) Improvement in 6MWT (5) Improvement in NYHA functional classification by $\geq$ 1 grade.
PE-20	Hahn, R. T.	2025	TRISCEND II [50]	NCT04482062	400	(1) Any cause (2) Durable implantation of a right ventricular assist device or heart transplantation (3) Tricuspid-valve surgery or percutaneous tricuspid intervention after any index intervention (4) Annualized rate of hospitalization for heart failure (5) An improvement of at least 10 points in the score on the KCCQ-OS (6) An improvement of at least one NYHA functional class (7) An increase in the 6-min walk distance of at least 30 m.
PE-21	Jaber, W. A.	2025	PEERLESS [51]	NCT05111613	550	(1) All-cause mortality (2) ICH (3) Major bleeding per International Society for Thrombosis and Hemostasis definition (4) Clinical deterioration and/or escalation to bailout therapy (5) Postprocedural ICU admission and length of stay.

(*Continued*)

Table 2. Continued

Analysis ID	Author	Year	Trial	Trial identifiers	Sample size	End point
PE-22	Mieghem, *N.* M. V.	2025	TAVR UNLOAD [52]	NCT02661451	178	(1) All-cause death (2) Disabling stroke (3) Hospitalizations related to HF symptomatic aortic valve disease or nondisabling stroke (4) The change in KCCQ at 1 y.
PE-23	Konstam, M. A.	2024	AT HOME-HF [53]	NCT04593823	51	(1) CV mortality (2) HF hospitalizations (3) Urgent ED/clinic visits for worsening HF over 30 d) (4) Change from baseline in NT-proBNP at day 7
WR-based prespecified secondary end point analysis						
SE-1	Butler, J.	2022	DIAMOND [54]	NCT03888066	878	(1) Cardiovascular death (2) Cardiovascular hospitalization (3) Severe hyperkalemia event (4) Moderate hyperkalemia event (5) Mild hyperkalemia event
SE-2	Hill, K. D.	2022	STRESS Network [11]	NCT03229538	1263	(1) Operative death (2) Heart transplantation before discharge (3) Severe kidney failure (4) Postoperative mechanical circulatory support (5) Surgical delay, reoperation, or unplanned catheterization (6) Cardiac arrest, organ failure, or mechanical ventilation support (7) Postoperative hospitalization period
SE-3	Kandzari, D. E.	2023	SPYRAL HTN-ON MED [13]	NCT02439775	337	(1) Change in 24-h ambulatory systolic blood pressure (SBP) from baseline to 6 mo, with a threshold of $\geq$5 mm Hg (2) Change in medication burden at 6 months from baseline
SE-4	Kosiborod, M. *N.*	2023	STEP-HFpEF [55]	NCT04788511	529	(1) Death (any cause) (2) No. of heart failure events (3) Time to heart failure events (4) Change in KCCQ-CSS $\geq$15 (5) Change in KCCQ-CSS $\geq$10 (6) Change in KCCQ-CSS $\geq$5 (7) Change in 6-min walk distance
SE-5	Mentz, R. J.	2023	PARAGLIDE-HF [56]	NCT03988634	466	(1) Cardiovascular death (2) HF hospitalizations (3) Urgent HF visits (4) Change in NT-proBNP.
SE-6	MS Maurer	2023	APOLLO-B [57]	NCT03997383	360	(1) Death from any cause (2) Cardiovascular events (3) Change from baseline in the 6-min walk test distance
SE-7	Kosiborod, M. *N.*	2024	STEP-HFpEF DM [58]	NCT04916470	616	(1) Death from any cause (2) Number and timing of heart failure events (3) Change in KCCQ-CSS from baseline to week 52
SE-8	Donal, E.	2025	Tri.Fr [59]	NCT04646811	300	(1) All-cause death (2) Tricuspid valve surgery (3) Time to heart failure hospitalization (4) Improvement of $\geq$15 points in KCCQ score

(*Continued*)

Table 2. Continued

Analysis ID	Author	Year	Trial	Trial identifiers	Sample size	End point
SE-9	Gao, C.	2025	REC-CAGEFREE II [60]	NCT04971356	1948	(1) All-cause death (2) Stroke (3) Myocardial infarction, (4) BARC type 3 bleeding (5) Revascularization (6) BARC type 2 bleeding
SE-10	Shefner, J. M.	2025	COURAGE-ALS [61]	NCT04944784	486	(1) ALSFRS-R total score (2) Time to respiratory insufficiency (3) Survival up to week 24 (4) Change from baseline to week 24 in percentage predicted FVC (5) ALSAQ-40 total score (6) Handgrip strength

PE, primary end point; SE, secondary end point; 6MWD, 6-min walk distance; 6MWT, 6-min walk test; ALSAQ, Amyotrophic Lateral Sclerosis Assessment Questionnaire; CV, cardiovascular; DVT, deep venous thrombosis; ED, emergency department; FVC, forced vital capacity; HF, heart failure; ICH, intracranial hemorrhage; ICU, intensive care unit; KCCQ, Kansas City Cardiomyopathy Questionnaire; KCCQ-CSS, KCCQ Clinical Summary Score; KCCQ-OS, KCCQ Overall Summary; KCCQ-TSS, KCCQ Total Symptom Score; LVAD, left ventricular assist device; MLHFQ, Minnesota Living with Heart Failure Questionnaire; NT-proBNP, N-terminal pro–B-type natriuretic peptide; NYHA, New York Heart Association functional class.

subgroup and pooled analyses that lacked comparable reference points) applied the WR method to reexamine trial outcomes. In all cases, the composite end points selected for reanalysis encompassed the original primary end point. Among these reanalyses, eight studies retained the original components and merely introduced a hierarchical structure, therefore, they can be considered to analyzing equivalent end points. In contrast, the remaining studies made modifications to the end point composition.

3.3. Component directional contribution

Due to the lack of component-level results in one study, a total of 57 studies were included for component-level directional contribution assessment. The component-level contribution of prespecified is illustrated in Figure 4, blue indicates directional concordance, gray indicates neutrality, and red indicates discordance. In prespecified primary end point analysis, only seven studies demonstrated complete

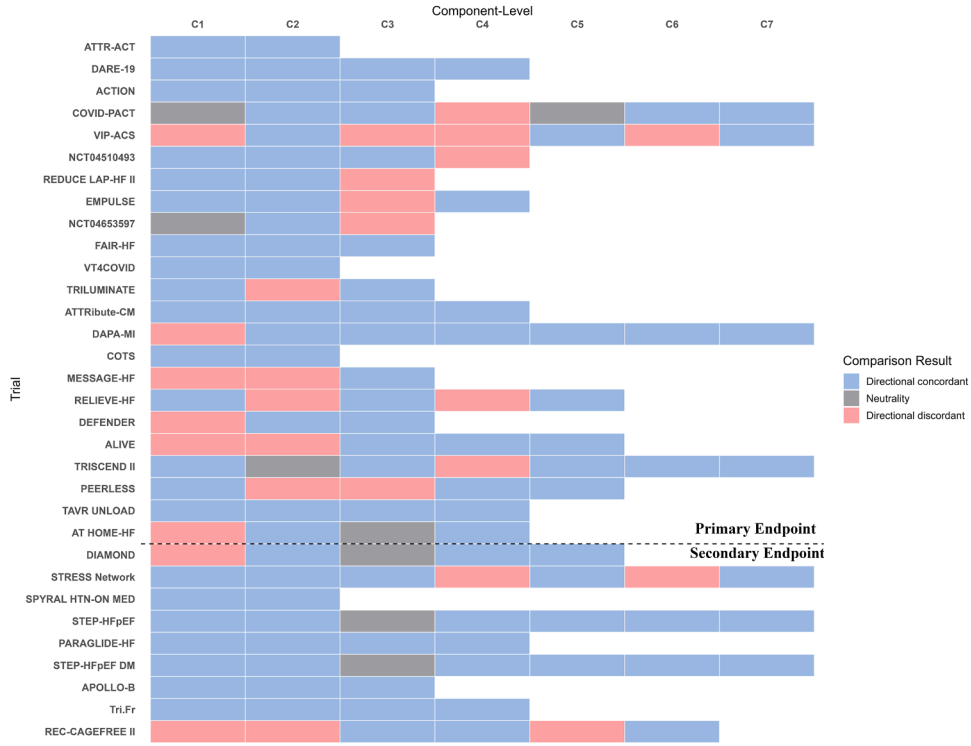


Figure 4. Component directional contribution with composite outcome in prespecified analysis.

agreement between all components and the end point. At the component level, from C1 to C7 (ranked by decreasing clinical priority), the counts of studies with discordance or neutrality were as follows: C1, 8/23; C2, 6/23; C3, 6/20; C4, 5/13; C5, 1/7; C6, 1/4; and C7, 0/4. The denominator at each position represents the total number of components mapped to that position, summed across studies for each level separately. Overall, 27 of 94 component-level comparisons (28.7%) showed discordance or neutrality. In the highest-priority component (C1), the rate of discordance or neutrality was 8/23 (34.8%), while in the lowest-priority component it was 3/23 (13.0%); however, the difference was not statistically significant (exact McNemar's test, $P = .180$). In secondary end point, only two out of ten studies showed any directional discordance among the components within their highest-priority composite. Among a total of 45 components, only six exhibited a direction opposite to that of the overall conclusion (nine when neutrality was included). The component directional concordance contribution for secondary end points was higher than that for primary end points; however, in the logistic generalized linear mixed-effects model with study as a random effect, the difference was not statistically significant ($P = .277$). In prespecified end point analyses, discordance or neutrality occurred in 36 of 139 components (25.9%). We further stratified trials based on whether they reported a positive overall conclusion. Among trials with positive conclusions, 20 of 80 components (25.0%) showed a direction opposite to the overall WR result, compared with 16 of 59 components (27.1%) in trials with negative conclusions ($P = .779$).

As shown in Figure 5, in WR-reported reanalyses, opposite directions were observed in three studies (3/25) for both the highest- and lowest-priority components. In total, 16 out of 94 components (17.0%) demonstrated a directional discordance or neutralit relative to the composite result. When compared with the results from prespecified primary or secondary end point analyses, this proportion did not differ significantly ($P = .114$).

The contribution of the highest- and lowest-priority components was restricted to studies in which level-specific win−loss data were available. Among the included studies, 44 provided component-level reporting; however, three presented only graphical displays without numerical data. Therefore, our analysis focused on 41 studies. The median contribution of the first-priority component was 23.8% (interquartile range, 10.8%−42.9%), whereas that of the last-priority component was 34.0% (18.4%−48.6%).

3.4. Conclusion-level consistency

In the prespecified secondary analyses and reanalyses we examined, 10 studies applied the WR to analyze secondary end points against prespecified primary end points, among which four yielded conclusions that were categorized as disagree. Among the 25 reanalysis studies, 4 (16%) were classified as possibly disagree, and 5 (20%)

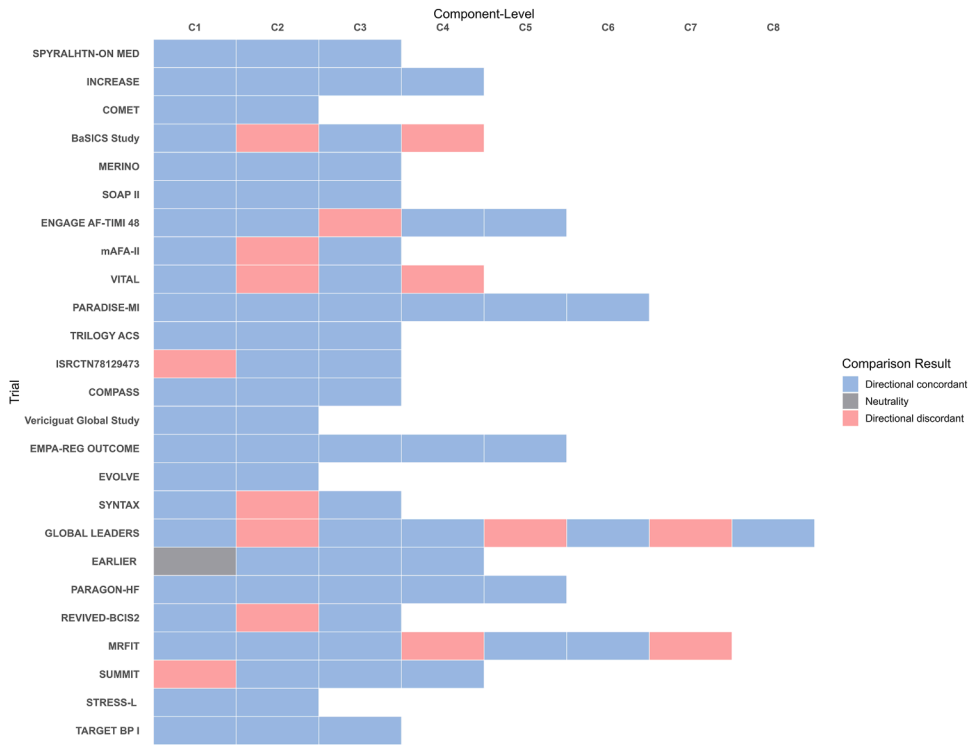


Figure 5. Component directional contribution with composite outcome in reanalysis.

studies were deemed disagree. Notably, in two of these five studies, the original trials concluded noninferiority, whereas the analysis summarized by the WR indicated superiority. In the remaining three, the WR analysis transformed originally nonsignificant findings into statistically significant results. Currently, eight reanalyses used identical end point components and differed only by the addition of a hierarchical structure, and yielded consistent results. In contrast, most of the disagreements arose in studies where the end point components had been changed, reflecting disagreements in conclusions across different analytic strategies despite analyzing the same trial, in the same population, and using clinically relevant outcome components (Fig 6).

4. Discussion

The application of the WR has shown a consistent upward trend. Most studies applying the WR have focused on diseases of the circulatory system, particularly within the cardiovascular domain. Furthermore, emerging evidence suggests that its use is extending into various therapeutic areas, with several studies applying the WR beyond cardiovascular settings and additional publications exploring its methodological applicability in other disease domains [14−18].

In practice, the generalized pairwise comparison framework underlying the WR allows hierarchical outcome structures and flexible composite end point definitions to be incorporated in a prespecified manner, helping align statistical analysis with clinically meaningful priorities.

In this study, we assessed the component directional contribution between individual outcome components and the overall WR result, and we adopted the concept of regulatory agreement to evaluate the extent to which conclusions derived from WR analyses aligned with those of the original studies [10]. It is important to note that our definition does not fully align with the traditional concept of regulatory agreement; it focuses on whether the application of the WR, including the introduction of hierarchical composite end points, led to conclusions that agree with the original studies within the same study population. We adopted this approach because we recognize that if these clinically meaningful end points had been designated as prespecified primary outcomes and yielded positive results, they could possibly influence decisions made by clinicians, patients, and regulatory authorities. A central motivation of this part of the study was to consider the implications of using WR as a prespecified primary analysis strategy. In our review, when the underlying clinical components remained unchanged and only hierarchical prioritization was introduced, the overall trial conclusions were largely consistent with those from the original analyses in most cases.

In our results, the application of the WR frequently revealed directional discordance across the components of composite end points. As shown in Figures 4 and 5, the most striking observation was that agreement tended to occur primarily in components with relatively lower clinical priority, although this difference was not statistically significant. In this comparison, we recognize that we are evaluating conditional rather than marginal effects. Although these are often directionally consistent in practice, the original component-specific results should also be interpreted with care. We also emphasize that we do not expect every hierarchical level to align with the overall result; instead, we focus on whether directional discordance tends to cluster at specific levels or in study contexts.

We also examined whether WR-based conclusions align with those of the original analyses (Fig 6), recognizing that interpretative differences may emerge even when analyzing the same data. As with traditional composite end points [19], the WR also poses certain interpretative challenges in clinical application [20]. Compared with conventional composites, WR addresses the issue of equal weighting

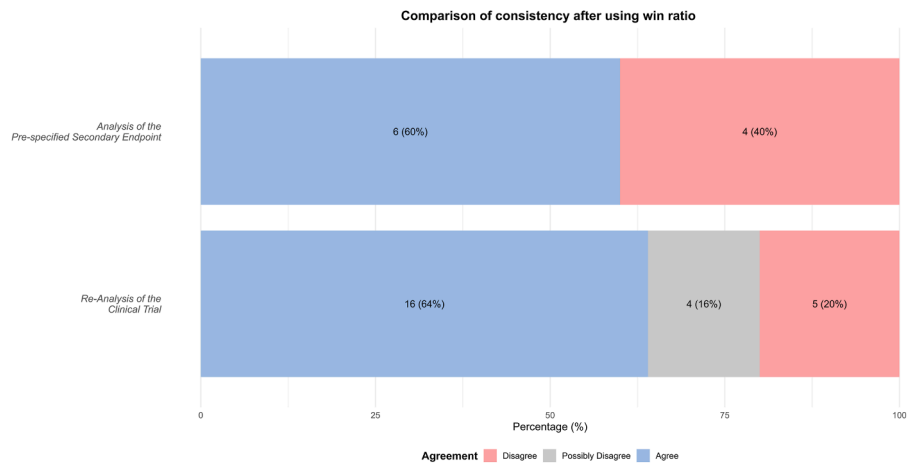


Figure 6. Consistency of study conclusions after applying the win ratio method included in prespecified secondary end point analysis and reanalysis.

among components by incorporating clinical prioritization, which aligns better with clinical decision-making logic [1]. Due to the prioritization structure of the WR, both PROs and pathophysiological measures can be incorporated into the primary end point [21−23]. This feature is essential for advancing patient-centered research and upholding the patient-centered philosophy [24]. However, when measures beyond the original primary end point yield statistical significance, it is critical to consider whether the conclusions remain valid within the original decision-making context, as this affects both interpretability and clinical applicability.

Within the generalized pairwise comparisons framework, the WR is only one of several possible effect summaries. Methodological work has highlighted that the WR can be sensitive to ties and to the conditional nature of hierarchical comparisons, and that its scale may yield misleading impressions of effect size in many realistic settings, particularly when higher-priority events are rare or when a large proportion of pairs are undecided [25]. Consequently, the WR is not universally recommended as the default summary measure. Alternative summaries such as net treatment benefit and its transformations (eg, win odds) or the probabilistic index have been advocated as more interpretable or robust in many applications [8].

Therefore, when applying the WR, researchers should carefully consider the appropriate use of the analytical framework and any clinical implications resulting from modifications to end point hierarchy and definitions, particularly the relative priority and interpretation of different components used to derive the overall clinical conclusion. The relative contribution of different components to the overall result also requires careful interpretation. Some visualization tools have been proposed to support this process [26,27]. Nevertheless, how to design hierarchical composite end points rationally and align them precisely with the clinical question remains a core challenge in WR application. Inappropriate use may lead to drawbacks, which is one reason why the WR has sometimes been controversial [28−30].

These findings highlight the need to be cautious about potential shifts in clinical meaning when applying the WR. It is also worth emphasizing that if the primary aim is to address limitations of traditional composite end points, such as treating all components as equally important or restricting component types, it is not necessarily required to add new outcome components. For example, the traditional major adverse cardiovascular and cerebrovascular events (MACCE) end point can be transformed into a hierarchical major adverse cardiovascular and cerebrovascular events without changing the original components, but by clearly ranking them in order of clinical importance, thereby, striking a balance between clinical relevance and methodological rigor.

5. Limitations

This study was not prospectively registered, which may limit transparency. However, we sought to reduce potential reporting bias by reporting all relevant findings identified during the review process. We systematically reviewed the application of the WR in clinical trials. To ensure the empirical nature of the included studies, we excluded those that used clinical trial data solely for methodological demonstration or exploratory assessments of feasibility. When studies compared multiple methods but ultimately focused on interpreting clinical trial results, we categorized them as reanalyses rather than methodological studies. We believe this reflects a pragmatic approach to capturing real-world applications of the WR. We recognized that the inclusion of studies analyzing multiple trials within a single analysis may be debatable in terms of whether they should be considered as addressing individual trials or broader methodological application. However, we observed that the conclusions of these studies primarily emphasized the application of the WR within specific clinical domains, rather than reporting results from individual trials. Therefore, we considered their exclusion to be reasonable. In addition, some included studies were research letters, which may not have undergone full peer review. Nevertheless, since they clearly applied the WR and were based on actual clinical trial data, their inclusion was considered reasonable. We did not search clinical trial registries (eg, ClinicalTrials.gov) as part of our systematic search strategy, because our primary objective was to assess the application of the WR method in published analyses of clinical trial results, rather than to identify protocols or planned studies that had not yet reported outcomes. Therefore, we focused our search on peer-reviewed literature databases to ensure the relevance and completeness of methodological information.

In our search, we did not include derivative metrics of the WR, such as win odds, as our focus was specifically on the application of the WR itself [31,32]. These related metrics warrant particular attention in future research. In the contribution determination process, when absolute win−loss counts were unavailable, we used the reported percentages for directional classification, which may have introduced minor classification uncertainty due to rounding. We did not infer the methods used to calculate confidence intervals or P values when not explicitly reported. Only methods clearly described or supported by cited statistical references in the original publication were recorded. Recommendations from methodological WR literature were not used to impute information. This approach may have led to a conservative estimate of unreported inferential methods; however, it underscores the importance of clearly specifying statistical procedures in hierarchical end point analyses.

6. Conclusion

The application of the WR in clinical research has shown a clear and steady upward trend. Its use is also expanding beyond traditional cardiovascular trials into other disease areas. In practice, hierarchical composite end points constructed using the WR may differ from conventional or conventional composite end points in terms of component selection. Discordance among components is not uncommon, and in WR-reported reanalyses, changes in the selected components may further influence the clinical interpretation of trial results. Therefore, careful attention must be paid to the appropriateness of the WR, the definition of end point components and the interpretation of results to ensure the clinical applicability and clarity of WR-reported analyses.

CRediT authorship contribution statement

Qize Li: Writing − review & editing, Writing − original draft, Visualization, Methodology, Conceptualization. **Yingxuan Zhu:** Writing − review & editing, Writing − original draft, Visualization, Methodology, Conceptualization. **Yanyan Zhao:** Writing − review & editing, Project administration, Methodology. **Chilie Danzeng:** Writing − review & editing, Visualization, Formal analysis. **Duoer Wang:** Writing − review & editing, Methodology. **Ru Jiang:** Writing − review & editing, Methodology. **FeiLong Chen:** Writing − review & editing, Methodology. **Xiaoying Lin:** Writing − review & editing, Validation. **Lingqian Hu:** Writing − review & editing, Validation. **Yang Wang:** Writing − review & editing, Project administration, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary data

Supplementary data related to this article can be found at https://doi,org/10.1016/j.jclinepi.2026.112315.

Data availability

Data will be made available on request.

References

[1] Pocock SJ, Ariti CA, Collier TJ, Wang D. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. Eur Heart J 2012;33(2):176−82. https://doi.org/10.1093/eurheartj/ehr352.

[2] Gregson J, Stone GW, Bhatt DL, Packer M, Anker SD, Zeller C, et al. Recurrent events in cardiovascular trials: JACC State-of-the-Art review. J Am Coll Cardiol 2023;82(14):1445−63. https://doi.org/10.1016/j.jacc.2023.07.024.

[3] Mao L, Kim K, Li Y. On recurrent-event win ratio. Stat Methods Med Res 2022;31(6):1120−34. https://doi.org/10.1177/09622802221084134.

[4] Redfors B, Gregson J, Crowley A, McAndrew T, Ben-Yehuda O, Stone GW, et al. The win ratio approach for composite endpoints: practical guidance based on previous experience. Eur Heart J 2020; 41(46):4391−9. https://doi.org/10.1093/eurheartj/ehaa665.

[5] Wang D, Pocock S. A win ratio approach to comparing continuous non-normal outcomes in clinical trials. Pharm Stat 2016;15(3): 238−45. https://doi.org/10.1002/pst.1743.

[6] Wang B, Zhou D, Zhang J, Kim Y, Chen LW, Dunnmon P, et al. Statistical power considerations in the use of win ratio in cardiovascular outcome trials. Contemp Clin Trials 2023;124:107040. https://doi.org/10.1016/j.cct.2022.107040.

[7] Finkelstein DM, Schoenfeld DA. Combining mortality and longitudinal measures in clinical trials. Stat Med 1999;18(11):1341−54.

[8] Buyse M, Verbeeck J, De Backer M, Deltuvaite-Thomas V, Saad ED, Molenberghs G. Handbook of generalized pairwise comparisons: methods for patient-centric analysis. Boca Raton (FL: CRC Press; 2025.

[9] Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. https://doi.org/10.1136/bmj.n71.

[10] Franklin JM, Pawar A, Martin D, Glynn RJ, Levenson M, Temple R, et al. Nonrandomized real-world evidence to support regulatory decision making: process for a randomized trial replication project. Clin Pharmacol Ther 2020;107:817−26. https://doi.org/10.1002/cpt.1633.

[11] Hill KD, Kannankeril PJ, Jacobs JP, Baldwin HS, Jacobs ML, O'Brien SM, et al. Methylprednisolone for heart surgery in infants-a randomized, controlled trial. N Engl J Med 2022;387(23): 2138−49. https://doi.org/10.1056/NEJMoa2212667.

[12] Lefaivre KA, Slobogean G, O'Hara NN, Canadian Orthopaedic Trauma Society (COTS) Investigators. Far cortical locking versus standard constructs for locked plate fixation in the treatment of acute, displaced fractures of the distal femur. J Bone Joint Surg Am 2024; 106(19):1739−49. https://doi.org/10.2106/JBJS.23.01390.

[13] Kandzari DE, Townsend RR, Kario K, Mahfoud F, Weber MA, Schmieder RE, et al. Safety and efficacy of renal denervation in patients taking antihypertensive medications. J Am Coll Cardiol 2023; 82(19):1809−23. https://doi.org/10.1016/j.jacc.2023.08.045.

[14] Ferreira JP, Jhund PS, Duarte K, Claggett BL, Solomon SD, Pocock S, et al. Use of the win ratio in cardiovascular trials. JACC Heart Fail 2020; 8(6):441−50. https://doi.org/10.1016/j.jchf.2020.02.010.

[15] Ferreira JP, Zannad F. The win ratio for composite outcomes in nephrology clinical trials. J Am Soc Nephrol 2024. https://doi.org/10.1681/asn.0000000597.

[16] Monzo L, Levy B, Duarte K, Baudry G, Combes A, Ouattara A, et al. Use of the win ratio analysis in critical care trials. Am J Respir Crit Care Med 2024;209(7):798−804. https://doi.org/10.1164/rccm.202309-1644CP.

[17] Gregson J, Donners SJA, Dippel D, Lingsma H, Dea Hazewinkel A, Brown MM, et al. Use of the win ratio for analysis of stroke trials: description, illustration, and planned use in the second European Carotid Surgery Trial (ECST-2). Stroke 2024. https://doi.org/10.1161/STROKEAHA.124.048689.

[18] Zampieri FG, Serpa-Neto A, Wald R, Bellomo R, Bagshaw SM. Hierarchical endpoints in critical care: a post-hoc exploratory analysis of the standard versus accelerated initiation of renal-replacement therapy in acute kidney injury and the intensity of continuous renal-replacement therapy in critically ill patients trials. J Crit Care 2024;82:154767. https://doi.org/10.1016/j.jcrc.2024.154767.

[19] Montori VM, Permanyer-Miralda G, Ferreira-González I, Busse JW, Pacheco-Huergo V, Bryant D, et al. Validity of composite end points in clinical trials. BMJ 2005;330(7491):594−6. https://doi.org/10.1136/bmj.330.7491.594.
[20] Ferreira-González I, Busse JW, Heels-Ansdell D, Montori VM, Akl EA, Bryant DM, et al. Problems with use of composite end points in cardiovascular trials: systematic review of randomised controlled trials. BMJ 2007;334(7597):786. https://doi.org/10.1136/bmj.39136.682083.
[21] Metz C, McCracken P, Hanmer J. Common patient-reported outcomes within the Food and Drug Administration voice of the patient reports. Value Health 2022;25(10):1743−51. https://doi.org/10.1016/j.jval.2022.03.024.
[22] Savarese G, Lindenfeld J, Stolfo D, Adams K, Ahmad T, Desai NR, et al. Use of patient-reported outcomes in heart failure: from clinical trials to routine practice. Eur J Heart Fail 2023;25(2):139−51. https://doi.org/10.1002/ejhf.2778.
[23] Reza N, Day SM, Owens AT. Patient-reported outcomes in clinical studies of patients with hypertrophic cardiomyopathy. Prog Cardiovasc Dis 2023;80:60−5. https://doi.org/10.1016/j.pcad.2023.08.007.
[24] Barry MJ, Edgman-Levitan S. Shared decision making–pinnacle of patient-centered care. N Engl J Med 2012;366(9):780−1. https://doi.org/10.1056/NEJMp1109283.
[25] Verbeeck J, De Backer M, Verwerft J, Salvaggio S, Valgimigli M, Vranckx P, et al. Generalized pairwise comparisons to assess treatment effects: JACC review topic of the week. J Am Coll Cardiol 2023;82(13):1360−72. https://doi.org/10.1016/j.jacc.2023.06.047.
[26] Karpefors M, Lindholm D, Gasparyan SB. The maraca plot: a novel visualization of hierarchical composite endpoints. Clin Trials 2023;20(1):84−8. https://doi.org/10.1177/17407745221134949.
[27] Finkelstein DM, Schoenfeld DA. Graphing the win ratio and its components over time. Stat Med 2019;38(1):53−61. https://doi.org/10.1002/sim.7895.
[28] Ajufo E, Nayak A, Mehra MR. Fallacies of using the win ratio in cardiovascular trials: challenges and solutions. JACC Basic Transl Sci 2023;8(6):720−7. https://doi.org/10.1016/j.jacbts.2023.05.004.
[29] Butler J, Stockbridge N, Packer M. Win ratio: a seductive but potentially misleading method for evaluating evidence from clinical trials. Circulation 2024;149(20):1546−8. https://doi.org/10.1161/circulationaha.123.067786.
[30] Davison BA, Cotter G, Koch GG. Hierarchical composite outcomes are here to stay, despite the limitations of the win ratio. Eur J Heart Fail 2025. https://doi.org/10.1002/ejhf.3590.
[31] Dong G, Huang B, Verbeeck J, Cui Y, Song J, Gamalo-Siebers M, et al. 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 2023;22(1):20−33. https://doi.org/10.1002/pst.2251.
[32] Song J, Verbeeck J, Huang B, Hoaglin DC, Gamalo-Siebers M, Seifu Y, et al. The win odds: statistical inference and regression. J Biopharm Stat 2023;33(2):140−50. https://doi.org/10.1080/10543406.2022.2089156.
[33] Kosiborod MN, Esterline R, Furtado RHM, Oscarsson J, Gasparyan SB, Koch GG, et al. Dapagliflozin in patients with cardiometabolic risk factors hospitalised with COVID-19 (DARE-19): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Diabetes Endocrinol 2021;9(9):586−94. https://doi.org/10.1016/s2213-8587(21)00180-7.
[34] Lopes RD, de Barros e Silva PGM, Furtado RHM, Macedo AVS, Bronhara B, Damiani LP, et al. Therapeutic versus prophylactic anticoagulation for patients admitted to hospital with COVID-19 and elevated D-dimer concentration (ACTION): an open-label, multicentre, randomised, controlled trial. The Lancet 2021;397(10291):2253−63. https://doi.org/10.1016/S0140-6736(21)01203-4.
[35] Bohula EA, Berg DD, Lopes MS, Connors JM, Babar I, Barnett CF, et al. Anticoagulation and antiplatelet therapy for prevention of venous and arterial thrombotic events in critically ill patients with COVID-19: COVID-PACT. CIRCULATION 2022;146(18):1344−56. https://doi.org/10.1161/CIRCULATIONAHA.122.061533.
[36] Fonseca HAR, Furtado RHM, Zimerman A, Lemos PA, Franken M, Monfardini F, et al. Influenza vaccination strategy in acute coronary syndromes: the VIP-ACS trial. Eur Heart J 2022;43(41):4378−88. https://doi.org/10.1093/eurheartj/ehac472.
[37] Hepprich M, Mudry JM, Gregoriano C, Jornayvaz FR, Carballo S, Wojtusciszyn A, et al. Canakinumab in patients with COVID-19 and type 2 diabetes − a multicentre, randomised, double-blind, placebo-controlled trial. eClinicalMedicine 2022;53. https://doi.org/10.1016/j.eclinm.2022.101649.
[38] Shah SJ, Borlaug BA, Chung ES, Cutlip DE, Debonnaire P, Fail PS, et al. Atrial shunt device for heart failure with preserved and mildly reduced ejection fraction (REDUCE LAP-HF II): a randomised, multicentre, blinded, sham-controlled trial. Lancet 2022;399(10330):1130−40. https://doi.org/10.1016/s0140-6736(22)00016-2.
[39] Voors AA, Angermann CE, Teerlink JR, Collins SP, Kosiborod M, Biegus J, et al. The SGLT2 inhibitor empagliflozin in patients hospitalized for acute heart failure: a multinational randomized trial. Nat Med 2022;28(3):568−74. https://doi.org/10.1038/s41591-021-01659-1.
[40] Freund Y, Viglino D, Cachanado M, Cassard C, Montassier E, Douay B, et al. Effect of noninvasive airway management of comatose patients with acute poisoning: a randomized clinical trial. JAMA 2023;330(23):2267−74. https://doi.org/10.1001/jama.2023.24391.
[41] Mentz RJ, Garg J, Rockhold FW, Butler J, De Pasquale CG, Ezekowitz JA, et al. Ferric carboxymaltose in heart failure with iron deficiency. N Engl J Med 2023;389(11):975−86. https://doi.org/10.1056/NEJMoa2304968.
[42] Richard JC, Terzi N, Yonis H, Chorfa F, Wallet F, Dupuis C, et al. Ultra-low tidal volume ventilation for COVID-19-related ARDS in France (VT4COVID): a multicentre, open-label, parallel-group, randomised trial. The Lancet Respir Med 2023;11(11):991−1002. https://doi.org/10.1016/S2213-2600(23)00221-7.
[43] Sorajja P, Whisenant B, Hamid N, Naik H, Makkar R, Tadros P, et al. Transcatheter repair for patients with tricuspid regurgitation. N Engl J Med 2023;388(20):1833−42. https://doi.org/10.1056/NEJMoa2300525.
[44] Gillmore JD, Judge DP, Cappelli F, Fontana M, Garcia-Pavia P, Gibbs S, et al. Efficacy and safety of acoramidis in transthyretin amyloid cardiomyopathy. N Engl J Med 2024;390(2):132−42. https://doi.org/10.1056/NEJMoa2305434.
[45] James S, Erlinge D, Storey RF, McGuire DK, de Belder M, Eriksson N, et al. Dapagliflozin in myocardial infarction without diabetes or heart failure. NEJM Evid 2024;3(2):EVIDoa2300286. https://doi.org/10.1056/EVIDoa2300286.
[46] Rohde LE, Rover MM, Hoffmann Filho CR, Rabelo-Silva ER, Silvestre OM, Martins SM, et al. Multifaceted strategy based on automated text messaging after a recent heart failure admission. JAMA Cardiol 2024;9(2):105−13. https://doi.org/10.1001/jamacardio.2023.4501.
[47] Stone GW, Lindenfeld J, Rodés-Cabau J, Anker SD, Zile MR, Kar S, et al. Interatrial shunt treatment for heart failure: the randomized RELIEVE-HF trial. Circulation 2024;150(24):1931−43. https://doi.org/10.1161/circulationaha.124.070870.
[48] Tavares CAM, Azevedo LCP, Rea-Neto Á, Campos NS, Amendola CP, Kozesinski-Nakatani AC, et al. Dapagliflozin for critically ill patients with acute organ dysfunction: the DEFENDER randomized clinical trial. JAMA 2024;332(5):401−11. https://doi.org/10.1001/jama.2024.10510.
[49] Estep JD, Ben-Yehuda O, Wechsler AS, Puri R, Kao AC, Heimes JK, et al. Transcatheter left ventricular reconstruction in heart failure patients with prior anterior myocardial infarction: the prospective ALIVE trial. JACC Heart Fail 2025;13(2):296−308. https://doi.org/10.1016/j.jchf.2024.09.023.
[50] Hahn RT, Makkar R, Thourani VH, Makar M, Sharma RP, Haeffele C, et al. Transcatheter valve replacement in severe tricuspid regurgitation. N Engl J Med 2025;392(2):115−26. https://doi.org/10.1056/NEJMoa2401918.

[51] Jaber WA, Gonsalves CF, Stortecky S, Horr S, Pappas O, Gandhi RT, et al. Large-Bore mechanical thrombectomy versus catheter-directed thrombolysis in the management of intermediate-risk pulmonary embolism: primary results of the PEERLESS randomized controlled trial. Circulation 2025;151(5):260−73. https://doi.org/10.1161/circulationaha.124.072364.

[52] Mieghem NMV, Elmariah S, Spitzer E, et al. Transcatheter aortic valve replacement in patients with systolic heart failure and moderate aortic stenosis TAVR UNLOAD. J Am Coll Cardiol 2025;85(9):878−90. https://doi.org/10.1016/j.jacc.2024.10.070.

[53] Konstam MA, Massaro J, Dhingra R, Walsh M, Ordway L, Pursley MS, et al. Avoiding treatment in hospital with subcutaneous furosemide for worsening heart failure. JACC Heart Fail 2024;12(11):1830−41. https://doi.org/10.1016/j.jchf.2024.07.015.

[54] Butler J, Anker SD, Lund LH, Coats AJS, Filippatos G, Siddiqi TJ, et al. Patiromer for the management of hyperkalemia in heart failure with reduced ejection fraction: the DIAMOND trial. Eur Heart J 2022;43(41):4362−73. https://doi.org/10.1093/eurheartj/ehac401.

[55] Kosiborod MN, Abildstrøm SZ, Borlaug BA, Butler J, Rasmussen S, Davies M, et al. Semaglutide in patients with heart failure with preserved ejection fraction and obesity. N Engl J Med 2023;389(12):1069−84. https://doi.org/10.1056/NEJMoa2306963.

[56] Mentz RJ, Ward JH, Hernandez AF, Lepage S, Morrow DA, Sarwat S, et al. Angiotensin-Neprilysin inhibition in patients with mildly reduced or preserved ejection fraction and worsening heart failure. J Am Coll Cardiol 2023;82(1):1−12. https://doi.org/10.1016/j.jacc.2023.04.019.

[57] Maurer MS, Kale P, Fontana M, Berk JL, Grogan M, Gustafsson F, et al. Patisiran treatment in patients with transthyretin cardiac amyloidosis. N Engl J Med 2023;389(17):1553−65. https://doi.org/10.1056/NEJMoa2300757.

[58] Kosiborod MN, Petrie MC, Borlaug BA, Butler J, Davies MJ, Hovingh GK, et al. Semaglutide in patients with obesity-related heart failure and type 2 diabetes. N Engl J Med 2024;390(15):1394−407. https://doi.org/10.1056/NEJMoa2313917.

[59] Donal E, Dreyfus J, Leurent G, Coisne A, Leroux PY, Ganivet A, et al. Transcatheter edge-to-edge repair for severe isolated tricuspid regurgitation. JAMA 2025;333(2):124−32. https://doi.org/10.1001/jama.2024.21189.

[60] Gao C, Zhu B, Ouyang F, Wen S, Xu Y, Jia W, et al. Stepwise dual antiplatelet therapy de-escalation in patients after drug coated balloon angioplasty (REC-CAGEFREE II): Multicentre, randomised, open label, assessor blind, non-inferiority trial. BMJ 2025;388:e082945. https://doi.org/10.1136/bmj-2024-082945.

[61] Shefner JM, Cudkowicz ME, Genge A, Hardiman O, Al-Chalabi A, Andrews JA, et al. Reldesemtiv in amyotrophic lateral sclerosis: results from the COURAGE-ALS randomized clinical trial. JAMA Neurol 2025. https://doi.org/10.1001/jamaneurol.2025.0241.