Statistical Considerations for Endpoints in Stimulant Use Disorder Clinical Trials

Vlad Dragalin, PhD

Johnson & Johnson

Disclaimer

Views presented are those of the author and may not coincide with company positions.

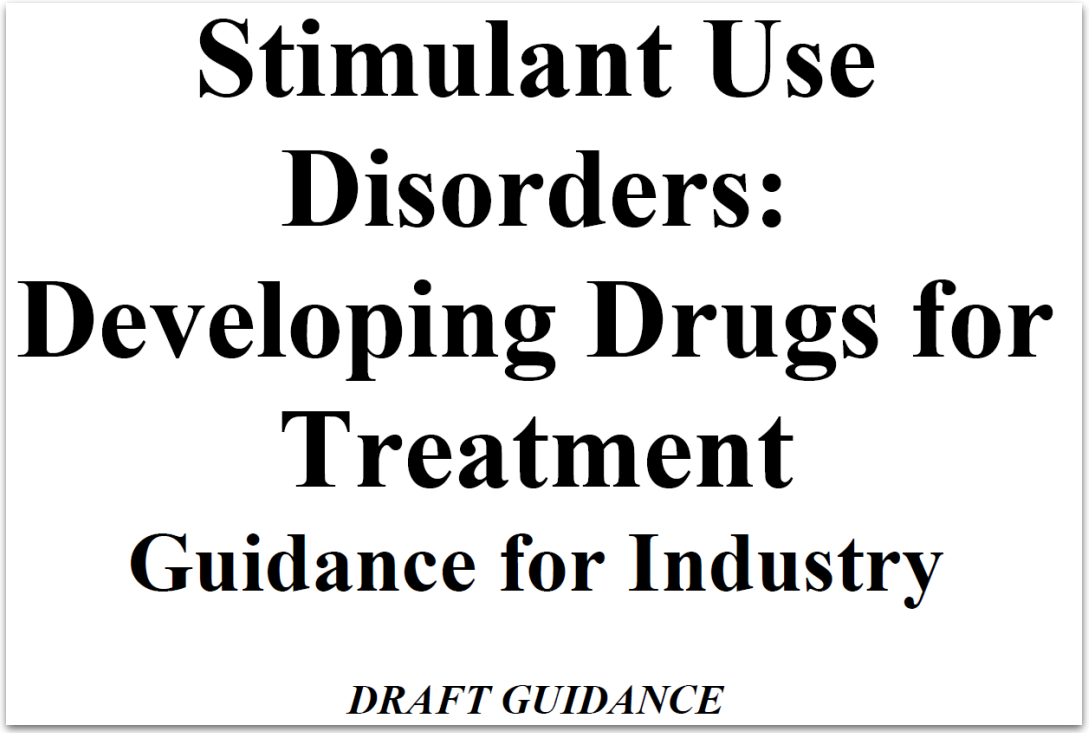
The case study example presented does not reflect any J&J study and was specifically developed for purpose of this presentation.

Disclosure: Shares in Johnson & Johnson stock

Measurements of Drug Use

To Assess Treatment Response

- Historically, results of urine toxicology testing are used to assess response to treatment
- But they are a surrogate measure, because they don't reflect how the subject *feels, functions, or survives*
- The number of subjects achieving complete abstinence based on sustained negative urine toxicology findings is not the only appropriate endpoint
- *Fewer uses per day* or *reduced amount of drug used per occasion of use* are impractical to measure in SUD



Endpoints

Change in Pattern of Stimulant Use

- **Pattern of Stimulant use:** the frequency of stimulant use per period of time (days of use per week or month)
- **Acceptable endpoint:** the proportion of subjects achieving a target pattern of use days per period of time – responder/non-responder
- **Not recommended:** the mean number of days free of use
- *Fewer uses per day* or *reduced amount of drug used per occasion of use* are impractical to measure in SUD

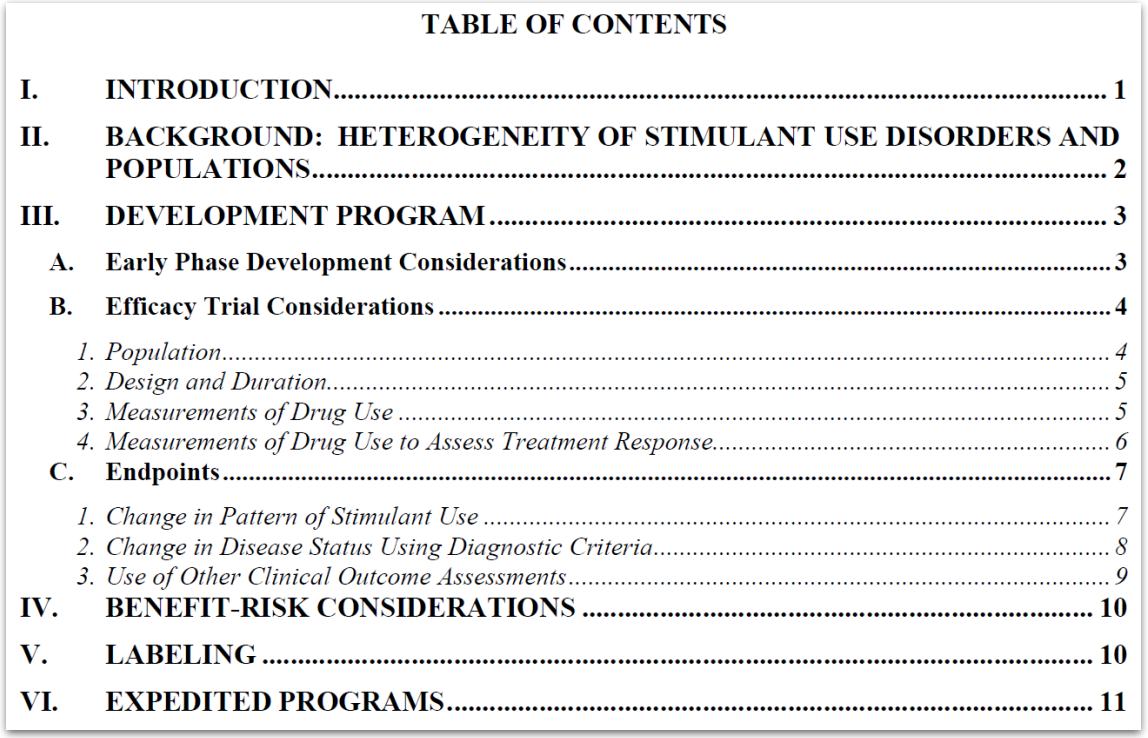


TABLE OF CONTENTS

Endpoints

Change in Disease Status Using Diagnostic DSM-5 Criteria

- **Early remission** is defined as meeting none of the criteria for stimulant use disorder for between 3 and 12 months
- **Sustained remission** is defined as meeting none of the criteria for stimulant use disorder for at least 12 months
- A suitable **primary endpoint** could be the proportion of subjects meeting criteria for early remission from stimulant use disorder at the end of the trial
- **Not recommended:** change in the number of DMC-5 diagnostic criteria endorsed

Stimulant Use Disorder

Diagnostic Criteria

A. A pattern of amphetamine-type substance, cocaine, or other stimulant use leading to clinically significant impairment or distress, as manifested by at least two of the following, occurring within a 12-month period:

1. The stimulant is often taken in larger amounts or over a longer period than was intended.
2. There is a persistent desire or unsuccessful efforts to cut down or control stimulant use.
3. A great deal of time is spent in activities necessary to obtain the stimulant, use the stimulant, or recover from its effects.
4. Craving, or a strong desire or urge to use the stimulant.
5. Recurrent stimulant use resulting in a failure to fulfill major role obligations at work, school, or home.
6. Continued stimulant use despite having persistent or recurrent social or interpersonal problems caused or exacerbated by the effects of the stimulant.
7. Important social, occupational, or recreational activities are given up or reduced because of stimulant use.
8. Recurrent stimulant use in situations in which it is physically hazardous.
9. Stimulant use is continued despite knowledge of having a persistent or recurrent physical or psychological problem that is likely to have been caused or exacerbated by the stimulant.
10. Tolerance, as defined by either of the following:
    a. A need for markedly increased amounts of the stimulant to achieve intoxication or desired effect.
    b. A markedly diminished effect with continued use of the same amount of the stimulant.

    Note: This criterion is not considered to be met for those taking stimulant medications solely under appropriate medical supervision, such as medications for attention-deficit/hyperactivity disorder or narcolepsy.
11. Withdrawal, as manifested by either of the following:
    a. The characteristic withdrawal syndrome for the stimulant (refer to Criteria A and B of the criteria set for stimulant withdrawal, p. 569).
    b. The stimulant (or a closely related substance) is taken to relieve or avoid withdrawal symptoms.

Endpoints

Use of Other Clinical Outcome Assessments

- **PRO**: patient-, observer-, or clinician-reported outcome measure to evaluate a direct effect on how patients feel or function
- **Stimulant craving**: ability of craving modification to predict clinical benefit to consider craving as a potential primary endpoint
- **Various adverse clinical outcomes**: reduced overall or overdose mortality or fewer hospitalizations
- Well-designed, appropriately justified **composite endpoints**

Moderate: Presence of 4–5 symptoms.
F15.20 Amphetamine-type substance
F14.20 Cocaine
F15.20 Other or unspecified stimulant
Moderate, In early remission
F15.21 Amphetamine-type substance
F14.21 Cocaine
F15.21 Other or unspecified stimulant
Moderate, In sustained remission
F15.21 Amphetamine-type substance
F14.21 Cocaine
F15.21 Other or unspecified stimulant
Severe: Presence of 6 or more symptoms.
F15.20 Amphetamine-type substance
F14.20 Cocaine
F15.20 Other or unspecified stimulant
Severe, In early remission
F15.21 Amphetamine-type substance
F14.21 Cocaine
F15.21 Other or unspecified stimulant
Severe, In sustained remission
F15.21 Amphetamine-type substance
F14.21 Cocaine
F15.21 Other or unspecified stimulant

Composite Outcome based on Prioritized Components

Motivation

- FDA recommendations follow traditional approach:
  - Responder/No Responder (not sensitive, loss of information)
  - Primary Endpoints (dual or co-primary), Key secondary endpoints (with multiplicity adjustments)
  - In the absence of prioritized components, for composite time-to-event endpoints, the outcome may not reflect the most important event (e.g. death)
  - Use patients to analyze outcomes rather than outcomes to analyze patients by comparing the experiences of trial participants in different treatment arms by the desirability of the overall patient outcome.

- In other TAs, FDA has been requesting newer methods in recent studies based on component ranking or weighting to address the drawbacks of traditional approach.
  - Ranking allows for the differentiation of the relative clinical importance of component from patient's or physician's perspectives, which will inform decision making for patient-centric drug development.
  - These methods can be extended to composites with a mixture of different types of outcomes (e.g., one time-to-event component and one binary component).
- The goal is a highly interpretable measure of treatment effect that properly weighs and considers all relevant available data into a single reportable estimate of treatment effect.

Methodology

Paradigm for Rank-Based Approach

Raw Data

- Multiple component outcomes per subject
- May be a mixture of components, e.g. continuous, binary, time-to-event
- May have recurrent events for time-to-event outcomes

"Ordinal" Data

- Determine component ranking
- Algorithm performs relative comparisons between subjects based on component hierarchy
- May be numbers of "wins", "losses", and "ties" between subjects
- May be a relative ordinal score used for subject-level comparison

Inference

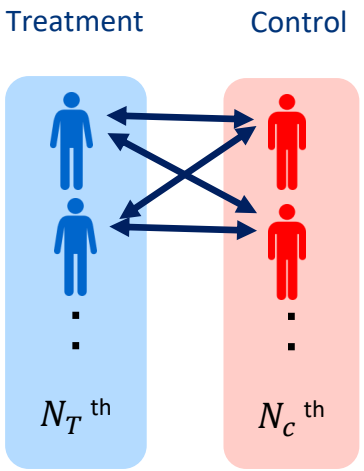
- Number of wins (better outcome in treatment vs control)
- Probability of winning
- Win ratio, Win Odds, Net Benefit
- Evaluate treatment effect based on "win statistics"
- P-value
- Point estimate
- Confidence interval

Methodology

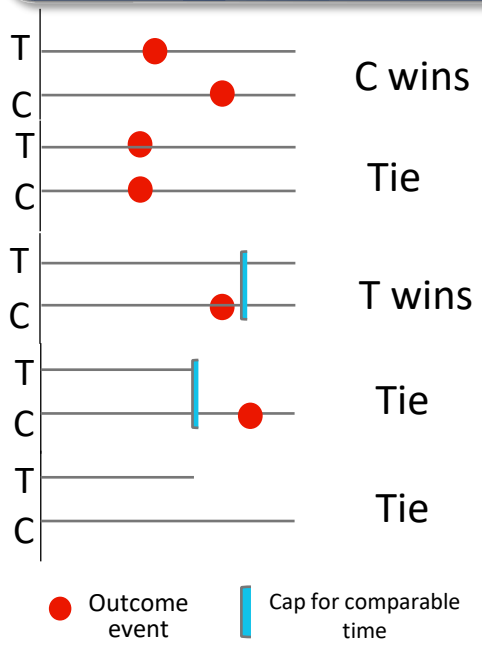
Generalized Pairwise Comparison (GPC) approach

Consider unique pairs of treatment vs control

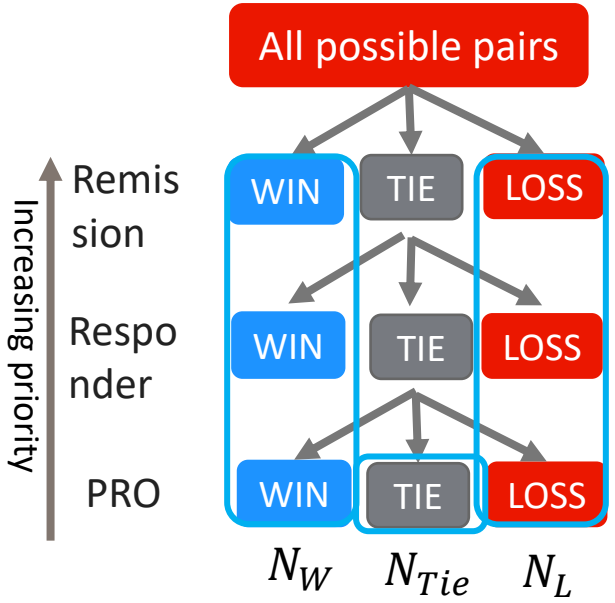
$N_P = N_T N_C$ possible pairs



Assign Win, Loss, or Tie for a single component



Determine Win, Loss, or Tie for multiple components



Win Statistics

- Win Ratio:

$$WR = \frac{N_W}{N_L}$$

- Win Odds:

$$WO = \frac{N_W + 0.5 * N_{Tie}}{N_L + 0.5 * N_{Tie}}$$

- Net Benefit:

$$NB = N_W/N_P - N_L\ /N_P$$

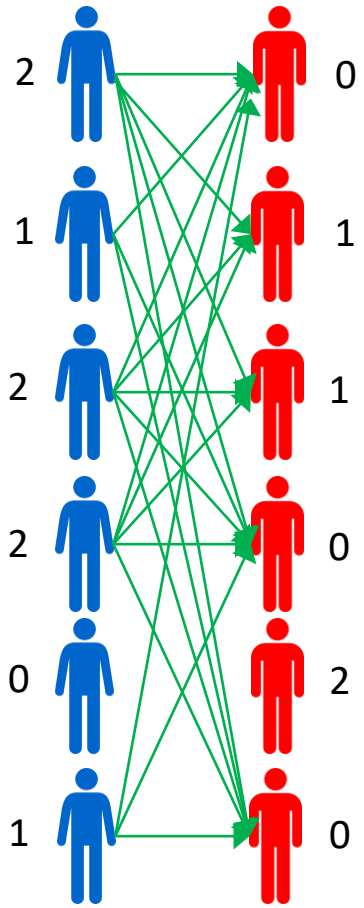
- Both have events-> later occurrence wins
- If only one has the event->no event within common follow-up time wins.
- Both are censored->ties

- Rank components (Remission>> Responder>>PRO)
- Comparisons starting with the most important.
- If ties, move on to the next in the rank.

Example

Generalized Pairwise Comparisons on Remission

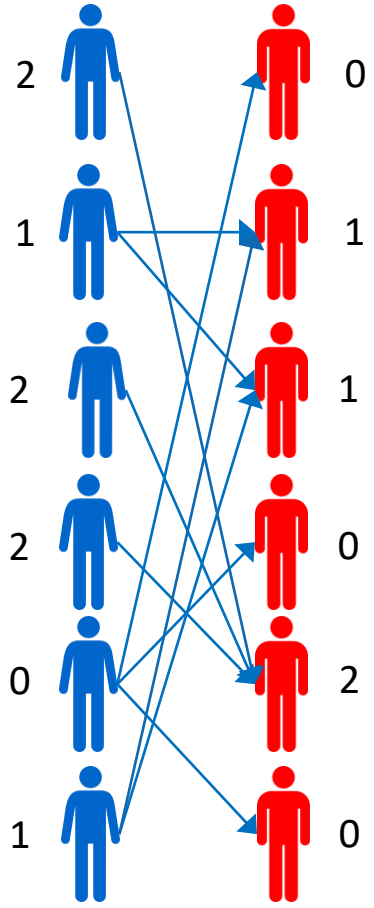
WINS=21



0 – no improvement
1 – early remission
2 – sustained remission

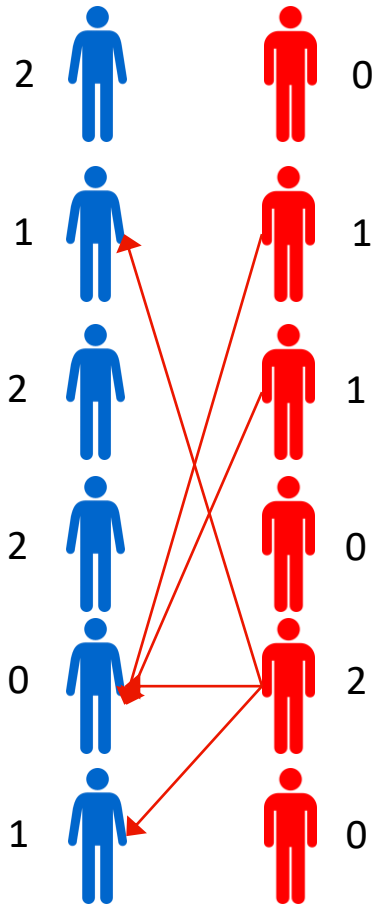
$$WR = \frac{N_W}{N_L} = \frac{21}{5} = 4.2$$

TIES=10



$$NB = N_W/N_P - N_L\ /N_P = 0.44$$

LOSSES=5

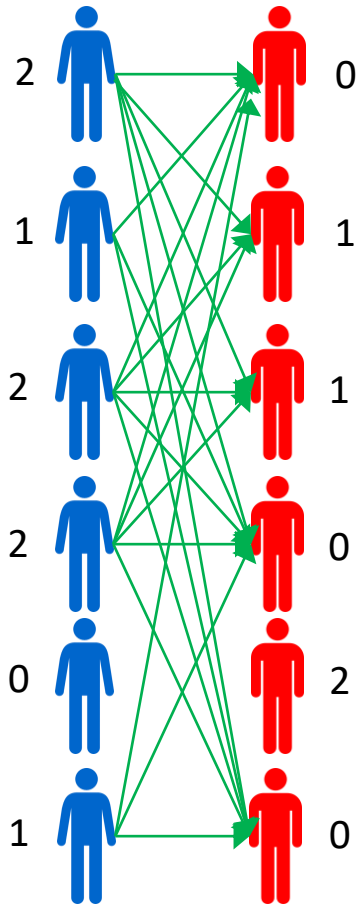


$$WO = \frac{N_W + 0.5 * N_{Tie}}{N_L + 0.5 * N_{Tie}} = \frac{21+5}{5+5} = 2.6$$

Example continued

GPC on Responders (achieving a target pattern of use days per period of time)

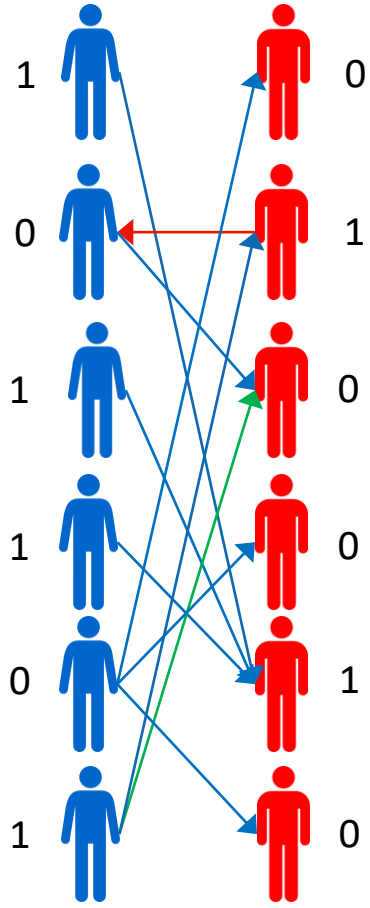
WINS=21



$$WR = \frac{N_W}{N_L} = \frac{21+1}{5+1} = 3.67$$

TIES=10

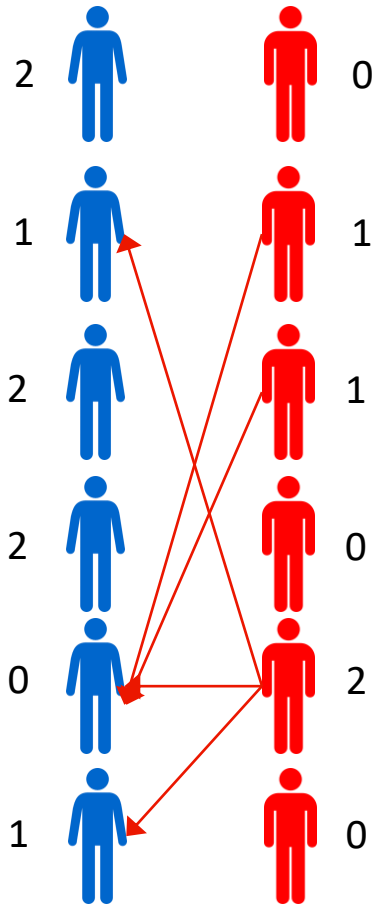
W=1, T=8, L=1



0 – non-responder
1 – responder

$$NB = N_W/N_P - N_L/N_P = 0.44$$

LOSSES=5

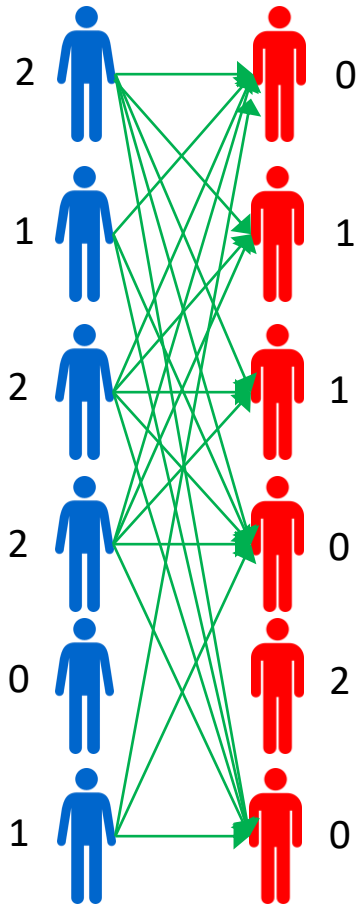


$$WO = \frac{N_W + 0.5 * N_{Tie}}{N_L + 0.5 * N_{Tie}} = \frac{22+4}{6+4} = 2.6$$

Example continued

GPC on Responders (achieving a target pattern of use days per period of time)

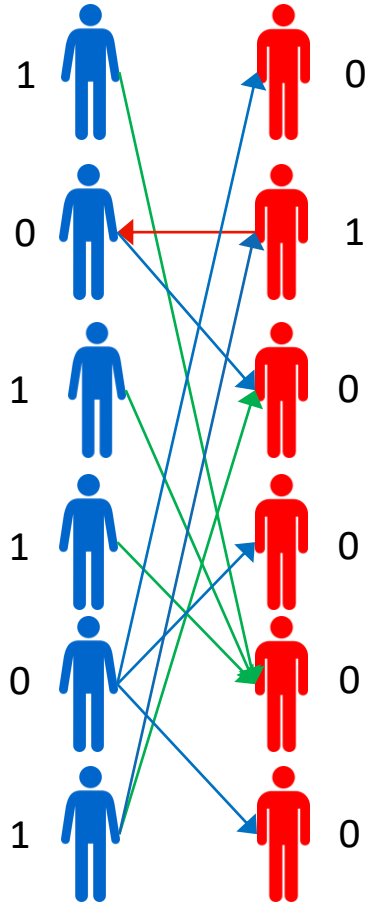
WINS=21



$$WR = \frac{N_W}{N_L} = \frac{21+4}{5+1} = 4.17$$

TIES=10

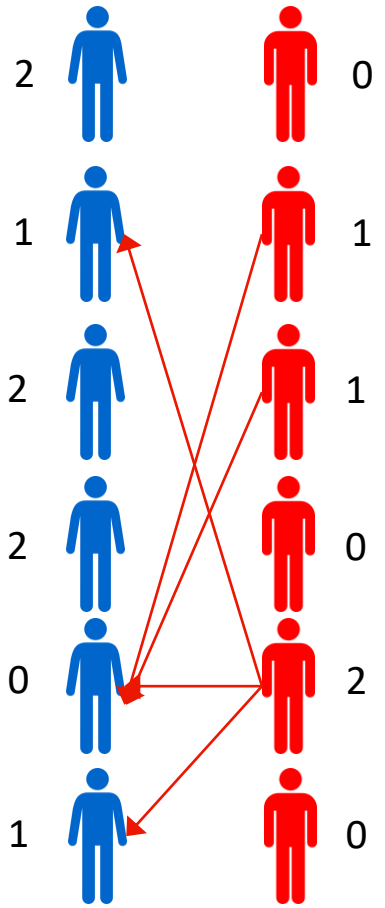
W=4, T=5, L=1



0 – non-responder
1 – responder

$$NB = N_W/N_P - N_L\ /N_P = 0.53$$

LOSSES=5



$$WO = \frac{N_W+0.5*N_{Tie}}{N_L+0.5*N_{Tie}} = \frac{25+2.5}{6+2.5} = 3.24$$

GPC Benefits

- **Increases flexibility of analyses**
- **Incorporates multiple outcomes without multiplicity adjustment**
- **Incorporates thresholds of clinical relevance**
- **May increase power as compared with single outcome**
- **Can be adapted to individual patient preferences**
- **Provides a unique measure of treatment effect that is meaningful to patients and caregivers**
- **Uses outcomes to analyze patients by comparing the experiences of trial participants in different treatment arms by the desirability of the overall patient outcome.**

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	Primary composite endpoint	Win ratio (95% CI)
ATTR-ACT[14]	Transthyretin amyloid cardiomyopathy	Tafamidis vs. placebo	All-cause mortality > number of heart failure hospitalizations	1.70 (1.26–2.29)
CHART-1[16]	LVEF ≤35%	Cardiopoietic stem cells vs. placebo	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	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	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	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	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	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.

Pocock et al. The win ratio approach for composite endpoints: practical guidance based on previous experience.
European Heart Journal (2020) 41, 4391–4399

Registered Trials in ClinicalTrials.gov

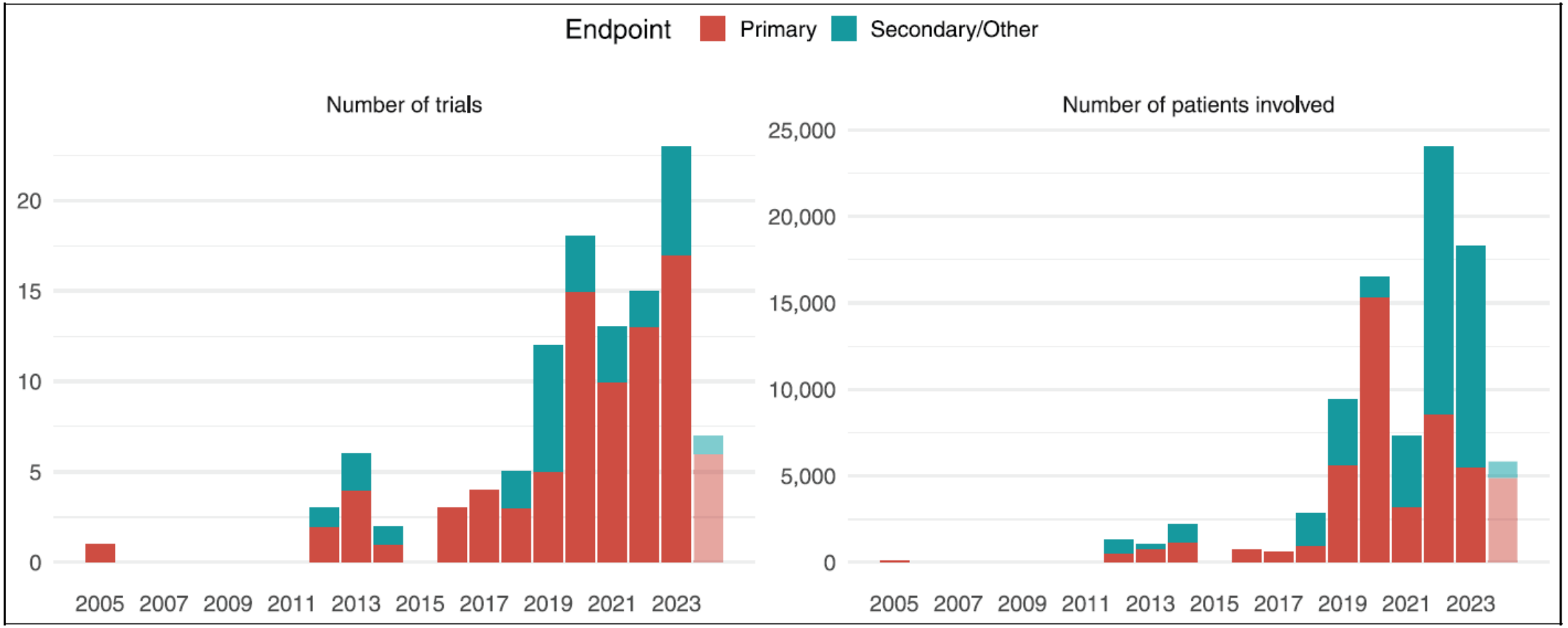


Figure 1. Registered trials (by start year) that specify win ratio-like approach to hierarchical composite endpoints in primary, secondary, or other analyses.
Source: ClinicalTrials.gov as of December 2023.

Mao (2024). Defining estimand for the win ratio: Separate the true effect from censoring. *Clinical Trials*. Vol. 21(5) 584–594

Conclusion Remarks

About rank-based approach

Advantages	Challenges
• Recognize clinical priority and timing of component events • Can be versatile, e.g., multiple event types, recurrence of events, different ranking approaches • Ranking can be determined through patient's preference survey for patient-centric decision making	• Ranking algorithm is subjective, so real interpretation can be difficult • Computational complexity can be quite high with large samples • Differential length of follow-up time and censoring patterns may lead to different conclusions (not unique for the newer method)

Recommendations

- Recent FDA Factor XI Workshop of May 2024 recommended rank-/weight-based methods for assessing both safety and efficacy endpoints in benefit-risk analysis
- Predefine and communicate early with HA the ranking algorithm
- Display contributions by each tie-breaking component in terms of win/loss/tie
- Present different win statistics together with win ratio, such as win odds (which considers ties), and net clinical benefit

References

References

Bakal, J. A., C. M. Westerhout, and P. W. Armstrong. 2012. "Impact of Weighted Composite Compared to Traditional Composite Endpoints for the Design of Randomized Controlled Trials." *Statistical Methods in Medical Research*. https://pubmed.ncbi.nlm.nih.gov/22275378/.

Bebu, Ionut, and John M. Lachin. 2016. "Large Sample Inference for a Win Ratio Analysis of a Composite Outcome Based on Prioritized Components." *Biostatistics*. https://pubmed.ncbi.nlm.nih.gov/26353896/.

Betts, M. B., P. Rane, E. Bergrath, M. Chitnis, M. K. Bhutani, C. Gulea, Y. Qian, and G. Villa. 2020. "Utility Value Estimates in Cardiovascular Disease and the Effect of Changing Elicitation Methods: A Systematic Literature Review." *Health Qual Life Outcomes*. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7385861/.

Brazier, J., R. Ara, I. Azzabi, J. Busschbach, H. Chevrou-Severac, B. Crawford, L. Cruz, et al. 2019. "Identification, Review, and Use of Health State Utilities in Cost-Effectiveness Models: An ISPOR Good Practices for Outcomes Research Task Force Report." *Value Health*. https://pubmed.ncbi.nlm.nih.gov/30832964/.

Buyse, M. 2010. "Generalized Pairwise Comparisons of Prioritized Outcomes in the Two-Sample Problem." *Statistics in Medicine*. https://pubmed.ncbi.nlm.nih.gov/21170918/.

Dodgson, J. S., M. Spackman, Pearman A., and L. D. Phillips. 2009. "Multi-Criteria Analysis: A Manual." *Department for Communities and Local Government, London*.

Dong, G., D. C. Hoaglin, B. Huang, Y. Cui, D. Wang, Y. Cheng, and M. Gamalo-Siebers. 2023. "The Stratified Win Statistics (Win Ratio, Win Odds, and Net Benefit)." *Pharm Stat*. https://pubmed.ncbi.nlm.nih.gov/36808217/.

Dong, G., D. C. Hoaglin, J. Qiu, R. A. Matsouaka, Y. W. Chang, J. Wang, and M. Vandemeulebroecke. 2019. "The Win Ratio: On Interpretation and Handling of Ties." *Statistics in Biopharmaceutical Research*. https://doi.org/10.1080/19466315.2019.1575279.

Dong, G., B. Huang, J. Verbeeck, Y. Cui, J. Song, M. Gamalo-Siebers, D. Wang, et al. 2022. "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." *Pharmaceutical Statistics*. https://pubmed.ncbi.nlm.nih.gov/35757986/.

Dong, G., D. Li, S. Ballerstedt, and M. Vandemeulebroecke. 2016. "A Generalized Analytic Solution to the Win Ratio to Analyze a Composite Endpoint Considering the Clinical Importance Order Among Components." *Pharmaceutical Statistics*. https://pubmed.ncbi.nlm.nih.gov/27485522/.

Dong, G., L. Mao, B. Huang, M. Gamalo-Siebers, J. Wang, G. Yu, and D. C. Hoaglin. 2020. "The Inverse-Probability-of-Censoring Weighting (IPCW) Adjusted Win Ratio Statistic: An Unbiased Estimator in the Presence of Independent Censoring." *J Biopharm Stat*. https://pubmed.ncbi.nlm.nih.gov/32552451/.

Dong, G., J. Qiu, D. Wang, and M. Vandemeulebroecke. 2018. "The Stratified Win Ratio." *J Biopharm Stat*. https://pubmed.ncbi.nlm.nih.gov/29172988/.

Duc, A. N., and M. Wolbers. 2016. "Weighted Analysis of Composite Endpoints with Simultaneous Inference for Flexible Weight Constraints." *Statistics in Medicine*. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5217097/.

Evans, S. R., D. Rubin, D. Follmann, G. Pennello, W. C. Huskins, J. H. Powers, D. Schoenfeld, et al. 2015. "Desirability of Outcome Ranking (DOOR) and Response Adjusted for Duration of Antibiotic Risk (RADAR)." *Clin Infect Dis*. https://pubmed.ncbi.nlm.nih.gov/26113652/.

Evans, S., and D. Follmann. 2016. "Using Outcomes to Analyze Patients Rather Than Patients to Analyze Outcomes: A Step Toward Pragmatism in Benefit-Risk Evaluation." *Statistics in Biopharmaceutical Research*. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5394932/.

Evans, S., M. Knutsson, P. Amarenco, G. W. Albers, P. M. Bath, H. Denison, P. Ladenvall, et al. 2020. "Methodologies for Pragmatic and Efficient Assessment of Benefits and Harms: Application to the SOCRATES Trial." *Clinical Trials*. https://pubmed.ncbi.nlm.nih.gov/32666831/.

Follmann, D., M. P. Fay, T. Hamasaki, and S. Evans. 2019. "Analysis of Ordered Composite Endpoints." *Statistics in Medicine*. https://pubmed.ncbi.nlm.nih.gov/31858640/.

Luo, X., J. Qiu, S. Bai, and H. Tian. 2017. "Weighted Win Loss Approach for Analyzing Prioritized Outcomes." *Stat Med*. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5490500/.

Luo, X., H. Tian, and W. Y. Mohanty S. & Tsai. 2014. "An Alternative Approach to Confidence Interval Estimation for the Win Ratio Statistic." *Biometrics*. https://pubmed.ncbi.nlm.nih.gov/25156540/.

Mao, L., and D. Y. Lin. 2016. "Semiparametric Regression for the Weighted Composite Endpoint of Recurrent and Terminal Events." *Biostatistics*. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4804115/.

Murray, T. A., P. F. Thall, and Y. Yuan. 2016. "Utility-Based Designs for Randomized Comparative Trials with Categorical Outcomes." *Statistics in Medicine*. https://pubmed.ncbi.nlm.nih.gov/27189672/.

Pocock, S., C. Ariti, T. Collier, and D. Wang. 2011. "The Win Ratio: A New Approach to the Analysis of Composite Endpoints in Clinical Trials Based on Clinical Priorities." *European Heart Journal*. https://pubmed.ncbi.nlm.nih.gov/21900289/.

Rauch, G., K. Kunzmann, M. Kiezer, K. Wegscheider, J. Konig, and C. Eulenburg. 2018. "A Weighted Combined Effect Measure for the Analysis of a Composite Time-to-First-Event Endpoint with Components of Different Clinical Relevance." *Statistics in Medicine*. https://pubmed.ncbi.nlm.nih.gov/29205425/.

Tervonen, T., J. Veldwijk, K. Payne, X. Ng, B. Levitan, L. G. Lackey, K. Marsh, et al. 2023. "Quantitative Benefit-Risk Assessment in Medical Product Decision Making: A Good Practices Report of an ISPOR Task Force." *Value Health*. https://pubmed.ncbi.nlm.nih.gov/37005055/.

Yu, R. X., and J. Ganju. 2022. "Sample Size Formula for a Win Ratio Endpoint." *Statistics in Medicine*. https://pubmed.ncbi.nlm.nih.gov/35084052/.

Yuan, Z., B. Levitan, H. Deng, M. Szarek, R. M. Bauersachs, S. D. Berkowitz, L. Haskell, E. S. Barnathan, and M. P. Bonaca. 2024. "Quantitative Benefit-Risk Evaluation of Rivaroxaban in Patients After Peripheral Arterial Revascularization." *J Am Heart Assoc*. https://pubmed.ncbi.nlm.nih.gov/38563380/.

Gregson et al 2023, JACC: Recurrent Events in Cardiovascular Trials

Wang et al 2023, Contemporary Clinical Trials: Statistical power considerations in the use of win ratio in cardiovascular outcome trials

Thank you