Win ratio: On interpretation and handling of ties

Gaohong Dong, iStats Inc.
Junshan Qiu, FDA/CDER
David C. Hoaglin, Univ. of Mass. Medical School
Roland Matsuoaka, Duke Univ. Sch. of Medicine
Yu-Wei (Victoria) Chang, Abbvie
Jiuzhou (Joe) Wang, ImmunoGen Inc.
Marc Vandemeulebroecke, Novartis Pharma AG

Regulatory-Industry Workshop, Sep. 25, 2019

Post-Workshop note

The following were added per some questions raised during the session:

- Algorithms (rules) defining winners (losers) and ties
- Win ratio estimands

Details of these two aspects can be found in the paper “The win ratio: On interpretation and handling of ties” (Dong et al., 2019)

STATISTICS IN BIOPHARMACEUTICAL RESEARCH
2019, VOL. 00, NO. 0, 1–8
https://doi.org/10.1080/19466315.2019.1575279

The Win Ratio: On Interpretation and Handling of Ties

Gaohong Dong[a], David C. Hoaglin[b], Junshan Qiu[c], Roland A. Matsouaka[d], Yu-Wei Chang[e], Jiuzhou Wang[f], and Marc Vandemeulebroecke[g]

JOURNAL OF BIOPHARMACEUTICAL STATISTICS
2018, VOL. 28, NO. 4, 778–796
https://doi.org/10.1080/10543406.2017.1397007

The stratified win ratio

Gaohong Dong [a], Junshan Qiu[b], Duolao Wang[c], and Marc Vandemeulebroecke[d]

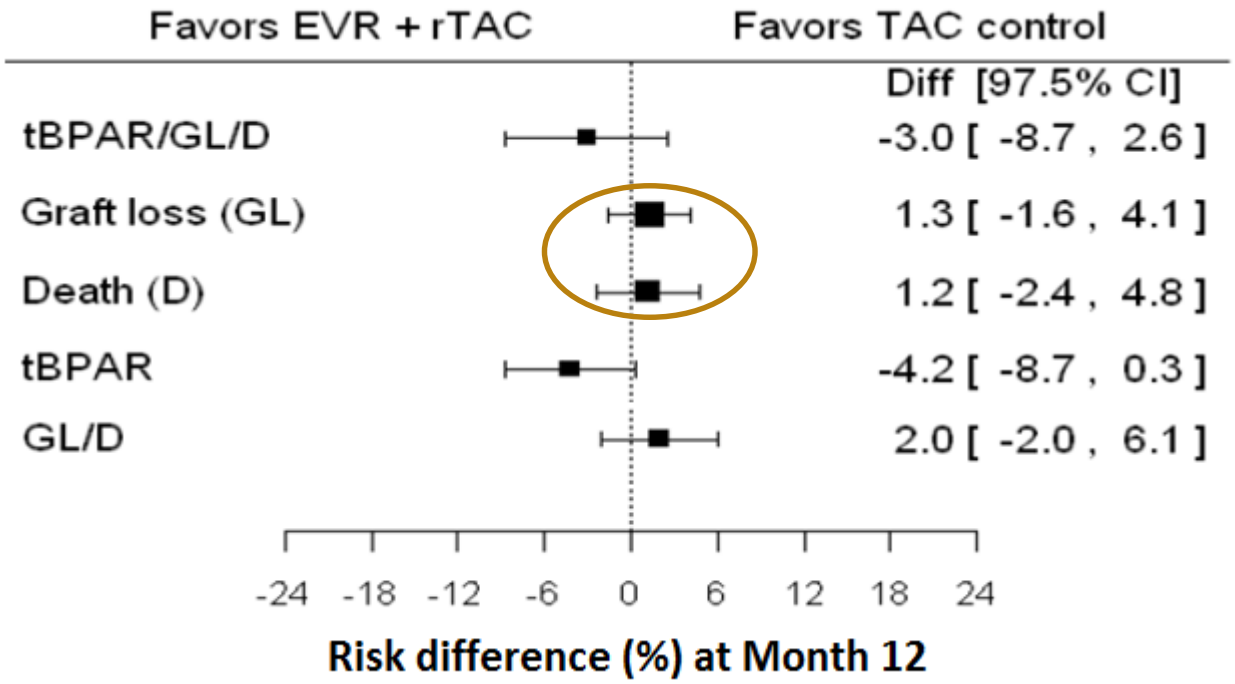
Outline

- Analyses of composite endpoints (two examples)
- Win ratio and stratified win ratio
- Algorithms (rules) defining winners (losers) and ties
- Interpretation: Win probability/proportion
- Interpretation: Connection to commonly used statistics
- Handling of ties and the win odds
- Win ratio estimands
- Best use of the win ratio
- Summary

Example 1: Phase III Liver Transplant Trial

- Phase III liver transplant trial: 488 = 243+245 patients randomized into the following two treatment groups
  - Treatment: EVR + reduced TAC exposure
  - Control: standard TAC exposure
- Primary endpoint is a composite of
  - Treated biopsy-proven acute rejection (tBPAR)
  - Graft loss (GL)
  - Death (D)
- This trial was completed in 2013

Example 1: Phase III Liver Transplant Trial (cont.)



Source: Dong et al. (2016), Saliba et al. (2013), De Simone et al. (2012), ClincalTrials.gov (NCT00622869)

Q1: Was the numerical difference in GL/D a chance finding?

Q2: Could we put more weight on more important outcomes GL and D?

Example 2: CHARM program

- CHARM program: included 3 separate randomized trials comparing candesartan with placebo in subjects with chronic heart failure (CHF).
- Primary endpoint: Composite of cardiovascular (CV) death or hospitalization for CHF.
- The three CHARM trials were completed in 2003 with 7599 subjects with median follow-up 3.14 years

Example 2: CHARM program (cont.)

	CHARM Added		CHARM Alternative		CHARM Preserved	
Adjusted HR	0.85		0.70		0.86	
95% CI	0.75–0.96		0.60–0.81		0.77–1.00	
P-value	0.010		<0.0001		0.051	
	C	Pl	C	Pl	C	Pl
No. of patients	1276	1272	1013	1015	1514	1509
No. with primary composite event	483	538	334	406	333	366
No. of these which were CV death[a]	174	182	127	120	92	90
Total no. with CV death[a]	302	347	219	252	170	170

Only 54% of CV deaths contributed to the composite

Q: Could all CV deaths be considered for the analysis?

Source: Pocock et al. (2012)

Composite endpoint

- Advantages:
  - Event rate is higher than those of the components alone.
    ==> Sample size can be reduced
  - Avoid multiplicity and competing risk
  - Captures multiple aspects of treatment effect

- Disadvantages:
  - All components are treated equally important
  - First event analyzed may not be the most important
  - Some components could show an opposite treatment effect
  - A less important component could be dominant

Win ratio

- Win ratio (Pocock et al., 2012)
  - Consider the most important outcome (e.g. death) first, then next important event, … etc.
  - Can handle a composite of multiple outcomes in any data type (e.g., time-to-event, ordinal, continuous, …)
  - Can handle non-proportional hazards situations
  - Enable project specific rules defining winners (losers) and ties
  - The name of 'win ratio' is intuitive

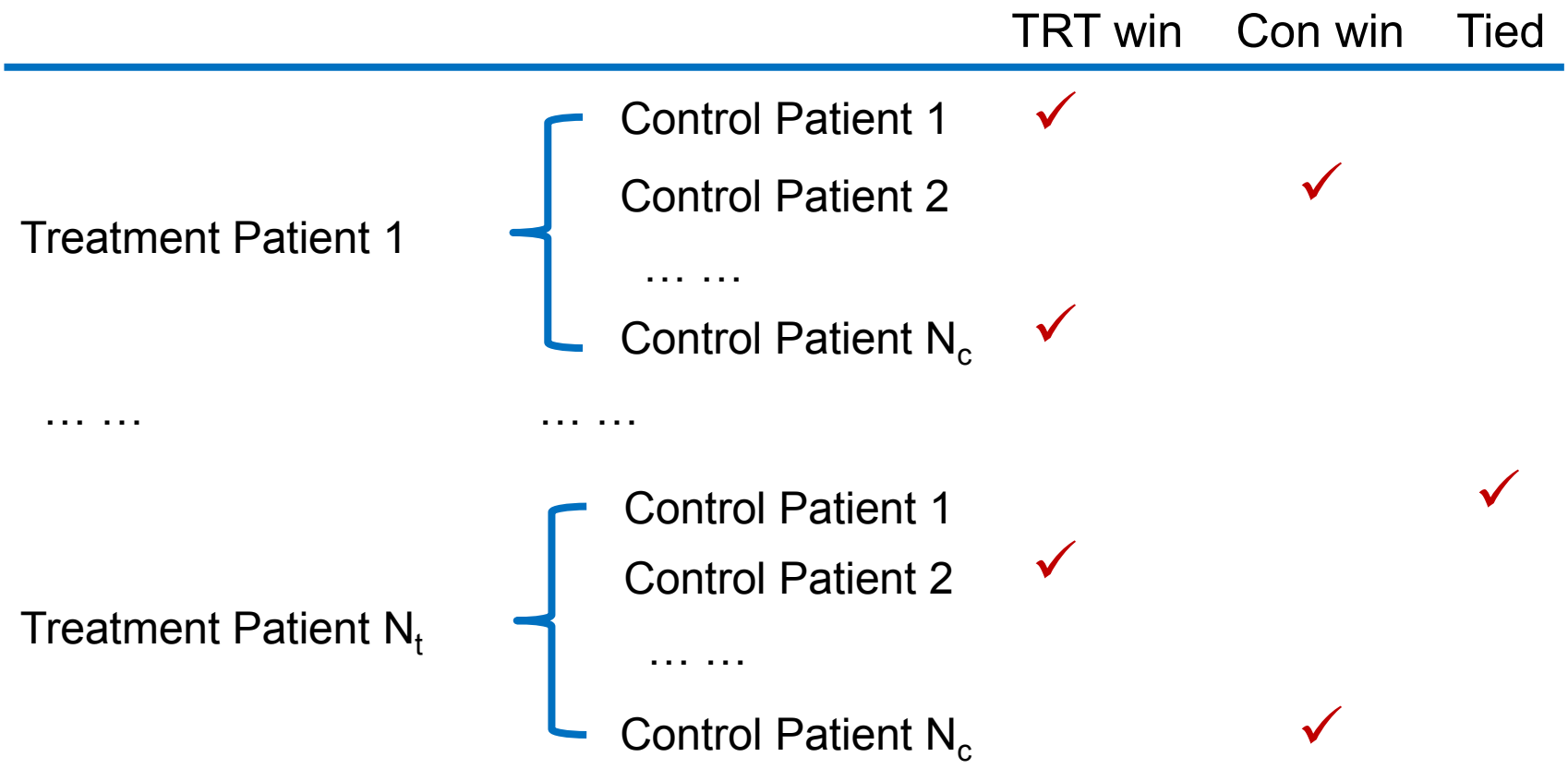
- Two closely related methods:
  - Net benefit (proportion in favor of treatment) by Buyse (2010).
  - Finkelstein and Schoenfeld (1999) : Combining mortality & longitudinal outcome

Recent methodological developments of the win ratio

- Pocock et al. (2012): Original proposal
- Luo et al. (2015): Closed-form variance estimator for a particular way of defining wins (or losses) and ties
- Bebu and Lachin (2016) and Dong et al (2016): Closed-form variance estimator for any way of defining wins (or losses) and ties
- Wang and Pocock (2016): Win ratio for non-normal outcomes
- Oakes (2016): Win and loss probabilities in a limited time c
- Luo et al. (2017): Weighted win and loss
- Dong et al. (2018): The stratified win ratio
- Finkelstein and Schoenfeld (2019): Win ratio as a function of time t.
- Mao (2019): on hypothesis of the win ratio
- Dong et al (2019): Win ratio: interpretation and handling of ties
- Dong et al (in press): Win ratio: Impact of censoring and follow-up time and use with non-proportional hazards

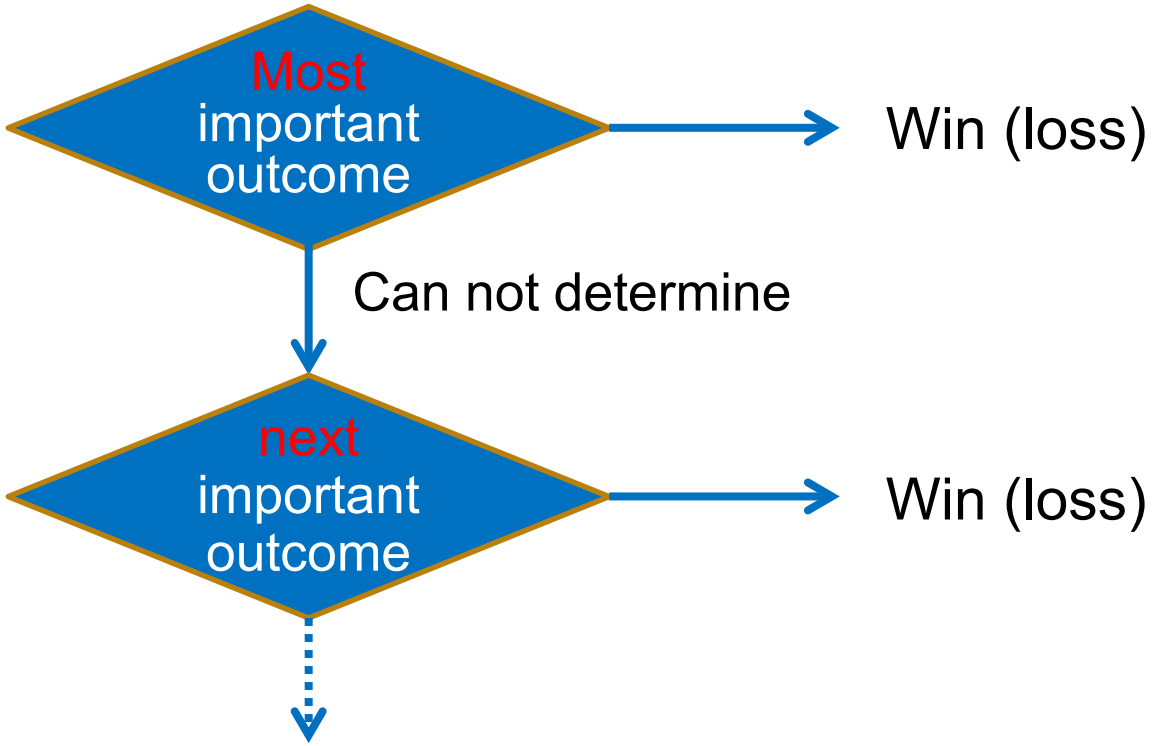
Win ratio

Based on pairwise comparisons: each patient in the Treatment group is compared with every patient in the Control group.



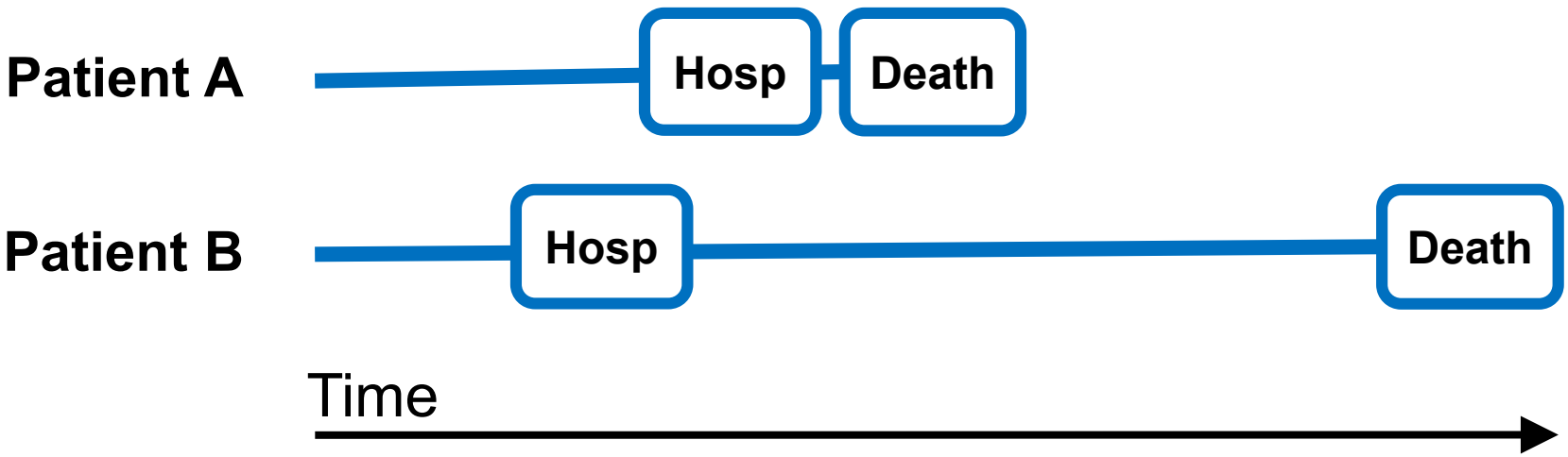
Win ratio (cont.)

- For each pair,



- $$Win\ ratio = \frac{Number\ of\ wins\ for\ Treatement}{Number\ of\ wins\ for\ Control}$$

Win ratio(cont.)



Who wins?

- Win ratio: Patient B wins on Death
- First-event analysis: Patient A wins on Hospitalization

Algorithms (rules) defining winners (losers) and ties

- No general rule can fit all disease indications
- Who is the winner?
  - Death on Day 100 vs. censored on Day 98
  - Hospitalized through the death on Day 100 vs. heathy and died on Day 98
- Examples:
  - Any difference is a difference
  - A minimal difference is required to be a win (Peron et al., 2016)
  - Dropouts were assumed alive given low mortality rate in the phase III liver transplant trial (Dong et al., 2016)
- Rules should be project specific
- Defining rules – a joint work of statistician and clinical team
- Rules should be prespecified in the analysis plan

Win ratio and stratified win ratio

- Let
  - X_i : patient i in Treatment $(i = 1, \ldots, N_t)$
  - Y_j : patient j in Control $(j = 1, \ldots, N_c)$
- Define

K(X_i , Y_j) = 1 if X_i wins over Y_j = 0 otherwise	L(X_i , Y_j) = 1 if Y_j wins over X_i = 0 otherwise
# of wins for the Treatment $n_t = \sum_{i=1}^{N_t} \sum_{j=1}^{N_c} [K(X_i, Y_j) = 1]$	# of wins for the Control $n_c = \sum_{i=1}^{N_t} \sum_{j=1}^{N_c} [L(X_i, Y_j) = 1]$

Win ratio and stratified win ratio (cont.)

$K(X_i, Y_j) = 1$ if X_i wins over Y_j $= 0$ otherwise	$L(X_i, Y_j) = 1$ if Y_j wins over X_i $= 0$ otherwise
# of wins for the Treatment $n_t = \sum_{i=1}^{N_t} \sum_{j=1}^{N_c} [K(X_i, Y_j) = 1]$	# of wins for the Control $n_c = \sum_{i=1}^{N_t} \sum_{j=1}^{N_c} [L(X_i, Y_j) = 1]$
Win ratio: $WR = n_t / n_c$	
Stratified win ratio: $WR = \frac{\sum_{m=1}^{M} w^{(m)} n_t^{(m)}}{\sum_{m=1}^{M} w^{(m)} n_c^{(m)}}$ where $w^{(m)}$ is the weight for m^{th} stratum (= 1, 2, ..., M).	

Stratified win ratio (cont.)

- We propose the stratified win ratio with $w^{(m)} = 1/N^{(m)}$, motivated by the Mantel-Haenszel method :

$$WR = \frac{\sum_{m=1}^{M} n_t^{(m)} / N^{(m)}}{\sum_{m=1}^{M} n_c^{(m)} / N^{(m)}}$$

- Why?
  - The win ratio can reduce to the odds ratio.
  - This stratified win ratio can reduce to the Mantel-Haenszel stratified odds ratio.
  - Mantel-Haenszel stratified odds ratio is similar or robust than other stratified odds ratios (e.g. logic, MLE).

Interpretation: Connections to commonly used statistics

Single outcome	Commonly used statistic	Connection
Binary	Odds ratio (OR)	WR = 1/OR
Time-to-event	Hazard ratio (HR)	WR = 1/HR, when the proportional hazards assumption is met.
Normally distributed continuous	Mean difference (δ)	WR = $f(\delta)$ ** Acion et al. (2006) and Buyse (2010)
Non-normal continuous	Mann-Whitney U	The number of wins (n_t and n_c) in the win ratio = Mann-Whitney U, when there are no ties.

Interpretation: Win probablity/proportion

- Win probability (Oakes, 2016)

$$\pi_t(x) = Prob\left(T^{(t)} > T^{(c)}\right) = -\int_0^x F^{(t)} dF^{(c)}$$

$$WR(x) = \frac{\pi_t(x)}{\pi_c(x)} = \frac{Prob\left(T^{(t)} > T^{(c)}\right)}{Prob\left(T^{(c)} > T^{(t)}\right)} = \frac{-\int_0^x F^{(t)} dF^{(c)}}{-\int_0^x F^{(c)} dF^{(t)}}$$

- Win proportions: estimates of win probabilities

$$\widehat{\pi_t} = P_t = n_t / N_t N_c, \quad \widehat{\pi_c} = P_c = n_c / N_t N_c$$

$$\widehat{WR} = P_t / P_c$$

- Proportions of "wins" are more readily interpretable than the number of "wins"
- Example: WR=1.2 means that a patient in the treatment group is 20% more likely to "win" than a patient in the control group

Interpretation: Win probablity/proportion(cont.)

- Example 1 of Phase III liver transplant trial: 243 vs. 245 patients randomized into Treatment vs. Control groups
- Number of pairs: $N_t N_c = 243 \times 245 = 59535$
- Numbers of wins and ties:

  $n_t = 5393$, $n_c = 3761$, and $n_{tie} = 50381$,

- Good to present win proportions instead:

$$P_t = \frac{n_t}{N_t N_c} = \frac{5393}{59535} = 9.1\% \text{ and } P_c = 6.3\%$$

$$\widehat{WR} = \frac{P_t}{P_c} = 1.4$$

- There are too many ties, but ignored in the win ratio calculation

Handling of ties and the win odds

- Mann-Whitney U:

$$U = \sum_{i=1}^{N_t}\sum_{j=1}^{N_c}\left[I(X_i \succ Y_j) + \frac{1}{2} I(X_i \approx Y_j) \right]$$

$$= \sum_{i=1}^{N_t}\sum_{j=1}^{N_c}\left[K(X_i, Y_j) + \frac{1}{2}\left(1 - K(X_i, Y_j) - L(X_i, Y_j)\right) \right]$$

$$U_2 = \frac{U}{N_t N_c}$$

- Win odds (WO) – an extension of the Mann-Whitney odds

$$WO = \frac{U_2}{1 - U_2} = \frac{n_t + 0.5 n_{tie}}{n_c + 0.5 n_{tie}} = \frac{P_t + 0.5 P_{tie}}{P_c + 0.5 P_{tie}}$$

Win Ratio vs. Win Odds

Win Ratio	Win Odds
$WR = \frac{n_t}{n_c} = \frac{P_t}{P_c}$	$WO = \frac{n_t + 0.5n_{tie}}{n_c + 0.5n_{tie}} = \frac{P_t + 0.5P_{tie}}{P_c + 0.5P_{tie}}$
WR=1.2 means that a patient in the treatment group is 20% more likely to "win" than a patient in the control group.	WO=1.2 means that a patient in the treatment group is 20% more likely to "win" or **"tie"** than a patient in the control group.

Example 3: PEACE trial

- PEACE trial: a cardiovascular study (reported in 2004) with low event rates and long follow-up time.
- Composite of CV death, MI, CABG, and PTCA.

Group	Stratum-specific win Proportion		Stratified win Prop.	Stratified win ratio (95% CI)	Stratified win odds (95% CI)	Hazard ratio in favor of treatment (95% CI)
	Female	Male				
Trandolapril	16.5%	18.1%	18.0%	1.03 (0.94, 1.13)	1.01 (0.98,1.04)	1.04 (0.94, 1.14)
Placebo	16.4%	17.5%	17.4%			

Example 4: ATTR-ACT trial (Pfizer)

- ATTR-ACT study: Compare tafamidis with placebo for the treatment of TTR-CM (transthyretin cardiomyopathy)
- Stratified by TTR genotype (variant vs. wild-type) and baseline disease severity (NYHA Class I/II vs. III)
- Primary composite endpoint of
  - All-cause mortality
  - Frequency of cardiovascular-related hospitalization
- Started: Dec 2013; Completed : Feb 2018
- Stratified win ratio (primary analysis with equal weight): 1.7 (95% CI:1.26, 2.29; P=0.0006 per Finkelstein and Schoenfeld (1999))
- Drug approval: Vyndaqel/Vyndamax (tafamidis) in early 2019

Source: Maurer et al. (2018), Pocock and Collier (2019)

Win ratio estimands

- Censoring impact on the win ratio
- What treatment effect is to be estimated
- Addendum of ICH E9 Guidelines
- Attributes of win ratio estimands:

1) Population: defined per inclusion/exclusion criteria
2) Variable (endpoint or outcome): a composite endpoint of *prioritized multiple outcomes* of interest
3) Intercurrent event: events that occur after treatment initiation, e.g., rescue medication, discontinuation, …etc, may cause informative censoring
4) Summary measure: the (stratified or weighted) win ratio statistic to compare the two treatment groups

Win ratio estimands (cont.)

- Intercurrent events:
  a) Dropouts:
     - Consider discontinuations as outcomes in the composite (composite strategy, ICH E9 Addendum)
     - Different discontinuations (e.g., AE vs. lack of efficacy) may be considered in different ways
  b) Rescue medication and switching study treatments: may be considered similarly to that for treatment discontinuation
  c) Missing endpoints: Perform imputations first, then calculate the win ratio
  d) etc..

Best use of the win ratio

- When the more important endpoints tend to occur later and are reasonably frequent
  - May provide a considerable benefit, when a substantial number of patients experience more than one endpoint
  - Would have minimal advantage, when more important outcomes are infrequent.
  - The follow-up time should be long enough to yield a useful number of occurrences of the more important endpoints
- **Censoring times** are similar in the two groups, unless the **event times** follow a proportional-hazards model.

Source: Dong et al. (in press): Win ratio: Impact of censoring and follow-up time and use with non-proportional hazards. *Pharmaceutical Statistics*

Summary

- Conventional analysis uses time to the first event. The first event analyzed may not be the most important outcome
- Win ratio is the ratio of numbers of wins for two groups, it is also the ratio of win probabilities (proportions) for two groups. It has a connection to some commonly used statistics.
- Win odds assigns 50% of ties to the numerator and the denominator, it is always closer to the null value 1.0, especially when the number of ties is large.
- Win ratio and the win odds are for time to the worst event. They consider the importance order of multiple outcomes. They provide an alternative way to analyze composite endpoints.
- Win ratio has been applied for study designs; one drug approval has been obtained, with the win ratio/Finkelstein-Schoenfeld test as the primary analysis.

Acknowledgements

- L.J. Wei (Harvard Univ. School of Public Health) for my first win ratio work
- Bo Huang (Pfizer)
- James Song (Beigene)
- Yodit Seifu (Merck)