

# Hierarchical testing of composite endpoints: applying the win ratio to percutaneous coronary intervention versus coronary artery bypass grafting in the SYNTAX trial



Milan Milojevic<sup>1</sup>, MD, MSc; Stuart J. Head<sup>1\*</sup>, MD, PhD; Eleni-Rosalina Andrinopoulou<sup>1,2</sup>, PhD; Patrick W. Serruys<sup>3</sup>, MD, PhD; Friedrich W. Mohr<sup>4</sup>, MD, PhD; Jan G. Tijssen<sup>5</sup>, PhD; A. Pieter Kappetein<sup>1</sup>, MD, PhD

1. Department of Cardiothoracic Surgery, Erasmus University Medical Center, Rotterdam, the Netherlands; 2. Department of Biostatistics, Erasmus University Medical Center, Rotterdam, the Netherlands; 3. Department of Cardiology, Erasmus University Medical Center, Rotterdam, the Netherlands; 4. Department of Cardiovascular Surgery, Herzzentrum Universität Leipzig, Leipzig, Germany; 5. Academic Medical Center, Amsterdam, the Netherlands

GUEST EDITOR: William Wijns, MD, PhD; The Lambe Institute for Translational Medicine and Curam, National University of Ireland, Galway and Saolta University Healthcare Group, Galway, Ireland

## KEYWORDS

- clinical trials
- composite endpoints
- death
- multivessel disease
- myocardial infarction
- win ratio

## Abstract

**Aims:** The goal of the study was to compare long-term outcomes of percutaneous coronary intervention (PCI) versus coronary artery bypass grafting (CABG), accounting for the clinical impact of individual components in the composite endpoints and prioritising these using the win ratio (Rw).

**Methods and results:** The win ratio was compared with conventional methods of analyses (hazard ratio [HR] and relative risk) in the SYNTAX trial (n=1,800). For the composite of death/stroke/myocardial infarction (MI), the win ratio favoured CABG and was 1.37 (95% CI: 1.10-1.77) for matched analysis, 1.28 (95% CI: 1.11-1.53) for unmatched analysis, while the conventional HR was 1.29 (95% CI: 1.11-1.53). The largest number of winners in favour of CABG over PCI were based on MI (n=39 vs. n=19, respectively). Death was significantly reduced with CABG in matched (Rw=1.39, 95% CI: 1.04-1.86) and unmatched win ratio analyses (Rw=1.27, 95% CI: 1.01-1.42) as compared with non-significant conventional analysis (HR 1.19, 95% CI: 0.92-1.56). In subgroups, matched win ratio analyses had a larger treatment effect in favour of CABG compared with conventional analyses, especially in patients with three-vessel disease and intermediate SYNTAX scores, while unmatched win ratios had a smaller point estimate, but with narrower confidence intervals than matched analyses findings.

**Conclusions:** This re-analysis of the SYNTAX trial using the win ratio shows that the most important benefit of CABG treatment is the reduction of hard clinical endpoints such as mortality and MI. Future trials using this approach can expect to maintain similar statistical power with smaller sample sizes, and thereby reduce the cost of a trial. ClinicalTrials.gov Identifier: NCT00114972

\*Corresponding author: Erasmus MC, 's-Gravendijkwal 230, 3015 CE Rotterdam, the Netherlands.

E-mail: s.head@erasmusmc.nl

## Abbreviations

<b>DM</b>	medically treated diabetes mellitus
<b>HR</b>	hazard ratio
<b>LM</b>	left main
<b>MACCE</b>	major adverse cardiac and cerebrovascular events
<b>MI</b>	myocardial infarction
<b>PCI</b>	percutaneous coronary intervention
<b>Rw</b>	win ratio
<b>3VD</b>	three-vessel disease

## Introduction

Coronary artery bypass grafting (CABG) and percutaneous coronary intervention (PCI) have been compared in many randomised clinical trials<sup>1</sup>. These trials often used composite endpoints to obtain higher event rates and provide more statistical power, thus requiring smaller sample sizes, shorter follow-up, or both<sup>2,3</sup>. However, composite endpoints are often criticised for having an intrinsic weakness, combining events with a very different impact on a patient's quality of life or life expectancy. The reporting of composite endpoints in clinical trials also has an inherent limitation in that it emphasises each patient's first event, which is often the outcome of lesser clinical importance.

The SYNTAX trial assessed the optimum revascularisation treatment for patients with *de novo* left main (LM) coronary disease and/or three-vessel disease (3VD), by randomly assigning patients to either CABG or PCI with a first-generation drug-eluting stent (DES). The primary endpoint was powered on non-inferiority of PCI versus CABG for the endpoint of major adverse cardiac and cerebrovascular events (MACCE), which is the composite of all-cause death, stroke, myocardial infarction (MI), and repeat revascularisation<sup>4</sup>. The difference in MACCE between PCI and CABG was largely driven by higher rates of repeat revascularisation with PCI, which is thought to be a softer, less important endpoint, while rates of all-cause death were not significantly different between CABG and PCI<sup>5,6</sup>.

To overcome this weakness of putting the same emphasis on individual components with a clinically different impact in the composite endpoint, a recent novel approach, the win ratio, has been introduced<sup>7</sup>. Based on clinical priorities, the win ratio methodology applies a hierarchical weighting to individual components in MACCE. This approach is also designed to combat two fundamental difficulties that may be present in typical efficacy studies: study population heterogeneity and important events that are censored. The method uses risk score stratification to select and match pairs with similar risk profiles from both treatment groups and provides a more patient-specific interpretation of composite endpoints in clinical trials.

The objective of this re-analysis of the SYNTAX trial was to compare PCI with CABG using different methods of analysis, accounting for the severity of the individual components and prioritising these using the win ratio approach as an informative estimate of treatment difference. Moreover, this paper evaluates the impact of applying the win ratio on the design of future trials.

## Methods

### STUDY DESIGN

The design and methods of the SYNTAX trial have been reported previously<sup>8</sup>. In SYNTAX, 1,800 patients with *de novo* LM or three-vessel coronary artery disease were randomly assigned to undergo CABG or PCI with a paclitaxel-eluting stent (TAXUS<sup>®</sup> Express<sup>™</sup>; Boston Scientific, Marlborough, MA, USA). Patients with anticipated clinical revascularisation equipoise through PCI and CABG were randomised (CABG n=897, PCI n=903). Five-year follow-up was 89.7% for CABG and 96.5% for PCI.

This study was carried out in accordance with the principles of the Declaration of Helsinki, and registered on the National Institutes of Health website with identifier NCT00114972.

### DEFINITIONS

The primary endpoint of the trial was MACCE, which included all-cause death, stroke, MI, or repeat revascularisation (subsequent CABG or PCI)<sup>9</sup>. Secondary endpoints consisted of: i) a composite safety endpoint of death/stroke/MI, and ii) the individual endpoint of all-cause death. Definitions of these endpoints have been published elsewhere<sup>8</sup>.

Medically treated diabetes was defined as treatment with oral hypoglycaemic agents or insulin at the time of enrolment.

### STATISTICAL ANALYSIS

All analyses were carried out according to the intention-to-treat principle. Conventional analyses were performed using: i) Cox proportional hazard analyses to provide hazard ratios (HRs), and ii) estimates of relative risk (RR) associated with PCI versus CABG treatment. The proportional hazards assumption was estimated using Schoenfeld's test and was found to have been met. Relative risks were calculated by dividing the Kaplan-Meier estimated rate of an event at five years in the PCI group by the event rate in the CABG group. The 95 percent confidence interval (CI) for the relative risk was calculated with the use of the standard errors from the Kaplan-Meier curve<sup>10</sup>. The significance of differences in event rates between treatment groups was assessed with the use of the log-rank test. Conventional analyses were performed using SPSS software, Version 20.0 (IBM Corp., Armonk, NY, USA).

Win ratio analyses were performed for all-cause death, the composite safety endpoint of death/stroke/MI, and for the composite of MACCE<sup>7</sup>. The hierarchy of events within MACCE was as follows: all-cause death, stroke, MI, repeat CABG, repeat PCI.

The win ratio can be used in a matched or unmatched fashion, depending on how the patients are compared. As recommended, priority was given to a matched approach versus an unmatched approach that dilutes the win ratio<sup>7</sup>. In the matched approach, each patient in the CABG group was matched to a patient in the PCI group based on a similar risk of death. A risk score to predict death was developed using 18 pre-selected baseline variables that are known to be associated with prognosis (**Table 1**). A receiver operating characteristic (ROC) curve was generated to assess the ability of the scoring system model to predict mortality,

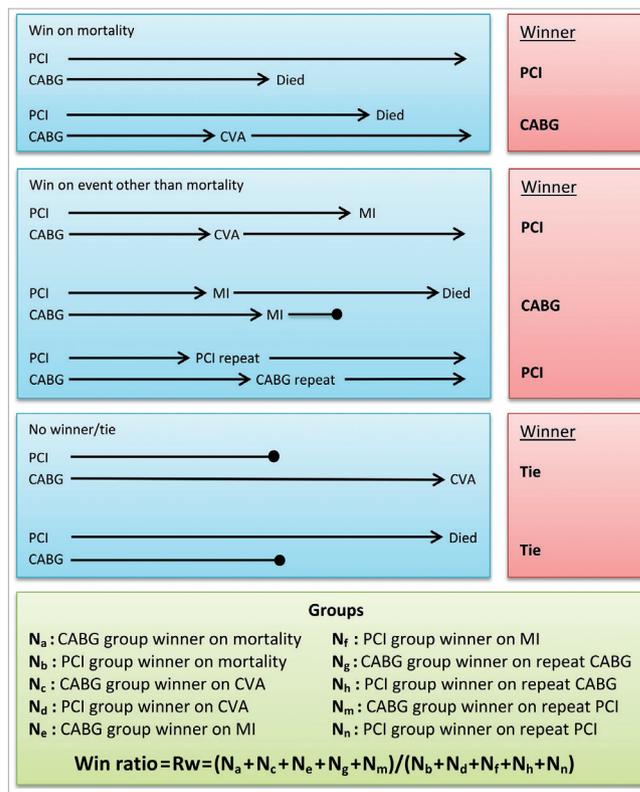
**Table 1. Pre-selected variables for risk score model.**

	Variable	Hazard ratio p-value*
Age	1.072	<0.001
Male sex	0.900	0.75
Body mass index	1.011	0.48
Current smoker	1.656	0.007
Medically treated diabetes	1.465	0.10
Previous myocardial infarction	1.370	0.030
Previous stroke	1.766	0.50
Previous transient ischaemic attack	0.825	0.80
Carotid artery disease	1.013	0.96
Congestive heart failure	1.166	0.88
Pulmonary hypertension	1.044	0.87
Peripheral vascular disease	2.513	<0.001
Creatinine >200 micromol/L	1.433	0.054
Dialysis	1.414	0.67
Chronic obstructive pulmonary disease	1.689	0.007
Ejection fraction moderate	1.405	0.040
Ejection fraction poor	1.371	0.006
SYNTAX score	1.010	0.093
Emergency treatment	0.707	0.58

\*p-value <0.05 is considered as statistically significant.

0.71 (95% CI: 0.67-0.75; p<0.0001). From the Cox proportional model’s coefficients, a risk score was calculated for each patient in the trial. Patients in the two treatment groups formed matched pairs based on their risk profiles and ranks. For each pair, the new treatment is a “winner” or “loser” according to who had died first (**Figure 1**). If no deaths occurred, a “winner” or “loser” was designated based on who first had a stroke, and so forth using the hierarchy of events. If one patient had an event but the follow-up period of the matched patient was shorter or if there were pairs without an event, they were considered “tied”. A “winner” patient had a more favourable outcome than his matched pair. The “win ratio” is the number of winners in the CABG group divided by the number of winners in the PCI group (**Figure 1**). An estimated win ratio >1 indicates a positive outcome of the CABG treatment compared to PCI while a win ratio <1 indicates a difference between treatment groups in favour of PCI. A corresponding 95% CI and p-values were calculated using dedicated statistical methods, as described by Pocock and co-authors<sup>7</sup>.

Unmatched analyses were performed for subgroups according to diabetes, LM disease and SYNTAX score to compare this result with the matched approach. Due to the fact of unequal treatment groups in subgroups, some patients had to be excluded randomly to provide equal numbers for matching. In smaller subgroups within pre-specified subgroups of patients, up to 17% had to be excluded. To determine the impact of randomly excluding patients, a repeated analysis was performed 10 times to examine



**Figure 1.** A conceptual diagram illustrating possible scenarios for the win ratio method. The determination of a “winner” is made using a predefined hierarchical outcome scheme. In the SYNTAX trial, mortality is considered the most important outcome followed by stroke, MI, repeat CABG revascularisation and repeat PCI revascularisation. The length of each arrow presents the duration of patient follow-up. Arrows ending in a solid circle denote either incomplete or shorter duration of follow-up.

whether the obtained results were affected: this was performed in the subgroup of patients with LM disease and an intermediate SYNTAX score (n=190). The results of the 10 analyses were very different, with the win ratio ranging from 0.31-0.50 with p-values ranging from 0.0050-0.1052 for all-cause death, and, respectively, 0.64-0.96 and 0.13-0.90 for MACCE. Therefore, only unmatched analyses were performed for the smaller SYNTAX score subgroups within subgroups of patients with LM/3VD and diabetes; no patients needed to be excluded for the unmatched analyses<sup>11</sup>. In the unmatched approach, a CI for the win ratio cannot be directly calculated: the bootstrap method with 1,000 samples was performed to determine significance and CIs, using R software version 3.2.4 (Institute for Statistics and Mathematics of WU, Vienna, Austria). A p-value <0.05 was considered to be statistically significant for all analyses.

While originally the win-ratio code for unmatched analyses was designed for composite endpoints that include two components (death and re-hospitalisation), in order to analyse the SYNTAX trial data with five endpoints, the statistical code was rewritten,

tested and validated according to the original statistical software provided directly by the authors of the win ratio approach<sup>7</sup>, when we requested and as they recommended. Now, this code can be used to calculate the win ratio with any number of components in the composite endpoint (for the code contact e.andrinopoulou@erasmusmc.nl).

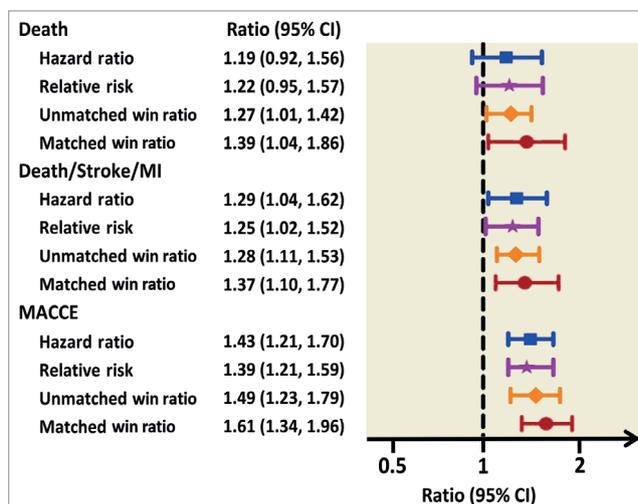
## Results

### OVERALL COHORT

Of the 1,800 patients randomly assigned to PCI or CABG, 880 matched pairs were computed based on the risk score.

For the primary outcome of MACCE at five years, 274 patients who underwent CABG won versus 170 patients who underwent PCI, corresponding to a matched win ratio of 1.61 (95% CI: 1.34-1.96;  $p < 0.0001$ ) (Table 2, Figure 2). In comparison with matched analyses, unmatched analyses tended to have a smaller ratio between CABG and PCI (Figure 2).

For the composite safety endpoint of death/stroke/MI, the win ratios and the conventional result were similar. The matched win ratio for the composite safety endpoint of death/stroke/MI was the analysis with the largest relative difference between CABG versus



**Figure 2.** Win ratio approach vs. conventional analyses for the overall SYNTAX randomised cohort. Different colours represent conventional time-to-event hazard ratio analyses (blue), relative risk (purple), unmatched win ratio (orange), matched win ratio (red), approaches with 95% CIs. MACCE: major adverse cardiac and cerebrovascular events

**Table 2.** The win ratio matched pairs approach, for the overall cohort and subgroups according to SYNTAX score terciles.

Matched pairs	SYNTAX trial	SYNTAX score $\leq 22$	SYNTAX score 23-32	SYNTAX score $\geq 33$
Death on PCI first	111	21	41	52
Death on CABG first	80	21	28	28
Stroke on PCI first	11	3	3	4
Stroke on CABG first	17	8	6	3
MI on PCI first	39	15	14	11
MI on CABG first	19	5	3	8
Repeat CABG on PCI first	24	9	10	9
Repeat CABG on CABG first	2	2	0	0
Repeat PCI on PCI first	89	29	27	29
Repeat PCI on CABG first	52	21	13	16
None of the above	437	137	149	127
Total number of pairs	880	272	294	287
<b>Win ratio for MACCE</b>	<b>1.61</b>	<b>1.38</b>	<b>1.90</b>	<b>1.91</b>
95% CI	1.34, 1.96	0.98, 1.87	1.37, 2.73	1.42, 2.75
Z-score	5.09	1.85	3.92	4.29
p-value	0.0001	0.064	0.0001	0.0001
<b>Win ratio for death/stroke/MI</b>	<b>1.39</b>	<b>1.15</b>	<b>1.57</b>	<b>1.72</b>
95% CI	1.10, 1.77	0.72, 1.83	1.05, 2.43	1.18, 2.62
Z-score	2.74	0.55	2.21	2.82
p-value	0.006	0.58	0.027	0.005
<b>Win ratio for all-cause death</b>	<b>1.39</b>	<b>1.00</b>	<b>1.46</b>	<b>1.86</b>
95% CI	1.04, 1.86	0.54, 1.86	0.92, 2.45	1.20, 3.07
Z-score	2.31	0	1.57	2.81
p-value	0.021	>0.99	0.12	0.005

CABG: coronary artery bypass grafting; CI: confidence interval; MACCE: major adverse cardiac and cerebrovascular events; MI: myocardial infarction; PCI: percutaneous coronary intervention

PCI in favour of CABG (n=116 vs. n=161, respectively;  $R_w=1.37$ , 95% CI: 1.10-1.77;  $p=0.006$ ). Out of 578 (65.7%) matched pairs that were tied for the composite of death/stroke/MI, 552 pairs did not have an event of death/stroke/MI during follow-up, and 26 patients were tied because of a different length of follow-up (**Table 2, Figure 2**).

Death occurred first after CABG in 80 patients and first after PCI in 111 patients; the unmatched win ratio for all-cause death was 1.27 (95% CI: 1.01-1.42) and the matched win ratio was 1.39 (95% CI: 1.04-1.86), which was statistically significant unlike conventional analyses that resulted in an HR of 1.19 (95% CI: 0.92-1.56) (**Table 2, Figure 2**).

## SUBGROUP ANALYSES

### SYNTAX SCORE

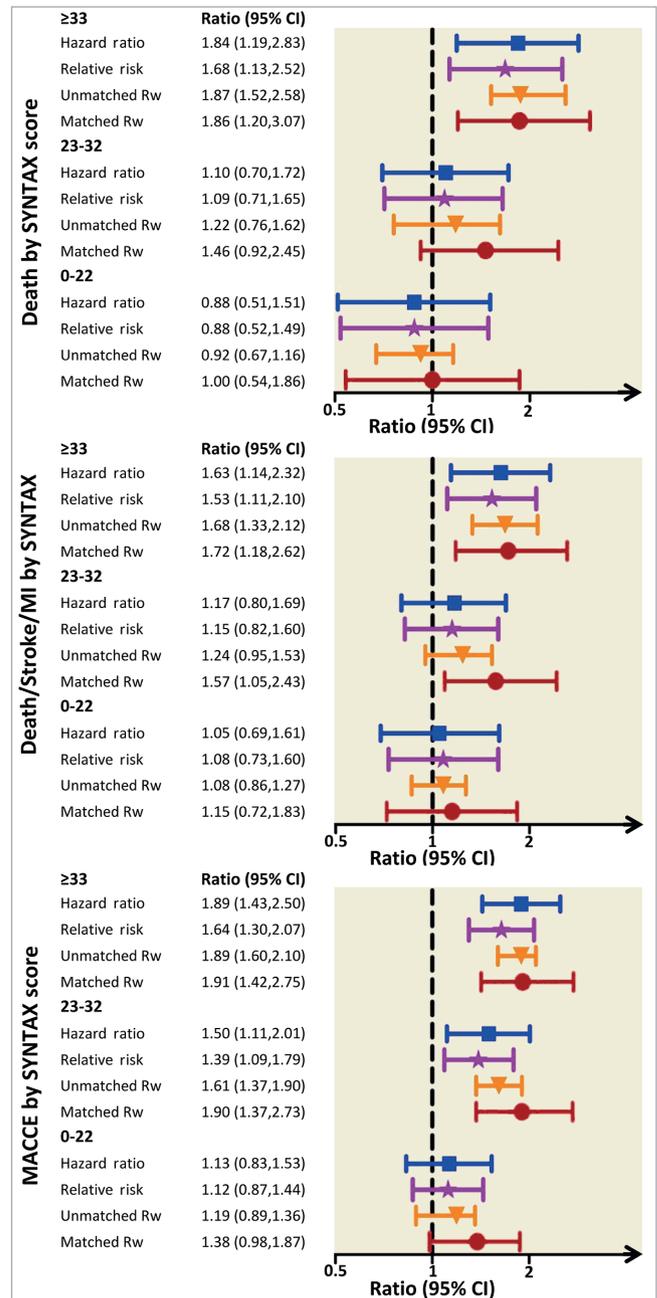
In patients with intermediate and high SYNTAX scores, the matched win ratio for MACCE at five years confirmed statistically significant better outcomes with CABG ( $R_w=1.90$ , 95% CI: 1.37-2.73 and  $R_w=1.91$ , 95% CI: 1.42-2.75, respectively) (**Figure 3, Table 2**). Nevertheless, the magnitude of the treatment effect was larger with the matched win ratio; even in the group of patients with a low SYNTAX score, there was a trend towards a difference.

For the composite endpoint of death/stroke/MI, the matched win ratio increased significantly in favour of CABG from low to intermediate to high SYNTAX scores ( $R_w=1.15$  vs.  $R_w=1.57$  vs.  $R_w=1.72$ ) as well as for all-cause death ( $R_w=1.00$  vs.  $R_w=1.46$  vs.  $R_w=1.86$ ) (**Figure 3, Table 2**). However, the treatment effect of PCI versus CABG was strongest with the matched win ratio, and particularly for subgroups of patients with intermediate SYNTAX scores where there was a clear increase in the treatment effect. In comparison with conventional analyses, the findings from the unmatched analyses were similar in patients with low and high SYNTAX scores, but were stronger in favour of CABG for patients with intermediate SYNTAX scores (**Figure 3**).

### LM/3VD

In patients with LM disease, the matched win ratio was not significantly different between CABG versus PCI: 1.27 (95% CI: 0.95-1.67) for MACCE, 1.02 (95% CI: 0.66-1.59) for the composite safety endpoint of death/stroke/MI, and 1.00 (95% CI: 0.70-1.44) for all-cause death (**Figure 4, Table 3**). In contrast, in patients with three-vessel disease, the matched win ratio for MACCE ( $R_w=1.92$ , 95% CI: 1.70-2.19), the composite safety endpoint of death/stroke/MI ( $R_w=2.00$ , 95% CI: 1.46-2.86), and all-cause death ( $R_w=2.15$ , 95% CI: 1.46-3.39) were all in favour of CABG. The unmatched approach supports the findings derived from conventional analyses in patients with LM, while unmatched analyses were stronger in favour of CABG among patients with three-vessel disease (**Figure 4**).

When separately analysing SYNTAX score subgroups for MACCE, differences between conventional analyses, the unmatched analyses, and matched analyses were only minimal (**Table 4**). Of note, there was no consistency in changes in PCI



**Figure 3.** Win ratio approach vs. conventional analyses by baseline SYNTAX score terciles. MACCE: major adverse cardiac and cerebrovascular events

versus CABG treatment effects choosing conventional or any win ratio analyses, although CIs appeared smaller when using win ratio analyses.

### DIABETES

In diabetic as well as non-diabetic patients, results using matched and unmatched win ratio approaches were comparable to those from conventional analyses (**Figure 4, Table 3**). In diabetic patients, MACCE was significantly lower in favour of CABG with a matched win ratio of 1.71 (95% CI: 1.19-2.52;  $p=0.003$ ), while all-cause death and the composite of death/stroke/MI were

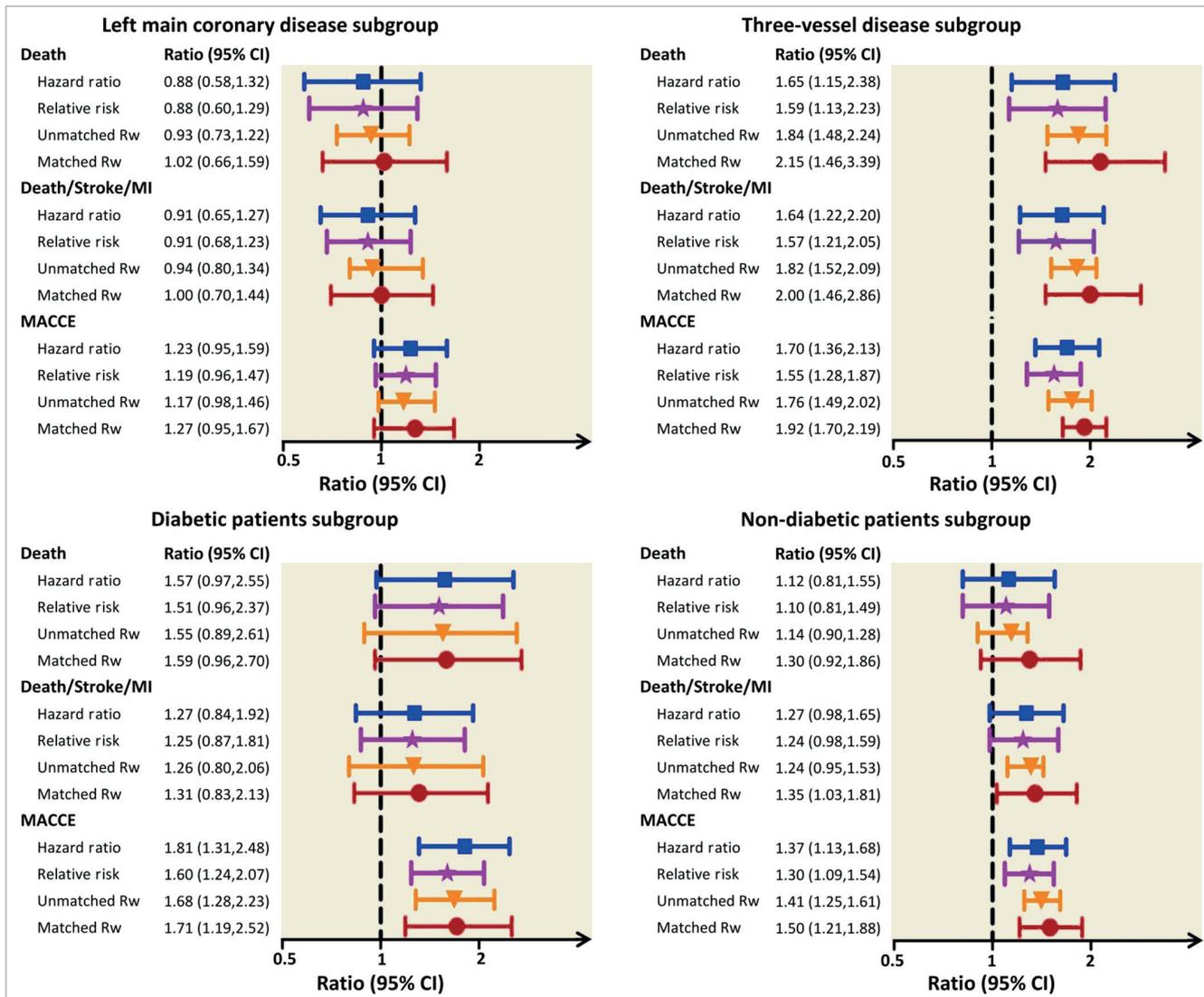


Figure 4. Win ratio approach vs. conventional analyses by baseline subgroups. MACCE: major adverse cardiac and cerebrovascular events

not significantly different between CABG and PCI. In non-diabetic patients, results with the matched win ratio approach slightly favoured CABG in comparison to conventional analyses, although these differences were minimal.

In separate analyses applying the unmatched win ratio approach to SYNTAX score terciles, the overall results of diabetics and non-diabetics were consistent with conventional analyses (Table 4). There was no consistency in increasing or decreasing the treatment effect of PCI versus CABG when using the win ratio approach.

## Discussion

The current analysis demonstrates that, by hierarchically prioritising events in the composite of MACCE, the treatment effect of CABG versus PCI is larger than with conventional analyses. In smaller subgroups of patients, for which unmatched win ratio analyses are necessary, differences between the win ratio approach and conventional analyses are minimal. These results provide

additional insights into the SYNTAX trial results and have several important implications for future trial conduct.

## PCI VERSUS CABG

The use of composite endpoints in trials is problematic because it may provoke controversy regarding their suitability<sup>12</sup>. Components are often unreasonably combined<sup>12,13</sup>, results are difficult to interpret<sup>14-16</sup>, and favourable outcomes or combinations of outcomes are cherry picked<sup>6,17</sup>. The criticism of the PCI versus CABG trials is that the superiority of CABG is primarily driven by repeat revascularisation which has less of a clinical impact than all-cause death, stroke and MI<sup>5</sup>. When repeat revascularisation is not part of the composite endpoint, there is no statistically significant difference between PCI and CABG in many trials. A meta-analysis of four trials comparing PCI with stents versus CABG also did not show a difference in rates of death/stroke/MI between CABG and PCI<sup>18</sup>. However, overall MACCE rates at five years were

**Table 3. The win ratio matched pairs approach, according to subgroups of left main disease and diabetes.**

Matched pairs	Diabetes		Coronary disease	
	DM	Non-DM	3VD	LM
Death on PCI first	38	74	71	42
Death on CABG first	24	57	33	41
Stroke on PCI first	0	7	8	3
Stroke on CABG first	6	11	8	10
MI on PCI first	4	34	25	15
MI on CABG first	2	17	11	9
Repeat CABG on PCI first	8	19	10	8
Repeat CABG on CABG first	0	3	2	1
Repeat PCI on PCI first	72	62	57	35
Repeat PCI on CABG first	13	43	35	20
None of the above	95	336	278	152
Total number of pairs	217	663	538	336
<b>Win ratio for MACCE</b>	<b>1.71</b>	<b>1.50</b>	<b>1.92</b>	<b>1.27</b>
95% CI	1.19, 2.52	1.21, 1.88	1.70, 2.19	0.95, 1.67
Z-score	3.01	3.70	5.45	1.64
p-value	0.003	0.0002	<0.0001	0.10
<b>Win ratio for death/stroke/MI</b>	<b>1.31</b>	<b>1.35</b>	<b>2.01</b>	<b>1.01</b>
95% CI	0.83, 2.13	1.03, 1.81	1.46, 2.86	0.70, 1.44
Z-score	1.17	2.15	4.42	0
p-value	0.24	0.032	<0.0001	>0.99
<b>Win ratio for all-cause death</b>	<b>1.59</b>	<b>1.30</b>	<b>2.15</b>	<b>1.02</b>
95% CI	0.96, 2.70	0.92, 1.86	1.46, 3.39	0.66, 1.59
Z-score	1.82	1.50	4.01	0.11
p-value	0.069	0.13	<0.001	0.91

CABG: coronary artery bypass grafting; CI: confidence interval; DM: medically treated diabetes; LM: left main coronary disease; MACCE: major adverse cardiac and cerebrovascular events; MI: myocardial infarction; PCI: percutaneous coronary intervention; 3VD: three-vessel disease

significantly lower in CABG patients as a result of persistently lower repeat revascularisation rates in those patients<sup>18</sup>.

The win-ratio approach addresses the limitations of softer clinical components in a composite endpoint by putting more emphasis on events with greater clinical importance. The win-ratio analysis takes into account not only the number of events, but also the

timing of the event. While there was no statistically significant difference in survival at longest follow-up in the SYNTAX study, the Kaplan-Meier curves showed a continuous higher all-cause mortality rate after PCI. In a conventional time-to-event analysis with log-rank testing, this difference is not reflected. The win-ratio analysis of the SYNTAX trial shows that the benefit of CABG

**Table 4. Unmatched win ratio approach for MACCE, according to subgroups of SYNTAX score within LM/3VD and diabetic subgroups.**

SYNTAX score	Type of analysis	Left main disease	Three-vessel disease	Diabetes	Non-diabetes
0-22	Hazard ratio	0.91 (0.56, 1.47)	1.28 (0.87, 1.90)	1.30 (0.74, 2.28)	1.03 (0.71, 1.47)
	Relative risk	0.96 (0.64, 1.44)	1.24 (0.89, 1.73)	1.26 (0.81, 1.97)	1.06 (0.77, 1.45)
	Unmatched Rw	0.93 (0.74, 1.89)	1.28 (1.02, 1.63)	1.26 (0.87, 2.58)	1.08 (0.95, 1.21)
23-32	Hazard ratio	0.94 (0.57, 1.55)	1.88 (1.29, 2.72)	1.56 (0.94, 2.64)	1.48 (1.07, 2.06)
	Relative risk	1.01 (0.67, 1.52)	1.68 (1.22, 2.31)	1.45 (0.92, 2.30)	1.38 (1.02, 1.85)
	Unmatched Rw	0.87 (0.55, 1.20)	1.83 (1.38, 2.87)	1.45 (0.80, 2.11)	1.41 (1.01, 2.11)
≥33	Hazard ratio	1.78 (1.21, 2.63)	2.02 (1.35, 3.03)	2.25 (1.51, 4.30)	1.62 (1.16, 2.25)
	Relative risk	1.57 (1.15, 2.14)	1.74 (1.24, 2.44)	2.22 (1.43, 3.45)	1.46 (1.11, 1.91)
	Unmatched Rw	1.68 (1.46, 2.27)	2.11 (1.62, 3.03)	2.70 (1.64, 2.89)	1.61 (1.20, 2.12)

Results are displayed as ratios with 95% CI between brackets. Rw: win ratio

over PCI is evident in terms of both lower MACCE and lower all-cause mortality rates ( $p=0.021$ ). Using the win ratio, this is the first time that a difference in all-cause mortality between PCI and CABG has been shown.

In patients with low SYNTAX scores, the win ratio for MACCE was not statistically different between treatment groups, but there was a considerable difference between the win-ratio and conventional analysis, suggesting that CABG may be favourable even in this subgroup of patients with a low SYNTAX score. This can be explained by the three times higher MI rates and the necessity for repeat CABG revascularisation in the PCI group. However, these findings are hypothesis-generating, and the preferred revascularisation method in the group with a low SYNTAX score remains a matter of debate that will need evaluation in future clinical trials.

### FUTURE CLINICAL TRIAL DESIGN

The win ratio proves to be an important method for analysing future randomised clinical trial data. The unmatched win ratio substantially increases statistical power, while the matched win ratio showed an even larger increase in treatment effect<sup>19,20</sup>. Using the win ratio for sample size calculations may therefore reduce the number of patients in a trial, with the obvious advantages of shorter enrolment and lower costs. Expanding the number of components and including components with a wide range of impact severity will increase event rates and reduce the sample size further. While this would be considered inappropriate for conventional analyses<sup>21</sup>, this is not an issue when applying the win ratio since events are prioritised based on their impact severity.

When using the win ratio, however, it is important to alternate between the matched and unmatched approaches. Although the matched approach is favoured, our subgroup analysis that was performed 10 times suggests that exclusion of patients from an analysis in order to produce matched pairs can create a selection bias, causing an incorrect estimate of the true treatment effect on the outcome of interest. Therefore, it is recommended to use the unmatched approach when matching two treatment groups for which a substantial number of patients (arbitrarily >10%) should be excluded for matching.

### Study limitations

Applying the win ratio has some limitations. First, there is no clear consensus on ranking the severity of the events in MACCE. In this study, we used the weighting scheme as proposed by Tong and co-authors<sup>17</sup>. In addition, a repeat CABG was rated as having more impact than repeat PCI. One may also argue that repeat CABG may have more impact than MI, due to its invasiveness and potential complications. Secondly, even within a single event, there are different degrees of severity, such as major MI with subsequent left ventricular dysfunction versus MI in the smaller branches of the coronary arteries, with less impact on a patient's quality of life and prognosis. Likewise, a severe MI may have more consequences than a minor stroke. Future validation and verifications of the win ratio should be conducted before it becomes widely

used in clinical trials. Moreover, the use of TAXUS stents in clinical practice was superseded by second-generation DES, which has been shown to improve long-term outcomes significantly. Therefore, the presented analyses must be considered observational and "hypothesis-generating".

It should be acknowledged that a hazard ratio (HR), relative risk (RR) and the win ratio (Rw) are different outcome measures, and it is therefore unclear whether they can be compared directly.

### Conclusions

The win ratio is a new method to analyse composite endpoints within clinical trials. It can be used effectively and provides a stronger estimate of a treatment effect than conventional analyses. Furthermore, it can easily be extended to analyse composite endpoints with multiple components and with a wider range of impact severity, while maintaining integrity. Based on these advantages, future trials adopting this approach can expect similar statistical power with smaller sample sizes, and lower trial costs.

In case of PCI versus CABG in the SYNTAX trial, this re-analysis bolstered the results of the conventional analysis and strengthened the finding in favour of CABG treatment for patients with complex coronary disease. This provides evidence that hard clinical outcomes in particular (e.g., death and MI) after CABG are less frequent as compared with PCI. It is important to emphasise that this analysis does not undermine the findings of the original conventional analysis based on a traditional pre-specified design; it does, however, more appropriately estimate the treatment effect of PCI versus CABG by prioritising hard clinical endpoints over softer endpoints.

### Impact on daily practice

This study demonstrates that the win ratio approach can be simply and efficiently used to analyse composite outcomes in clinical trials that have combined several components with different clinical importance into a single measure. The obtained results provide a valuable framework to clinicians for meaningful outcome analysis following percutaneous coronary intervention and coronary artery bypass grafting. The win ratio has several advantages over conventional analyses and may be pre-specified in future trial designs.

### Guest Editor

This paper was guest edited by William Wijns, MD, PhD; The Lambe Institute for Translational Medicine and Curam, National University of Ireland, Galway and Saolta University Healthcare Group, Galway, Ireland.

### Funding

The SYNTAX trial was supported by funds from Boston Scientific Corporation.

## Conflict of interest statement

The authors have no conflicts of interest to declare. The Guest Editor has no conflicts of interest to declare.

## References

1. Deb S, Wijeyesundera HC, Ko DT, Tsubota H, Hill S, Fremes SE. Coronary artery bypass graft surgery vs percutaneous interventions in coronary revascularization: a systematic review. *JAMA*. 2013;310:2086-95.
2. Lim E, Brown A, Helmy A, Mussa S, Altman DG. Composite outcomes in cardiovascular research: a survey of randomized trials. *Ann Intern Med*. 2008;149:612-7.
3. Ong AT, Serruys PW, Mohr FW, Morice MC, Kappetein AP, Holmes DR Jr, Mack MJ, van den Brand M, Morel MA, van Es GA, Kleijne J, Koglin J, Russell ME. The SYnergy between percutaneous coronary intervention with TAXus and cardiac surgery (SYNTAX) study: design, rationale, and run-in phase. *Am Heart J*. 2006;151:1194-204.
4. Serruys PW, Morice MC, Kappetein AP, Colombo A, Holmes DR, Mack MJ, Stähle E, Feldman TE, van den Brand M, Bass EJ, Van Dyck N, Leadley K, Dawkins KD, Mohr FW; SYNTAX Investigators. Percutaneous coronary intervention versus coronary-artery bypass grafting for severe coronary artery disease. *N Engl J Med*. 2009;360:961-72.
5. Mantovani V, Lepore V, Mira A, Berglin E. Non-inferiority randomized trials, an issue between science and ethics: the case of the SYNTAX study. *Scand Cardiovasc J*. 2010;44:321-4.
6. Cordoba G, Schwartz L, Woloshin S, Bae H, Gotzsche PC. Definition, reporting, and interpretation of composite outcomes in clinical trials: systematic review. *BMJ*. 2010;341:c3920.
7. 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;33:176-82.
8. Head SJ, Davierwala PM, Serruys PW, Redwood SR, Colombo A, Mack MJ, Morice MC, Holmes DR Jr, Feldman TE, Stähle E, Underwood P, Dawkins KD, Kappetein AP, Mohr FW. Coronary artery bypass grafting vs. percutaneous coronary intervention for patients with three-vessel disease: final five-year follow-up of the SYNTAX trial. *Eur Heart J*. 2014;35:2821-30.
9. Kappetein AP, Feldman TE, Mack MJ, Morice MC, Holmes DR, Stähle E, Dawkins KD, Mohr FW, Serruys PW, Colombo A. Comparison of coronary bypass surgery with drug-eluting stenting for the treatment of left main and/or three-vessel disease: 3-year follow-up of the SYNTAX trial. *Eur Heart J*. 2011;32:2125-34.
10. Laarman GJ, Suttrop MJ, Dirksen MT, van Heerebeek L, Kiemeneij F, Slagboom T, van der Wieken LR, Tijssen JG, Rensing BJ, Patterson M. Paclitaxel-eluting versus uncoated stents in primary percutaneous coronary intervention. *N Engl J Med*. 2006;355:1105-13.
11. Finkelstein DM, Schoenfeld DA. Combining mortality and longitudinal measures in clinical trials. *Stat Med*. 1999;18:1341-54.
12. Le Henanff A, Giraudeau B, Baron G, Ravaud P. Quality of reporting of noninferiority and equivalence randomized trials. *JAMA*. 2006;295:1147-51.
13. Freemantle N, Calvert M, Wood J, Eastaugh J, Griffin C. Composite outcomes in randomized trials: greater precision but with greater uncertainty? *JAMA*. 2003;289:2554-9.
14. Ferreira-Gonzalez I, Busse JW, Heels-Ansdell D, Montori VM, Akl EA, Bryant DM, Alonso-Coello P, Alonso J, Worster A, Upadhye S, Jaeschke R, Schünemann HJ, Permyer-Miralda G, Pacheco-Huergo V, Domingo-Salvany A, Wu P, Mills EJ, Guyatt GH. Problems with use of composite end points in cardiovascular trials: systematic review of randomised controlled trials. *BMJ*. 2007;334:786.
15. Kip KE, Hollabaugh K, Marroquin OC, Williams DO. The problem with composite end points in cardiovascular studies: the story of major adverse cardiac events and percutaneous coronary intervention. *J Am Coll Cardiol*. 2008;51:701-7.
16. Cutlip DE, Windecker S, Mehran R, Boam A, Cohen DJ, van Es GA, Steg PG, Morel MA, Mauri L, Vranckx P, McFadden E, Lansky A, Hamon M, Krucoff MW, Serruys PW; Academic Research Consortium. Clinical end points in coronary stent trials: a case for standardized definitions. *Circulation*. 2007;115:2344-51.
17. Tong BC, Huber JC, Ascheim DD, Puskas JD, Ferguson TB Jr, Blackstone EH, Smith PK. Weighting composite endpoints in clinical trials: essential evidence for the heart team. *Ann Thorac Surg*. 2012;94:1908-13.
18. Daemen J, Boersma E, Flather M, Booth J, Stables R, Rodriguez A, Rodriguez-Granillo G, Hueb WA, Lemos PA, Serruys PW. Long-term safety and efficacy of percutaneous coronary intervention with stenting and coronary artery bypass surgery for multivessel coronary artery disease: a meta-analysis with 5-year patient-level data from the ARTS, ERACI-II, MASS-II, and SoS trials. *Circulation*. 2008;118:1146-54.
19. Moye LA, Davis BR, Hawkins CM. Analysis of a clinical trial involving a combined mortality and adherence dependent interval censored endpoint. *Stat Med*. 1992;11:1705-17.
20. McMahon RP, Harrell FE. Power calculation for clinical trials when the outcome is a composite ranking of survival and a non-fatal outcome. *Control Clin Trials*. 2000;21:305-12.
21. Head SJ, Kaul S, Bogers AJ, Kappetein AP. Non-inferiority study design: lessons to be learned from cardiovascular trials. *Eur Heart J*. 2012;33:1318-24.