Bias Attributable to Composite Outcome

Fredi Alexander Diaz-Quijano, MD, Ph.D.[1]

[1] Department of Epidemiology - Laboratório de Inferência Causal em Epidemiologia (LINCE-USP). School of Public Health, University of São Paulo. São Paulo, Brazil. E-mail: frediazq@usp.br; Twitter: @DiazQuijanoFA

Abstract

Background: Little guidance is available on how composite outcomes should be interpreted, especially in situations of varied direction in the association across the event subtypes. I proposed an index to evaluate the bias attributable to composite outcomes (BACO) and applied it in recently published clinical trials.

Methods: I defined the BACO index as the ratio between logarithms of the association measures of both a composite outcome and its most relevant component (e.g., any-cause mortality). By using the non-linear combination of parameters, based on the delta method, I calculated the confidence intervals and performed Wald-type tests for the null hypotheses (BACO index = 1). I applied this method in systematically selected clinical trials, and in two other preselected trials which I considered "positive controls". These last trials have been recognized as examples of primary composite outcomes that were disregarded because of inconsistency with the treatment effect on mortality.

Results: BACO index values different from one were classified according to whether the use of composite outcomes overestimated (BACO index >1), underestimated (BACO index between zero and <1), or inverted (BACO index <0) the association between exposure and prognosis. In three of 23 clinical trials and the two positive controls, the BACO indices were significantly lower than one (using $p < 0.005$ as a preset cutoff).

Conclusion: Based on the BACO index testing, researchers could predefined rules to make impartial decisions about maintaining a composite outcome as the primary endpoint or to state cautions regarding its interpretation.

Keywords: Composite outcomes, Mortality, Bias, Clinical trials, Delta method.

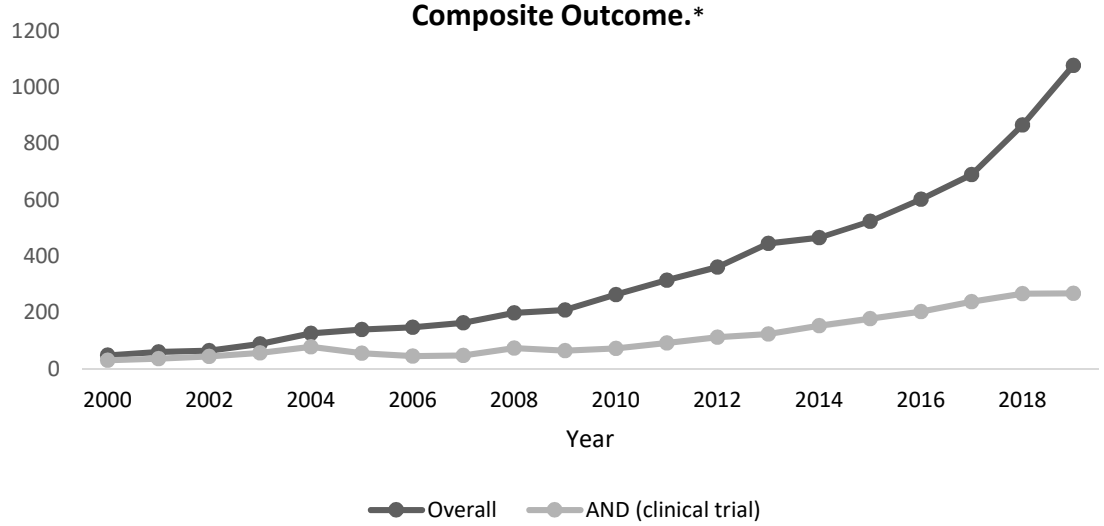
Key Messages

- Discrepancies between the effects on composite outcomes and those on their most critical components make the interpretation of research results a challenge.
- An index based on the ratio of association measures can be used to evaluate the correspondence between the composite outcome and its most critical component.
- This index could help to preset rules to make decisions for interpretation of clinical studies.

Introduction

Composite outcomes, particularly those that define an event when at least one of a group of component endpoint occurs, are increasingly used in clinical research.[1–3] In the last two decades, publications using terms such as "composite outcome" or "composite endpoint," increased progressively in PubMed, going from less than one hundred per year, before 2004, to more than one thousand in 2019 (Figure 1). About a third of these publications are associated with the terms "clinical" and "trial". These figures highlight the importance of this kind of outcome to evaluate interventions in clinical practice.

Figure 1. Annual counts of publications using terms related to Composite Outcome.*



* Output in Pubmed using the terms: "composite outcome" OR "composite endpoint" OR "composite end-point".

The use of composite outcomes can have different methodological purposes.[3] When several potentially eligible endpoints exist, a composite outcome can avoid the need to choose a simple one and prevent problems associated with multiple comparisons.[1] Moreover, it can be a way of dealing with competitive risks.[4] In other cases, a composite outcome can deliberately integrate events of a different nature that are not expected to be strongly related to each other. For example, it may simultaneously incorporate indicators of effectiveness and safety of an intervention.[5]

In addition, composite outcomes are often used to represent severity more broadly than with a single event (e.g., death). In this way, the composite outcomes facilitate a higher number of events for analysis.[1,2] This may increase the power of the study, reduce its costs, and provide a faster response to a research question. However, the results of these studies are often misinterpreted as the effect of exposure on each of the elements of the composite outcome.[2,6] This is problematic when the overall effect on the composite outcome does not follow the same direction as on its most critical components.

In some cases, intervention effects on primary composite outcomes have been ruled out because they differ greatly from those on mortality.[7–9] This may occur because the composite outcome includes events generated by different mechanisms. For example, some components may be directly related to severity (heart attacks, death), while others may be more influenced by medical decisions and resource availability (hospitalization, catheterization). In this way, some component endpoints can introduce bias affecting the estimation of the overall effect.

Despite its importance for the conclusions of the clinical studies, little guidance is available on how composite outcomes should be interpreted, especially in situations of varied direction in the association across the event subtypes.[5] Moreover, there is a lack of statistical tools supporting the decision to accept or rule out the use of the composite outcome. In this paper, I proposed an index to evaluate the bias attributable to the composite outcome (BACO). Moreover, I applied this BACO index in a group of clinical trials recently published in major medical journals.

Methods

BACO Index

I proposed to compare the association measure of a composite outcome with that obtained with a component endpoint that indisputably represents the study target. Here, I chose the any-cause mortality as the indisputable target.[10] To describe the magnitude and sense of the association of an intervention (or exposure) with both the composite outcome and death, I used the natural logarithm (Ln) of ratio-based measures (φ). Consequently, I defined the BACO index as follows:

$$BACO\ index = \frac{Ln(\varphi_c)}{Ln(\varphi_d)}$$

Where φ_c and φ_d are the association measures for a composite outcome and death, respectively. Depending on the study design and outcome characteristics, φ could correspond to the relative risk (RR), hazard ratio, incidence rate ratio, or so on.

A BACO index equal to one indicates that there is no bias attributable to the use of a composite outcome, taking the association measure for mortality as the reference. In another way, a BACO index higher than one would indicate that the association with the composite outcome is stronger than that with death. A value between zero but less than one would suggest that the association with the composite outcome is biased towards nullity. On the other hand, a negative value would result when the bias leads to an inversion of the association.

These interpretations can be applied regardless of the reference group in the comparative study. That is because if the comparison groups were inverted, the signs of both the numerator and denominator would also be inverted.

Simulation

To illustrate different results of the BACO index, I simulated a comparative study in which a group of a thousand people was exposed to an experimental intervention and presented mortality of 4.8% during the follow-up. This group was compared with a reference group integrating by another thousand people, which showed an 8% mortality. Consequently, the RR for death was 0.6 (4.8%/8%), or, in other words, the intervention had an efficacy of 40% for reducing mortality.

Now consider that researchers have four probable composite outcomes that included within their definition the occurrence of either death or other specific events. Among these composite outcomes, the first one maintained the proportionality in the risks for both groups (RR = 0.6). The second outcome led to an overestimation of the effect of the intervention with a RR of 0.3 (efficacy of 70%). On the contrary, the third outcome led to an underestimation of the effect with a RR of 0.8 (efficacy of 20%). Finally, the fourth composite outcome led to an inversion of the association measure, suggesting that the intervention duplicates the risk of the event (Table 1).

Table 1. Distribution of simulated outcomes in two hypothetic groups.

Outcome	Intervention (n=1000)	Reference (n=1000)	RR
Death	48	80	0.6
Composite outcome 1 (unbiased)	120	200	0.6
Composite outcome 2 (overestimated effect)	60	200	0.3
Composite outcome 3 (underestimated effect)	160	200	0.8
Composite outcome 4 (inverted effect)	320	160	2.0

Based on these assumptions, I elaborated a base by simulating individual data. For each of the outcomes, I calculated the natural logarithm of the RR using the Poisson regression, recognized as an alternative for calculating this measure to avoid problems of convergence of the Log-Binomial regression.[11] Using the non-linear combination of parameters, I estimated the BACO index as the ratio between the Poisson regression coefficient of the composite outcome and that of mortality.

Since each coefficient was calculated in a separate model, I stored and combined the estimates and the (co)variance matrices by using the seemingly unrelated estimation. This tool combines the estimation results into one parameter vector and simultaneous (co)variance matrix of the sandwich/robust type, which is appropriate even if the estimates were obtained on the same or overlapping data.[12]

Figure 2: BACO index for four simulated composite outcomes

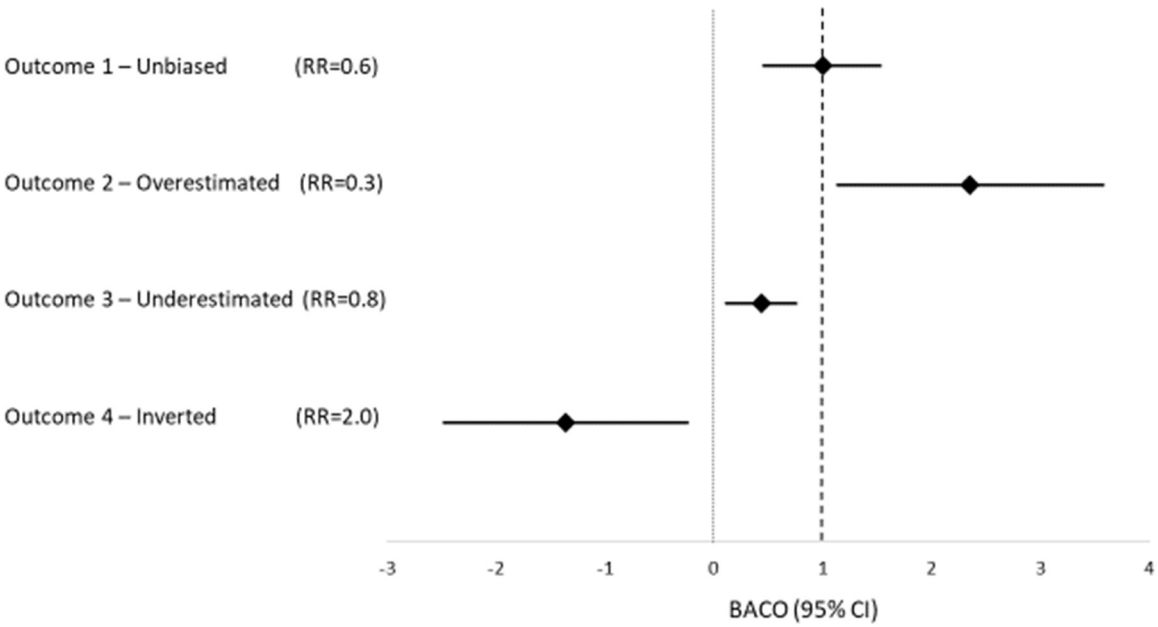


Figure 2 illustrates the results of this simulation. Specifically, for the first composite outcome (unbiased), the BACO index was equal to one (95% CI: 0.46 to 1.54). For the second composite outcome, which overestimated the effect of the intervention on prognosis, the BACO index was 2.36 (95% CI: 1.14 to 3.58). The third outcome, which underestimated the effect, presented a BACO index of 0.44 (95% CI: 0.11 to 0.76). For the last simulated composite outcome, which had an inverted association measure, the BACO index was negative: -1.36 (95% CI: -2.48 to -0.24).

Application

Although the BACO index could be applied for observational studies, in this work, I restricted it to randomized clinical trials. Therefore, I obviated the need for adjustments to control confusion. I conducted a review of 2-groups, parallel-design clinical trials published in journal groups of JAMA, NEJM, and Lancet in 2019 (updated on 07/01/2020).

In the PubMed database, I used the following word combination: composite primary (endpoint OR outcome OR ("end-point")) (mortality OR death) (randomised OR randomized) (trial) (JAMA OR NEJM OR Lancet). After, I selected those studies whose primary outcome was a composite, binary, and included all-cause mortality within its components. Secondary subgroup analyses and studies with five or fewer fatal events were excluded. A study was also excluded whose outcome results were mainly based on imputations because it presented substantial losses during the follow-up.[13]

I reviewed each article and built a database to reproduce individual data concerning the variables of intervention, composite outcome, and death. When necessary, I contacted the corresponding author to ask for data not provided in the article.

Data Analysis.

For each of the trials, regardless of the association measure preferred by the authors, I calculated the RR for both the composite outcome (RR_c) and death from any cause (RR_d), using the corresponding intervention as the independent variable. Following the same procedures described for the simulation, I used the non-linear combination of parameters to estimate the BACO index as the ratio of the Poisson regression coefficients. Moreover, I performed Wald-type tests, based on the delta method, for the null hypothesis that the BACO index is equal to one. These analyses were conducted using Stata software (version 15.0, Stata Corp LP, College Station, TX, USA), and the commands employed were described in the appendix.

This work intended to reproduce what the application of the BACO index would have been independently in each of the studies. Therefore, I did not consider the number of trials to adjust the level of significance. However, I preset 0.005 as a level to define a statistically significant BACO.[14] This value was chosen and not a higher one (e.g., 0.05), assuming that the composite outcomes were purposefully defined to be consistent with mortality, thus expecting a low pre-test probability that the BACO index is different than one. However, I used the term "suggestive" for p values between 0.005 and 0.05.[15]

I applied these analysis procedures in the selected studies from 2019 and two other clinical trials, the studies CAPRICORN and EXPEDITION.[16,17] CAPRICORN investigated carvedilol in patients with left-ventricular dysfunction after acute myocardial infarction;[16] and EXPEDITION evaluated intravenous caripode in high-risk coronary artery bypass graft surgery patients.[17] I included these last two studies as "positive controls" because they have been described as examples of primary composite outcomes that were disregarded for not being consistent with the treatment effect on mortality.[8,9]

Results

From a total of 82 references in Pubmed, I selected 23 clinical trials. Most of them were about cardiovascular diseases. Besides mortality, the composite outcomes integrated diverse components, often including cardio-cerebrovascular events, such as myocardial infarction or stroke, and other related to the use of health services, such as hospital admission and vascular interventions. The sample sizes of these trials ranged from 240 to 536233. While the number of composite outcomes ranged from 31 to 4067, the number of deaths ranged from 6 to 1140 (Table 2).

In six of the studies,[18–23] the RRs of the composite outcomes were further from the null value than the corresponding RRs of death (Table 3). Consequently, these studies had a BACO index greater (although not statistically different) than 1. One study had the same association measure for both the composite outcome and any-cause death (BACO index = 1).[24]

BACO index was lower than one in the other 16 studies,[25,26,35–40,27–34] and the BACO was statistically significant in three of them. These works included the study by Yasuda et al. in which a monotherapy with rivaroxaban (a non–vitamin K antagonist oral anticoagulant) was compared to a combination therapy with rivaroxaban plus a single antiplatelet agent.[28] The apparent effect of the monotherapy on the composite outcome (RR_c: 0.74; 95% CI: 0.57 to 0.96) was lower than the effect on mortality (RR_d: 0.56; 95% CI: 0.39 to 0.82), in patients with atrial fibrillation and coronary disease. Thus, the BACO index was 0.53 (95% CI: 0.21 to 0.86; p=0.0048).

The second study with statistically significant BACO was that by Lanz et al., in which they compared a self-expanding versus a balloon-expandable bioprosthesis for transcatheter aortic valve replacement in patients with symptomatic severe aortic stenosis.[33] In this trial, the self-expanding bioprosthesis group presented an incidence of the composite outcome about 43% higher than observed in the balloon-expandable group (RR_c: 1.43; 95% CI: 1.06 – 1.92). However, mortality in the first group was about three-times that of the second one (RR_d: 2.96; 95% CI: 0.81 to 10.8). In this case, the BACO index was 0.33 (95% CI: -0.1 to 0.76; p=0.002).

In the other study, Onland et al. evaluated the effect of systemic hydrocortisone compared with placebo on a composite outcome of death or bronchopulmonary dysplasia in very preterm infants.[34] The intervention was not associated with a significant change in the composite outcome incidence (RR_c= 0.95; 95%CI: 0.84 to 1.08). However, the hydrocortisone group exhibited lower mortality compared to placebo (RR_d: 0.65; 95% CI: 0.42 to 0.99). BACO index in this study was 0.11 (95% CI: -0.17 to 0.39; p<0.001).

Three other works were suggestive of a significant BACO in which the RR of the composite outcome was towards nullity compared to the RR of mortality.[29,30,35] These works included the clinical trials by Schuetz et. al. (BACO index: 0.52; 95%CI: 0.07 to 0.98; p=0.04); Stone et al. (BACO index: 0.52; 95% CI: 0.07 to 0.98, p=0.03); and Nagel et al., which exhibited a negative point estimate of the BACO index (-0.06; 95% CI: -1.1 to 0.98, p=0.04).

Regarding the positive controls, the CAPRICORN study showed that carvedilol, compared with placebo, was not significantly associated with the composite outcome (RR = 0.94; 95% CI: 0.84 to 1.06) but was associated with a 22% reduction in mortality (BACO index: 0.24; 95% CI: -0.15 to 0.64). On the other hand, in the EXPEDITION study, the use of cariporide was associated with an 18% lower incidence of the composite outcome, but with a 53% higher mortality, compared with placebo (BACO index: -0.48; 95% CI: -1.03 to 0.08). The BACO indices of the two studies were significantly lower than one (p <0.001).

Table 2. General description of the study population and outcomes of clinical trials selected.

First Author	Characteristics of the study population	Composite elements other than all-cause mortality	n	Composite /Deaths
Holm NR	Left main coronary artery disease requiring revascularization.	Non-procedural myocardial infarction, repeat revascularization, or stroke.	1201	275/104
Katheria A	Preterm infants (born at 23-31 week gestation).	Severe intraventricular hemorrhage.	474	49/32
Schüpke S	ACS for whom invasive evaluation was planned.	Myocardial infarction or stroke at one year.	4018	321/163
Brott TG	Moderate or severe atherosclerotic symptomatic stenosis at the carotid bifurcation.	Stroke within 120 days after randomization or subsequent ipsilateral stroke up to 10 years after randomization.	4754	447/39
Tomaniak M	Acute coronary syndromes (ACS) beyond one month after a percutaneous coronary intervention (PCI).	New Q-wave myocardial infarction.	7487	130/89
Wu Y	ACS in resource-constrained hospitals.	Reinfarction/myocardial infarction or nonfatal stroke.	29346	1214/1140
Hahn JY	Patients undergoing PCI.	Myocardial infarction, or stroke, at 12 months after the index procedure.	2993	78/39
Packer DL	Symptomatic AF, aged ≥65 years; or younger than 65 years with one or more risk factors for stroke.	Disabling stroke, serious bleeding, or cardiac arrest.	2204	190/125
Macdougall IC	Adults undergoing maintenance hemodialysis.	Nonfatal myocardial infarction, nonfatal stroke, or hospitalization for heart failure.	2141	658/515
Mack MJ	Severe aortic stenosis and low surgical risk.	Stroke, or rehospitalization at one year.	1000	110/16
Yasuda S	Aged ≥75 years, diagnosis of atrial fibrillation (AF) and stable coronary artery disease (patients with AF who had undergone PCI or coronary-artery bypass grafting (CABG) >1 year earlier or who had angiographically confirmed coronary artery disease not requiring revascularization).	Stroke, systemic embolism, myocardial infarction, unstable angina requiring revascularization.	2215	210/114
Schuetz P	Nutritional risk and with an expected length of hospital stay of >4 days.	Admission to intensive care, non-elective hospital readmission, major complications, or decline in functional status at 30 days	2088	504/173
Stone GW	Left main coronary artery disease of low or intermediate anatomical complexity.	Stroke, or myocardial infarction.	1905	379/208
Zenati MA	Patients undergoing CABG.	Nonfatal myocardial infarction, or repeat revascularization.	1150	169/83

First Author	Characteristics of the study population	Composite elements other than all-cause mortality	n	Composite /Deaths
Iversen K	Adults in stable condition who had endocarditis on the left side of the heart.	Unplanned cardiac surgery, embolic events, or relapse of bacteremia with the primary pathogen, from randomization until six months after antibiotic treatment was completed.	400	42/20
Lanz J	Aged ≥75 years, undergoing transfemoral Transcatheter aortic valve replacement (TAVR) for treatment of symptomatic severe aortic stenosis, and who were deemed to be at increased surgical risk.	Any stroke, life-threatening or disabling bleeding, major vascular complications, coronary artery obstruction requiring intervention, acute kidney injury (stage 2 or 3), rehospitalization for valve-related symptoms or congestive heart failure, valve-related dysfunction requiring repeat procedure, moderate or severe prosthetic valve regurgitation, or prosthetic valve stenosis within 30 days of the procedure.	739	147/12
Onland W	Preterm infants (gestational age <30 weeks or birth weight <1250 g) who were ventilator dependent.	Bronchopulmonary dysplasia (BPD) assessed at 36 weeks' postmenstrual age.	372	268/73
Nagel E	Typical angina and either ≥2 cardiovascular risk factors or a positive exercise treadmill test.	Nonfatal myocardial infarction, or target-vessel revascularization within one year.	851	31/6
Ho KM	Severely injured patients who had a contraindication to anticoagulant agents.	Symptomatic pulmonary embolism at 90 days after enrollment.	240	34/27
Van Spall HGC	Adult patients hospitalized for heart failure (HF).	All-cause readmission, or emergency department (ED) visit at three months.	2494	1243/247
Kozhuharov N	Patients hospitalized for acute heart failure.	Rehospitalization for AHF at 180 days.	788	228/116
Nguyen HQ	40 years or older who had any acute care use related to chronic obstructive pulmonary disease, in the previous 12 months.	All-cause hospitalizations, observation stays, emergency department visits.	2707	1747/234
Vousden N	Users of maternity care.	Eclampsia, or emergency hysterectomy.	536233	4067/998
Positive controls:				
Dargie HJ (CAPRICORN)	A proven acute myocardial infarction and a left-ventricular ejection fraction of </=40%.	Hospital admission for cardiovascular problems.	1959	705/267
Mentzer R. Jr (EXPEDITION)	High-risk CABG surgery patients.	Myocardial infarction (MI).	5761	1064/106

Table 3. Relative risks of composite and death and BACO index in clinical trials.

First author	RR_c (95% CI)	RR_d (95% CI)	BACO index (95% CI)	p value*
Holm NR	1.51 (1.22 to 1.87)	1.09 (0.75 to 1.57)	4.85 (-14.7 to 24.4)	0.70
Katheria A	1.46 (0.85 to 2.51)	1.14 (0.58 to 2.23)	2.84 (-8.44 to 14.1)	0.75
Schüpke S	1.34 (1.08 to 1.66)	1.23 (0.91 to 1.66)	1.41 (-0.12 to 2.95)	0.60
Brott TG	0.71 (0.59 to 0.85)	0.78 (0.42 to 1.47)	1.41 (-2.08 to 4.9)	0.82
Tomaniak M	0.73 (0.52 to 1.03)	0.74 (0.49 to 1.13)	1.05 (0.22 to 1.89)	0.90
Wu Y	0.87 (0.78 to 0.97)	0.87 (0.78 to 0.97)	1 (0.8 to 1.21)	0.99
Hahn JY	1.17 (0.75 to 1.81)	1.17 (0.63 to 2.19)	1 (-1.85 to 3.85)	1
Packer DL	0.87 (0.66 to 1.14)	0.86 (0.61 to 1.2)	0.89 (-0.31 to 2.08)	0.85
Macdougall IC	0.91 (0.8 to 1.03)	0.88 (0.75 to 1.02)	0.74 (0.22 to 1.26)	0.32
Mack MJ	0.61 (0.04 to 0.88)	0.45 (0.16 to 1.28)	0.62 (-0.16 to 1.4)	0.34
Yasuda S	0.74 (0.57 to 0.96)	0.56 (0.39 to 0.82)	0.53 (0.21 to 0.86)	0.0048
Schuetz P	0.84 (0.72 to 0.98)	0.72 (0.54 to 0.96)	0.52 (0.07 to 0.98)	0.04
Stone GW	1.16 (0.97 to 1.4)	1.35 (1.04 to 1.75)	0.51 (0.07 to 0.94)	0.03
Zenati MA	1.12 (0.84 to 1.48)	1.25 (0.82 to 1.89)	0.5 (-0.45 to 1.44)	0.30
Iversen K	1.35 (0.75 to 2.4)	1.88 (0.76 to 4.6)	0.47 (-0.24 to 1.18)	0.15
Lanz J	1.43 (1.06 to 1.92)	2.96 (0.81 to 10.8)	0.33 (-0.1 to 0.76)	0.002
Onland W	0.95 (0.84 to 1.08)	0.65 (0.42 to 0.99)	0.11 (-0.17 to 0.39)	<0.001
Nagel E	0.96 (0.48 to 1.91)	2.04 (0.38 to 11.1)	-0.06 (-1.1 to 0.98)	0.04
Ho KM	0.97 (0.52 to 1.8)	1.41 (0.68 to 2.9)	-0.1 (-2.11 to 1.91)	0.28
Van Spall HGC	0.98 (0.91 to 1.06)	1.03 (0.81 to 1.3)	-0.63 (-7.64 to 6.38)	0.65
Kozhuharov N	1.1 (0.88 to 1.37)	0.94 (0.67 to 1.31)	-1.48 (-12 to 9.05)	0.64
Nguyen HQ	1.02 (0.96 to 1.07)	0.99 (0.78 to 1.27)	-2.27 (-88.3 to 83.7)	0.94
Vousden N	0.92 (0.86 to 0.97)	1.04 (0.92 to 1.18)	-2.35 (-11.2 to 6.51)	0.46
Positive Controls:				
Dargie HJ (CAPRICORN)	0.94 (0.84 to 1.06)	0.78 (0.62 to 0.97)	0.24 (-0.15 to 0.64)	<0.001
Mentzer R. Jr (EXPEDITION)	0.82 (0.73 to 0.91)	1.53 (1.04 to 2.26)	-0.48 (-1.03 to 0.08)	<0.001

* p-value for the hypothesis of the BACO index is equal to one.

RR_c: Relative Risk for the composite outcome; RR_d: Relative Risk for any-cause death; BACO index: $Ln\ (RR_c)/Ln\ (RR_d)$

Discussion

Composite outcomes can prove challenging for the interpretation of results.[1,7] Differential effects on their less critical but more frequent components may result in a misleading impression about the impact of a treatment.[2,41] Therefore, it has been recommended that if there is a great variation between the effects on the components, the composite outcome should be abandoned.[7]

However, assessing this is difficult, considering the asymmetric distribution of association measures and random variations. Some authors have evaluated the differences between the associations of both the composite outcome and mortality based on disagreement in statistical significance.[1] This type of comparison induces a bias because of the fatal outcome is a sub-element of the composite and will always have fewer events. Therefore, there will be less power to evaluate an association with mortality.

The BACO index summarizes the relationship between the associations of the composite outcome and its most critical endpoint. Being based on the logarithms, the comparison is more consistent with the association measure distributions. On the other hand, the integration into a single index allows a unique statistical test for the null hypothesis.

Other authors stated they had planned to calculate a ratio between the efficacy for the composite outcome and that for mortality.[6] However, they considered that it would be problematic because the observations were not independent, as the death contributes to the composite. In that sense, the methodology proposed to obtain the BACO index solves this problem by considering overlapping observations. The non-linear combination of parameters using the delta method is a versatile and powerful tool to calculate confidence intervals and perform hypothesis tests based on the ratio of coefficients.[12,42,43]

Despite this, a limitation for an index based on the ratio of efficacy measures is that denominators close to zero lead to unstable or seemingly inflated results.[6,44] Hence, a BACO index should not be computed together with another whose reference effect is different. In other words, comparisons between BACO indices only make sense to contrast two or more composite outcomes when they have the same reference value (e.g., of the RR for mortality). In other circumstances, it is prudent to interpret the BACO indices only by classifying them into the three categories of overestimation, underestimation, and effect inversion; and considering the null hypothesis test.

In practice, the BACO index proved to be a simple measure to validate the composite outcome in clinical trials. Three of the analyzed studies had an index significantly lower than one, suggesting that the composite outcome underestimated the association between the intervention and the prognosis. Additionally, three other studies were suggestive of a similar trend, i.e., the composite outcomes seemed to dilute the associations that were stronger for mortality.

When this analysis was applied to the positive controls, the BACO index was significantly different from one. These studies have been well recognized in the literature as examples of bias associated with the composite outcome. After many discussions about the results of the CAPRICORN and EXPEDITION trials, the composite outcome was disregard and the conclusions were based on the effects on all-cause mortality.[8,9] This would be consistent with the results of the BACO index.

In the absence of clear guidelines, the BACO index could be a statistical tool contributing to the interpretation of composite outcomes. The level of significance could be adjusted depending on the sensitivity to identify a BACO that the authors want. I chose a level of 0.005, but the researchers could predefine a more sensitive cut-off point (e.g., 0.05 or 0.10), especially

if they were afraid that the outcome might have elements that introduce a bias in the representation of the disease severity.

This work is based on the expectation that the composite outcome must go in the same direction as its most relevant component. This is important when what is sought is that the composite outcome offers more events to increase power. In other cases, it may not be necessary for the composite outcome to have a similar association magnitude as mortality. For example, when an intervention is expected to improve the quality of life without affecting total survival. In these cases, a composite outcome could be an option to deal with the problem of competitive risks.[4] However, a significant BACO could lead to reconsider the necessity of using a composite outcome or to adequately disaggregate the estimations for the component endpoints in both results and conclusions.

The BACO index calculation could be incorporated into the analysis plan of clinical studies. Thus, based on a predefined rule, researchers could make impartial decisions regarding maintaining or replacing a composite outcome as the primary endpoint. Even if the researchers decide to keep their conclusions based on the composite outcome, a significant BACO should lead to the caution that the association of the composite is stronger, weaker, or even opposite than that of its most critical component.

Acknowledgments

I thank my children, Alejandro and Ana Maria, for helping me organize the bibliographic material. I also thank Dr. Kai Arzheimer for his guidance on Stata commands.

Funding

This work had no specific funding. However, the author was granted a fellowship for research productivity from the Brazilian National Council for Scientific and Technological Development – CNPq.

Competing interests

The author has no conflict of interest related to this study.

References

1. Freemantle N, Calvert M, Wood J, Eastaugh J, Griffin C. Composite Outcomes in Randomized Trials Greater Precision But With Greater Uncertainty? *JAMA*. 2003;**289**(19):2554–2559.

2. Ferreira-González I, Busse JW, Heels-Ansdell D, et al. Problems with use of composite end points in cardiovascular trials: Systematic review of randomised controlled trials. *Br Med J*. 2007 Apr 14;**334**(7597):786–788.

3. Sankoh AJ, Li H, D'agostino RB. Use of composite endpoints in clinical trials. *Stat Med*. 2014;**33**:4709–4714.

4. Manja V, AlBashir S, Guyatt G. Criteria for use of composite end points for competing risks-a systematic survey of the literature with recommendations. *J Clin Epidemiol* [Internet]. Elsevier; 2017 Feb 1 [cited 2019 Oct 16];**82**:4–11. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27965044

5. Armstrong PW, Westerhout CM. Composite End Points in Clinical Research. *Circulation*. 2017;**135**:2299–2307.

6. Cordoba G, Schwartz L, Woloshin S, Bae H, Gøtzsche PC. Definition, reporting, and interpretation of composite outcomes in clinical trials: systematic review. *BMJ*. 2010;**341**:c3920.

7. Montori VM, Permanyer-Miralda G, Ferreira-González I, et al. Validity of composite end points in clinical trials. *Br. Med. J.* 2005. p. 594–596.

8. Pocock SJ, Stone GW. The Primary Outcome Fails — What Next? *N Engl J Med*. 2016;**375**(9):861–870.

9. Pocock SJ, Stone GW. The Primary Outcome Is Positive — Is That Good Enough? *N Engl J Med*. 2016;**375**(10):971–979.

10. Gelijns AC. Modern Methods of Clinical Investigation. Gelijns AC, editor. *Mod. Methods Clin. Investig.* Washington (DC): NATIONAL ACADEMY PRESS; 1990.

11. Barros AJD, Hirakata VN. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. *BMC Med Res Methodol* [Internet]. 2003;**3**:21. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=521200&tool=pmcentrez&rendertype=abstract

12. Stata Press. STATA BASE REFERENCE MANUAL Release 15. 15th ed. StataCorp LLC, editor. Texas: Stata Press; 2017.

13. Popma JJ, Michael Deeb G, Yakubov SJ, et al. Transcatheter aortic-valve replacement with a self-expanding valve in low-risk patients. *N Engl J Med*. 2019;**380**(18):1706–1715.

14. Benjamin DJ, Berger JO, Johannesson M, et al. Redefine statistical significance. *Nat Hum Behav* [Internet]. Nature Publishing Group; 2018 Jan 1 [cited 2018 Apr 3];**2**(1):6–10. Available from: http://www.nature.com/articles/s41562-017-0189-z

15. Benjamin DJ, Berger JO. Three Recommendations for Improving the Use of *p* -Values. *Am Stat* [Internet]. 2019 Mar 29 [cited 2020 Jan 21];**73**(sup1):186–191. Available from: https://www.tandfonline.com/doi/full/10.1080/00031305.2018.1543135

16. Dargie HJ. Effect of carvedilol on outcome after myocardial infarction in patients with left-ventricular dysfunction: The CAPRICORN randomised trial. *Lancet*. Elsevier Limited; 2001 May 5;**357**(9266):1385–1390.

17. Mentzer RM, Bartels C, Bolli R, et al. Sodium-Hydrogen Exchange Inhibition by Cariporide to Reduce the Risk of Ischemic Cardiac Events in Patients Undergoing Coronary Artery Bypass Grafting: Results of the EXPEDITION Study. *Ann Thorac Surg*. 2008 Apr;**85**(4):1261–1270.

18. Holm NR, Mäkikallio T, Lindsay MM, et al. Percutaneous coronary angioplasty versus coronary artery bypass grafting in the treatment of unprotected left main stenosis: updated 5-year outcomes from the randomised, non-inferiority NOBLE trial. *Lancet* [Internet]. 2019 [cited 2020 Jan 8];(19). Available from: https://linkinghub.elsevier.com/retrieve/pii/S0140673619329721

19. Katheria A, Reister F, Essers J, et al. Association of umbilical cord milking vs delayed umbilical cord clamping with death or severe intraventricular hemorrhage among preterm infants. *JAMA - J Am Med Assoc*. American Medical Association; 2019 Nov 19;**322**(19):1877–1886.

20. Schüpke S, Neumann F-J, Menichelli M, et al. Ticagrelor or Prasugrel in Patients with

Acute Coronary Syndromes. *N Engl J Med* [Internet]. 2019 Oct 17 [cited 2020 Jan 8];**381**(16):1524–1534. Available from: http://www.nejm.org/doi/10.1056/NEJMoa1908973

21. Brott TG, Calvet D, Howard G, et al. Long-term outcomes of stenting and endarterectomy for symptomatic carotid stenosis: a preplanned pooled analysis of individual patient data. *Lancet Neurol*. Lancet Publishing Group; 2019 Apr 1;**18**(4):348–356.

22. Tomaniak M, Chichareon P, Onuma Y, et al. Benefit and risks of aspirin in addition to ticagrelor in acute coronary syndromes: A post hoc analysis of the randomized global leaders trial. *JAMA Cardiol*. American Medical Association; 2019 Nov 1;**4**(11):1092–1101.

23. Wu Y, Li S, Patel A, et al. Effect of a Quality of Care Improvement Initiative in Patients with Acute Coronary Syndrome in Resource-Constrained Hospitals in China: A Randomized Clinical Trial. *JAMA Cardiol*. American Medical Association; 2019 May 1;**4**(5):418–427.

24. Hahn JY, Song Y Bin, Oh JH, et al. Effect of P2Y12 Inhibitor Monotherapy vs Dual Antiplatelet Therapy on Cardiovascular Events in Patients Undergoing Percutaneous Coronary Intervention: The SMART-CHOICE Randomized Clinical Trial. *JAMA - J Am Med Assoc*. American Medical Association; 2019 Jun 25;**321**(24):2428–2437.

25. Packer DL, Mark DB, Robb RA, et al. Effect of Catheter Ablation vs Antiarrhythmic Drug Therapy on Mortality, Stroke, Bleeding, and Cardiac Arrest among Patients with Atrial Fibrillation: The CABANA Randomized Clinical Trial. *JAMA - J Am Med Assoc*. 2019;**321**(13):1261–1274.

26. Macdougall IC, White C, Anker SD, et al. Intravenous Iron in Patients Undergoing Maintenance Hemodialysis. *N Engl J Med* [Internet]. 2019 Jan 31 [cited 2020 Jan 8];**380**(5):447–458. Available from: http://www.nejm.org/doi/10.1056/NEJMoa1810742

27. Mack MJ, Leon MB, Thourani VH, et al. Transcatheter Aortic-Valve Replacement with a Balloon-Expandable Valve in Low-Risk Patients. *N Engl J Med* [Internet]. 2019 May 2 [cited 2019 Oct 9];**380**(18):1695–1705. Available from: http://www.nejm.org/doi/10.1056/NEJMoa1814052

28. Yasuda S, Kaikita K, Akao M, et al. Antithrombotic Therapy for Atrial Fibrillation with Stable Coronary Disease. *N Engl J Med* [Internet]. 2019 Sep 19 [cited 2020 Jan 8];**381**(12):1103–1113. Available from: http://www.nejm.org/doi/10.1056/NEJMoa1904143

29. Schuetz P, Fehr R, Baechli V, et al. Individualised nutritional support in medical inpatients at nutritional risk: a randomised clinical trial. *Lancet (London, England)* [Internet]. 2019 Jun 8 [cited 2019 Oct 9];**393**(10188):2312–2321. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0140673618327764

30. Stone GW, Kappetein AP, Sabik JF, et al. Five-Year Outcomes after PCI or CABG for Left Main Coronary Disease. *N Engl J Med* [Internet]. 2019 Nov 7 [cited 2020 Jan 8];**381**(19):1820–1830. Available from: http://www.nejm.org/doi/10.1056/NEJMoa1909406

31. Zenati MA, Bhatt DL, Bakaeen FG, et al. Randomized Trial of Endoscopic or Open Vein-Graft Harvesting for Coronary-Artery Bypass. *N Engl J Med* [Internet]. 2019 Jan 10 [cited 2020 Jan 8];**380**(2):132–141. Available from: http://www.nejm.org/doi/10.1056/NEJMoa1812390

32. Iversen K, Ihlemann N, Gill SU, et al. Partial Oral versus Intravenous Antibiotic Treatment of Endocarditis. *N Engl J Med* [Internet]. 2019 Jan 31 [cited 2020 Jan 7];**380**(5):415–424.

Available from: http://www.nejm.org/doi/10.1056/NEJMoa1808312

33. Lanz J, Kim WK, Walther T, et al. Safety and efficacy of a self-expanding versus a balloon-expandable bioprosthesis for transcatheter aortic valve replacement in patients with symptomatic severe aortic stenosis: a randomised non-inferiority trial. *Lancet*. Lancet Publishing Group; 2019 Nov 2;**394**(10209):1619–1628.

34. Onland W, Cools F, Kroon A, et al. Effect of Hydrocortisone Therapy Initiated 7 to 14 Days after Birth on Mortality or Bronchopulmonary Dysplasia among Very Preterm Infants Receiving Mechanical Ventilation: A Randomized Clinical Trial. *JAMA - J Am Med Assoc*. American Medical Association; 2019 Jan 29;**321**(4):354–363.

35. Nagel E, Greenwood JP, McCann GP, et al. Magnetic Resonance Perfusion or Fractional Flow Reserve in Coronary Disease. *N Engl J Med* [Internet]. 2019 Jun 20 [cited 2019 Oct 9];**380**(25):2418–2428. Available from: http://www.nejm.org/doi/10.1056/NEJMoa1716734

36. Ho KM, Rao S, Honeybul S, et al. A Multicenter Trial of Vena Cava Filters in Severely Injured Patients. *N Engl J Med* [Internet]. 2019 Jul 25 [cited 2020 Jan 8];**381**(4):328–337. Available from: http://www.nejm.org/doi/10.1056/NEJMoa1806515

37. Spall HGC Van, Lee SF, Xie F, et al. Effect of Patient-Centered Transitional Care Services on Clinical Outcomes in Patients Hospitalized for Heart Failure: The PACT-HF Randomized Clinical Trial. *JAMA* [Internet]. 2019 Feb 26 [cited 2019 Oct 9];**321**(8):753–761. Available from: http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.2019.0710

38. Kozhuharov N, Goudev A, Flores D, et al. Effect of a Strategy of Comprehensive Vasodilation vs Usual Care on Mortality and Heart Failure Rehospitalization Among Patients With Acute Heart Failure: The GALACTIC Randomized Clinical Trial. *JAMA*. NLM (Medline); 2019 Dec 17;**322**(23):2292–2302.

39. Nguyen HQ, Moy ML, Liu I-LA, et al. Effect of Physical Activity Coaching on Acute Care and Survival Among Patients With Chronic Obstructive Pulmonary Disease: A Pragmatic Randomized Clinical Trial. *JAMA Netw open* [Internet]. 2019 Aug 2 [cited 2019 Oct 9];**2**(8):e199657. Available from: https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2748055

40. Vousden N, Lawley E, Nathan HL, et al. Effect of a novel vital sign device on maternal mortality and morbidity in low-resource settings: a pragmatic, stepped-wedge, cluster-randomised controlled trial. *Lancet Glob Heal*. Elsevier Ltd; 2019 Mar 1;**7**(3):e347–e356.

41. Lim E, Brown A, Helmy A, Mussa S, Altman DG. Composite outcomes in cardiovascular research: A survey of randomized trials. *Ann. Intern. Med.* American College of Physicians; 2008. p. 612–617.

42. Oehlert GW. A Note on the Delta Method. *Am Stat*. 1992;**46**(1):27–29.

43. Arzheimer K. nlcom and the Delta Method [Internet]. 2012 [cited 2020 Jan 27]. Available from: https://www.kai-arzheimer.com/delta-method-nlcom/amp/

44. Krogsbøll LT, Hróbjartsson A, Gøtzsche PC. Spontaneous improvement in randomised clinical trials: meta-analysis of three-armed trials comparing no treatment, placebo and active intervention. *BMC Med Res Methodol* [Internet]. 2009 Dec 5 [cited 2020 Jan 29];**9**(1):1. Available from: http://bmcmedresmethodol.biomedcentral.com/articles/10.1186/1471-2288-9-1

Appendix: Stata code for BACO index estimation and testing.

Suppose we have a dataset with exposure or treatment variable "t", composite outcome variable "c", and the most critical endpoint variable "d" (i.e., death). For this example, take the data of the CAPRICORN study, which compared 975 allocated to experimental intervention vs. 984 patients in the control group.[16] The composite outcome occurred in 340 and 365 patients, including 116 and 151 deaths, respectively. Consequently, the RRs were 0.94 (95% CI: 0.84 – 1.06) for the composite outcome and 0.78 (95% CI: 0.62 – 0.97) for death. Before calculating the BACO index based on the relative risks, we should store and integrate the regression coefficients as described in the code that follows:

```
poisson c t
est store coefcomp
poisson d t
est store coefdeath
suest  coefcomp coefdeath
```

In this example, the coefficients generated by the regression models were stored under the names "coefcomp" and "coefdeath". Then, by using "suest" (seemingly unrelated estimation) command, we combine the stored parameter estimates and associated (co)variance matrices. Now we can use "nlcom" as described in the code that follows, the output (presented in the box) will include the estimate of *BACO* index, its standard error, and a 95% confidence interval.

```
nlcom [coefcomp_c]_b[t]/([coefdeath_d]_b[t])
```

```
       _nl_1:  [coefcomp_c]_b[t]/([coefdeath_d]_b[t])

------------------------------------------------------------------------
     |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+------------------------------------------------------------------
_nl_1|   .2426837   .2023317     1.20   0.230    -.1538791    .6392466
------------------------------------------------------------------------
```

Please note that the previous p-value (0.23) refers to the null hypothesis of the BACO index equals to zero (therefore, it should not be considered).

To calculate the p-value for the null hypothesis of the BACO index equals one, we can use the "testnl" command:

```
testnl ([coefcomp_c]_b[t]/([coefdeath_d]_b[t]))=1
```

```
(1)  ([coefcomp_c]_b[t]/([coefdeath_d]_b[t]))=1
               chi2(1) =       14.01
           Prob > chi2 =        0.0002
```