TOOLS N TECHNIQUES

Novel approaches to composite endpoints in clinical trials

Cynthia M. Westerhout[1]*, PhD; Jeffrey A. Bakal[2], PhD, PStat

1. Canadian VIGOUR Centre at the University of Alberta, Edmonton, Alberta, Canada; 2. Patient Health Outcomes and Clinical Effectiveness Unit, Division of General Internal Medicine, University of Alberta, Edmonton, Alberta, Canada

"If you always do what you always did, you will always get what you always got."
Albert Einstein

The choice of endpoint is an important decision in the design of clinical trials. This choice has been compounded by the rapid expansion in the spectrum of endpoints, particularly in cardiovascular clinical trials. Hence, in lieu of a single event there are now multiple types of event often combined to form a single outcome as a composite endpoint. The first mega trial in acute myocardial infarction, Global Utilization of Streptokinase and Tissue plasminogen activator for Occluded coronary arteries (GUSTO-I), was designed on the primary endpoint of 30-day death[1]. In contrast, most contemporary trials are based on a composite endpoint of at least two and often four or more event types. This transformation has been precipitated by both wanting to reflect the patient experience and decreasing rates for single events, such as mortality, thereby impacting on the projected sample size needs, and, *pari passu*, challenging the operation and financing of such trials. Combining multiple types of clinically relevant events into a single endpoint is one strategy to buffer this new reality. However, as each of these individual rates declines, this has led to the inclusion of increasingly disparate composites, encompassing death and recurrent ischaemia, for example, as an equivalent outcome. It may also be argued that patient-oriented outcomes have long been absent from the evaluation of new treatments.

Despite this evolution in the choice(s) of endpoint, the approach to their analysis has not kept pace. Traditional analysis of endpoints, both single and composite, is time-to-first-event. However, this approach does not take into account multiple events in a single patient, their differential clinical impact, sequencing and time course, thereby potentially wasting valuable and clinically meaningful data. Additionally, this approach biases the analysis towards using the earlier and potentially less severe events. There are also important assumptions underlying the use and analysis of composite endpoints, which are infrequently satisfied **(Table 1)**.

Table 1. Assumptions of composite endpoints.

Assumptions
– Each component of the endpoint should be meaningful
– Each event should be equivalent in severity and
– The treatment effect should be similar for each component of the composite endpoint

Alternative methodologies were developed in the late 1980s and 1990s to handle multiple occurrences of the same event[2-4]. These were intended for the analysis of multiple events of the same type more often seen in areas such as oncology (e.g., recurrence of tumours) rather than multiple events of different types typical of cardiology trials. While these approaches deal with the multiplicity of events, the relative value of event types is not accounted for (i.e., all events are valued equally). There is recent and increasing discomfort with this status quo, and new methods are emerging in cardiovascular research[5-7]. One such approach is the weighted composite endpoint **(Figure 1)**[8-12]. In this approach, the relative severity weights for each component of the composite endpoint are first determined based on a survey of stakeholders. In our previous work, the stakeholders were experienced clinician-investigators; however, more recent efforts have sought the perspective of patients[12]. The results based on a contemporary ST-segment elevation myocardial infarction (STEMI) trial generated the following weights: death (1.0), cardiogenic shock (0.5), congestive heart failure (0.3), and reinfarction (0.2)[8]. Patients with multiple events were treated multiplicatively. Using this approach, each patient has a residual weight on each day; if a patient had no events, their weight was 1.0; if events occurred, their score was reduced according to the type and number of events, thereby providing a real-time weighted event rate. For example, a patient with cardiogenic shock on day 1 would have a residual weight of 0.5 at the end of day 1 (1-0.5=0.5) **(Figure 2A)**. If a reinfarction then occurred on day 3, the residual score would be reduced to 0.1 (0.5*0.2=0.1). In **Figure 2B**,

*Corresponding author: Canadian VIGOUR Centre, University of Alberta, 2-132 Li Ka Shing Centre for Health Research Innovation, Edmonton, Alberta, T6G 2E1, Canada. E-mail: cindy.westerhout@ualberta.ca
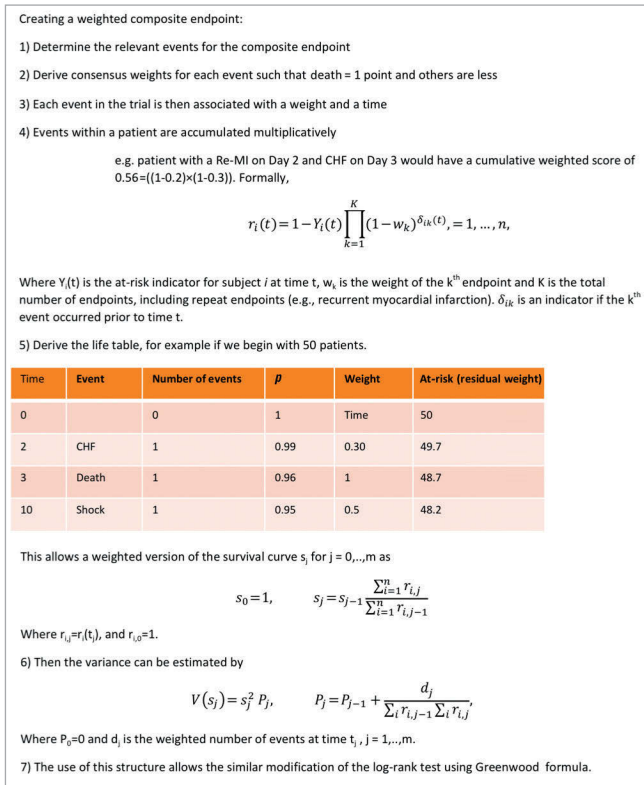
Creating a weighted composite endpoint:

1) Determine the relevant events for the composite endpoint

2) Derive consensus weights for each event such that death = 1 point and others are less

3) Each event in the trial is then associated with a weight and a time

4) Events within a patient are accumulated multiplicatively

e.g. patient with a Re-MI on Day 2 and CHF on Day 3 would have a cumulative weighted score of 0.56=((1-0.2)×(1-0.3)). Formally,

$$r_i(t) = 1 - Y_i(t) \prod_{k=1}^{K} (1 - w_k)^{\delta_{ik}(t)}, = 1, \dots, n,$$

Where $Y_i(t)$ is the at-risk indicator for subject i at time t, w_k is the weight of the kth endpoint and K is the total number of endpoints, including repeat endpoints (e.g., recurrent myocardial infarction). δ_{ik} is an indicator if the kth event occurred prior to time t.

5) Derive the life table, for example if we begin with 50 patients.

Time	Event	Number of events	p	Weight	At-risk (residual weight)
0		0	1	Time	50
2	CHF	1	0.99	0.30	49.7
3	Death	1	0.96	1	48.7
10	Shock	1	0.95	0.5	48.2

This allows a weighted version of the survival curve s_j for j = 0,..,m as

$$s_0 = 1, \qquad s_j = s_{j-1} \frac{\sum_{i=1}^{n} r_{i,j}}{\sum_{i=1}^{n} r_{i,j-1}}$$

Where $r_{i,j}=r_i(t_j)$, and $r_{i,0}=1$.

6) Then the variance can be estimated by

$$V(s_j) = s_j^2 P_j, \qquad P_j = P_{j-1} + \frac{d_j}{\sum_i r_{i,j} \sum_i r_{i,j}},$$

Where $P_0=0$ and d_j is the weighted number of events at time t_j, j = 1,..,m.

7) The use of this structure allows the similar modification of the log-rank test using Greenwood formula.

Figure 1. *Creating a weighted composite endpoint.*

A Residual weight at the end of each day of follow-up

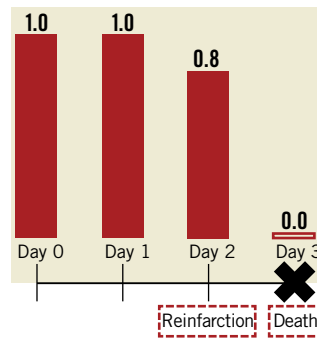**B** Residual weight at the end of each day of follow-up

Figure 2. *Illustration of residual weights at the end of each day of follow-up in a sample patient with cardiogenic shock on day 1 and reinfarction on day 3 (A), and in another patient with reinfarction on day 2 and death on day 3 (B).*

a patient experienced reinfarction on day 2 and then death on day 3 of follow-up. Note that the occurrence of death fully eliminates any residual weight. Based on this weighted event rate, the association between the study treatment and the composite endpoint would be analysed with a modified Kaplan-Meier survival analysis and modified log-rank score as the test between treatment groups.

A *post hoc* application of the weighted composite analysis of the Trial of Routine Angioplasty and Stenting after Fibrinolysis to Enhance Reperfusion in Acute Myocardial Infarction (TRANSFER-AMI) trial, comparing early PCI to standard treatment of STEMI, revealed two important observations which may not be uncommon in the traditional time-to-first-event analysis[10]. Recall that one assumption of the traditional composite endpoint is that the direction of association of treatment with all components of the composite should be the same. This assumption, however, was violated in the TRANSFER-AMI trial as the direction of death and cardiogenic shock was opposed to that of reinfarction and recurrent ischaemia; secondly, recurrent ischaemia clearly drove the benefit for early PCI. However, this event should or could be considered the least clinically relevant among the other components in the composite endpoint. Introducing the relative weights of the events calibrated the analysis such that no significant difference between the treatment groups was observed, unlike the standard time-to-first-event analysis **(Figure 3)**[13].

Another advantage of the weighted composite endpoint approach is the full use of the data collected, as all events, not only the first, are counted. This enhances precision and provides the potential to increase the statistical power. In a simulation study, we
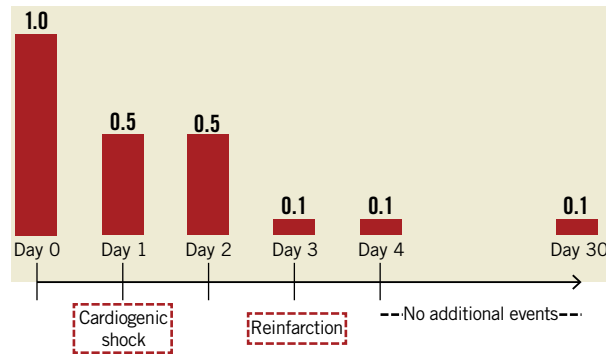
demonstrated that this approach not only allowed more events per patient in the analysis but also in turn reduced the variance of the estimate and provided additional discriminative power[9].

With the introduction of this novel approach to composite endpoints, we encourage current and future trialists to report the frequency of all events and to consider this more comprehensive analysis in the pre-specified analysis plan, including the *a priori* stakeholder evaluation of the relative weights. In addition to clinician-investigators, these stakeholders may include patients, sponsors, regulators and journal editors. This may have a particular impact in investigations with longer-term follow-up where non-fatal events have a greater window of opportunity to occur[11]. Beyond the single-study context, an extension of this approach to the synthesis of multiple studies should be considered in the future. By modifying our traditional approaches (and learning from Einstein's statement) we can make the best use of all clinical data collected and advance our understanding of treatments and the diseases we study.

Acknowledgements

Conflict of interest statement

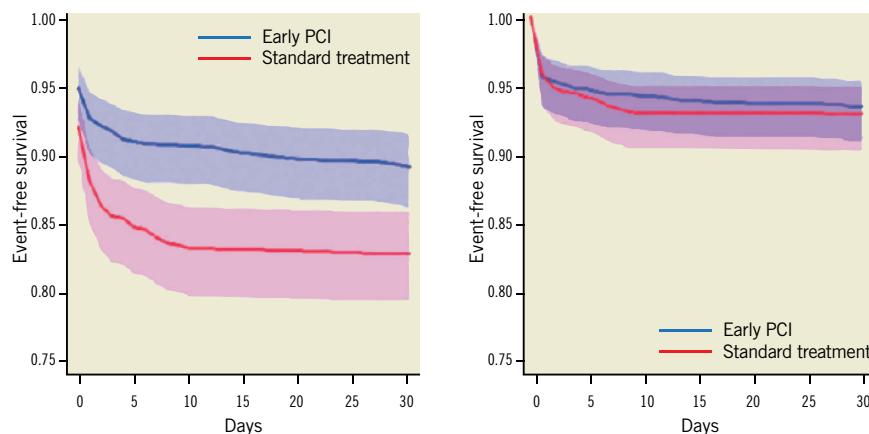The authors have no conflicts of interest to declare.

Figure 3. *Analysis of patients enrolled in the TRANSFER-AMI trial. Left panel: traditional time-to-first-event (30-day death [1.0], cardiogenic shock [1.0], congestive heart failure [1.0], reinfarction [1.0], and recurrent ischaemia [1.0]) comparing early PCI (blue: - line: Kaplan-Meier estimate; shaded area: 95% confidence region) to standard treatment (red: - line: Kaplan-Meier estimate; shaded area: 95% confidence region). Right panel: weighted composite endpoint (30-day death [1.0], cardiogenic shock [0.5], congestive heart failure [0.3], reinfarction [0.2], and recurrent ischaemia [0.1]) comparing early PCI (blue: - line: Kaplan-Meier estimate; shaded area: 95% confidence region) to standard treatment (red: - line: Kaplan-Meier estimate; shaded area: 95% confidence region).*

References

1. [No authors listed]. An international randomized trial comparing four thrombolytic strategies for acute myocardial infarction. The GUSTO Investigators. *N Engl J Med.* 1993;329:673-82.

2. Andersen PK, Gill RD. Cox's regression model for counting processes: a large sample study. *Ann Stat.* 1982;10:1100-20.

3. Prentice RL, Williams BJ, Peterson AV. On the regression analysis of multivariate failure time data. *Biometrika.* 1981;68:373-9.

4. Wei LJ, Lin DY, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *J Am Stat Assoc.* 1989;84:1065-73.

5. Bueno H, Armstrong PW, Buxton MJ, Danchin N, Lubsen J, Roland E, Verheugt FW, Zalewski A, Jackson N, Komajda M, Steg PG; Cardiovascular Round Table Clinical Trials ThinkTank participants. The future of clinical trials in secondary prevention after acute coronary syndromes. *Eur Heart J.* 2011;32:1583-9.

6. Zannad F, Garcia AA, Anker SD, Armstrong PW, Calvo G, Cleland JG, Cohn JN, Dickstein K, Domanski MJ, Ekman I, Filippatos GS, Gheorghiade M, Hernandez AF, Jaarsma T, Koglin J, Konstam M, Kupfer S, Maggioni AP, Mebazaa A, Metra M, Nowack C, Pieske B, Piña IL, Pocock SJ, Ponikowski P, Rosano G, Ruilope LM, Ruschitzka F, Severin T, Solomon S, Stein K, Stockbridge NL, Stough WG, Swedberg K, Tavazzi L, Voors AA, Wasserman SM, Woehrle H, Zalewski A, McMurray JJ. Clinical outcome endpoints in heart failure trials: a European Society of Cardiology Heart Failure Association consensus document. *Eur J Heart Fail.* 2013;15:1082-94.

7. Armstrong PW, Westerhout CM. The power of more than one. *Circulation.* 2013;127:665-7.

8. Armstrong PW, Westerhout CM, Van de Werf F, Califf RM, Welsh RC, Wilcox RG, Bakal JA. Refining clinical trial composite outcomes: an application to the Assessment of the Safety and Efficacy of a New Thrombolytic-3 (ASSENT-3) trial. *Am Heart J.* 2011;161:848-54.

9. Bakal JA, Westerhout CM, Armstrong PW. Impact of weighted composite compared to traditional composite endpoints for the design of randomized controlled trials. *Stat Methods Med Res.* 2012 Jan 24. [Epub ahead of print].

10. Bakal JA, Westerhout CM, Cantor WJ, Fernández-Avilés F, Welsh RC, Fitchett D, Goodman SG, Armstrong PW. Evaluation of early percutaneous coronary intervention vs. standard therapy after fibrinolysis for ST-segment elevation myocardial infarction: contribution of weighting the composite endpoint. *Eur Heart J.* 2013;34:903-8.

11. Bakal JA, Roe MT, Ohman EM, Goodman SG, Fox KA, Zheng Y, Westerhout CM, Hochman JS, Lokhnygina Y, Brown EB, Armstrong PW. Applying novel methods to assess clinical outcomes: insights from the TRILOGY ACS trial. *Eur Heart J.* 2015;36:385-92a.

12. Stafinski T, Menon D, Nardelli A, Bakal J, Ezekowitz J, Tymchak W, Welsh R, Gyenes G, Armstrong PW. Incorporating patient preferences into clinical trial design: results of the opinions of patients on treatment implications of new studies (OPTIONS) project. *Am Heart J.* 2015;169:122-31.e22.

13. Cantor WJ, Fitchett D, Borgundvaag B, Ducas J, Heffernan M, Cohen EA, Morrison LJ, Langer A, Dzavik V, Mehta SR, Lazzam C, Schwartz B, Casanova A, Goodman SG; TRANSFER-AMI Trial Investigators. Routine early angioplasty after fibrinolysis for acute myocardial infarction. *N Engl J Med.* 2009;360:2705-18.