

Composite Outcomes in Cardiovascular Research: A Survey of Randomized Trials

Eric Lim, MBChB, MSc, MD; Adam Brown, BSc, MBBChir; Adel Helmy, MA, MBBChir; Shafi Mussa, MA; and Douglas G. Altman, DSc

Background: Composite end points are common in clinical trials.

Objective: To describe how composite outcomes are used, constructed, and reported in cardiovascular trials and to evaluate the contribution of individual end points to the composite estimate of effect in those trials.

Design: Review of 2-group, parallel-design, randomized cardiovascular trials that used composite end points and were published in 14 clinical journals from 1 January 2000 to 1 January 2007.

Setting: Published randomized trials in cardiovascular medicine and surgery.

Participants: Two-group, parallel-design trials published in 14 leading general medical, cardiology, and cardiothoracic surgery journals from 1 January 2000 to 1 January 2007.

Measurements: The types and numbers of individual events included in the composite outcome and *P* values and risk estimates for the composite outcome.

Results: Of 304 trials published that used composite outcomes, 221 (73%) reported a composite primary outcome and 83 (27%) reported a composite secondary outcome. Composite outcomes

comprised a median of 3 (interquartile range, 3 to 4) individual outcomes; death was the most common individual outcome. The total number of individual events and the total number of events represented by the composite outcome differed in 79% of trials. *P* values for composite outcomes were less than 0.050 more frequently than they were 0.050 or greater. Death as an individual end point made a relatively minimal contribution to estimates of effect summarized by the trials' composite end points, whereas revascularization made a greater contribution.

Limitation: All-cause and cardiovascular mortality were not distinguished, and the findings might not apply to trials in other fields.

Conclusion: Composite outcomes in cardiovascular trials are frequent and commonly comprise 3 to 4 individual end points that vary in clinical significance. Discrepancies between the total number of individual events in a trial and those reported for composite outcomes are common. Individual outcomes do not contribute equally to composite measures, so the overall estimate of effect for a composite measure cannot be assumed to apply equally to each of its individual outcomes.

Ann Intern Med. 2008;149:612-617.

For author affiliations, see end of text.

www.annals.org

Clinical trials often use composite outcomes to provide an overall estimate of the effect of an intervention. Composite outcomes are appropriate under many circumstances, such as when no single most important outcome exists, when several end points are thought to be important, and when 2 end points are clinically related (such as nonfatal myocardial infarction and death from fatal myocardial infarction). Also, because death is an uncommon occurrence, very large sample sizes are required to detect differences in deaths among trial groups. Composite outcomes are thought to facilitate the evaluation of the effects of treatment on infrequent outcomes, such as death in smaller trials, and are a convenient way to represent a broader range of the beneficial effects of an intervention (1, 2).

Composite outcomes, however, require that the constituent outcomes are clinically meaningful and interpretable (3). Ideally, those individual outcomes should share the same biological effect: Each should contribute equally

to the composite, and patients should attach a similar importance to each individual outcome (4). In addition, clinicians and statisticians have noted that composite outcomes are difficult to interpret when results are inconsistent for the individual outcomes, composite outcomes often overstate the results of the individual outcomes, and composite outcomes are used even if each individual event is not considered equally important (4–7).

To better understand these issues, we reviewed published reports of randomized trials in cardiovascular medicine and surgery to assess how composite outcomes are currently used, constructed, and reported and to quantitatively assess the contribution of common, clinically important individual outcomes to the composite outcome.

METHODS

Study Identification

To identify published cardiovascular clinical trials, we hand-searched each issue of 14 leading general medical, cardiology, and cardiothoracic surgery journals (*New England Journal of Medicine*, *Lancet*, *JAMA*, *BMJ*, *Archives of Internal Medicine*, *Annals of Internal Medicine*, *Circulation*, *Journal of the American College of Cardiology*, *American Heart Journal*, *Heart*, *European Heart Journal*, *Journal of Thoracic and Cardiovascular Surgery*, *Annals of Thoracic*

See also:

Print

Editors' Notes 613

Web-Only

Conversion of graphics into slides

Context

Many trials combine several events into a single composite outcome when reporting the effect of an intervention.

Contribution

Of cardiovascular trials published over 7 years, 37% used composite outcomes. The individual events making up the composite differed in their clinical significance and in their influence on the composite estimate of effect.

Caution

The findings may not apply to trials in other specialties.

Implication

Composite outcomes are used commonly in cardiovascular trials and rely on end points that vary in their clinical significance. The overall estimate of effect for a composite measure cannot be assumed to apply equally to each individual outcome.

—The Editors

Surgery, and *European Journal of Cardiothoracic Surgery*) from 1 January 2000 to 1 January 2007. We identified a cardiovascular trial on the basis of any reported drug treatment or mechanical intervention aimed at improving myocardial, heart valve, or coronary artery function. We accepted any trial that was described as “randomized” by the authors as randomized. A composite outcome was any outcome that combined 2 or more different outcome measures. Death as an outcome could refer to all-cause mortality or cardiovascular mortality.

Data Abstraction

We identified all potentially relevant abstracts from hand-searched journals and retrieved full reports of manuscripts for review. We included all 2-group, parallel-design trials and excluded studies with more than 2 groups and those that did not provide a breakdown of the composite outcome (by not specifying the individual outcomes or by reporting the composite outcomes in a time-to-event analysis).

Four researchers abstracted the data, and we resolved differences by consensus. We compared the number of composite outcome events with the number of events for each individual outcome for each trial. When composite outcomes were used for both primary and secondary outcomes, we included the primary composite outcome only so that each trial contributed only 1 composite outcome to the analysis. When the same group of composite outcomes was reported more than once, we used results from the longest reporting interval. When more than 1 composite outcome was reported, we used the composite outcome with the largest number of component outcomes. Because more than 1 end point can occur in a single patient and occasionally the same end point can occur more than once, we screened each article to determine whether each end

point was counted as a separate event or whether a hierarchy of end points was used. A study with a hierarchy of end points would contribute a single event for each patient, either the first to occur or the most important.

Statistical Analysis

We measured how frequently composite outcomes are used in cardiovascular trials and how frequently specific individual outcomes are used and included as part of a composite outcome. We also quantified the number of individual outcomes used in the composite, and the proportion of the composite outcome contributed by each individual outcome. We obtained bias-corrected and accelerated CIs by using bootstrapping with 1000 replications for these descriptive data.

We used linear regression to ascertain the relationship between study sample size (dependent variable) and the number of end points (explanatory variable). For outcomes of death and revascularization, we determined how the individual outcome influences the estimated strength of association between the intervention and outcome. We did this by creating a log–log, inverse variance, weighted circle plot of the estimated relative risk for the composite outcome against the estimated relative risk for the composite outcome without death or revascularization as an individual outcome. The larger circles in those plots represent studies in which the number of events is greater and the variance is smaller, usually resulting from estimates derived from larger sample sizes. If the relative risk for the composite outcome with and without the end point (death or revascularization) remained exactly the same, the circle would fall on a diagonal line from the bottom left to the top right of the plot. We arbitrarily defined a clinically

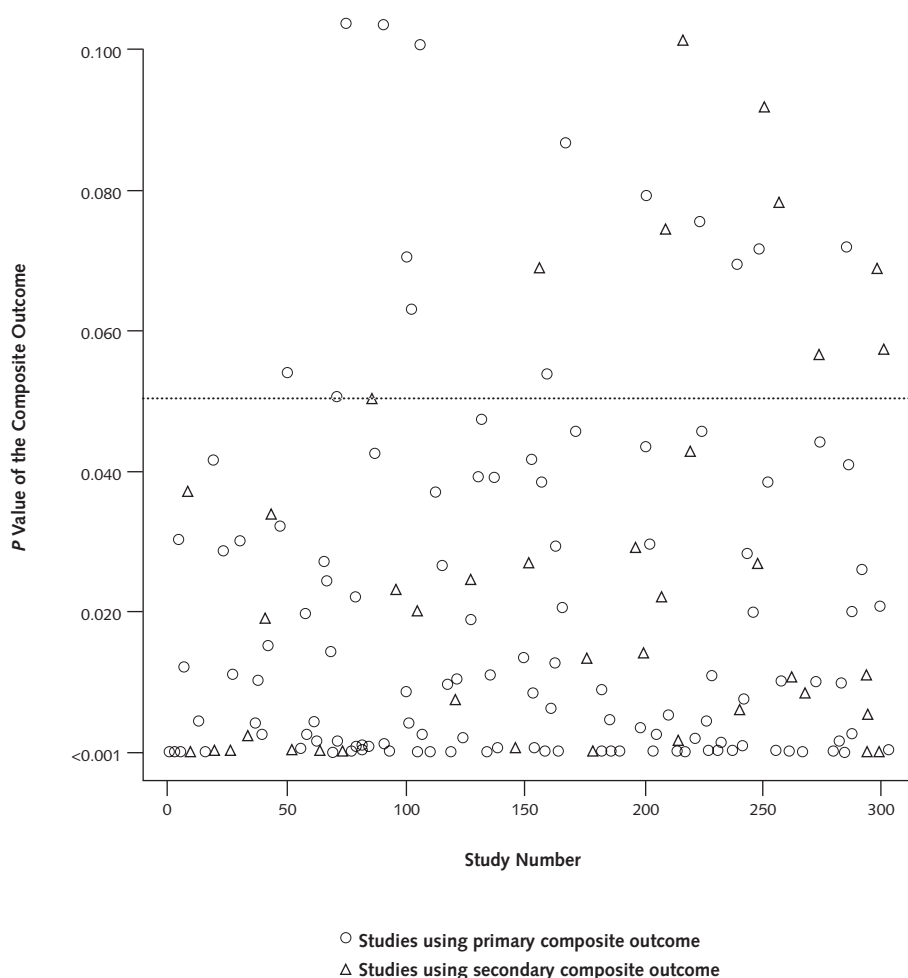
Table. Frequency and Contribution of End Points

| Outcome | Trials, n (%) | Mean Percentage of Composite (95% CI) |
|------------------------------------------------|---------------|---------------------------------------|
| Composite outcome* | 304 (100) | 100 (NA) |
| Individual end points | | |
| Death | 298 (98) | 27 (24–29) |
| Myocardial infarction | 280 (92) | 37 (34–40) |
| Reintervention | 164 (54) | 58 (53–63) |
| Stroke | 98 (32) | 15 (13–18) |
| Angina | 30 (10) | 44 (35–54) |
| Hospitalization | 36 (12) | 50 (41–59) |
| Cardiac failure | 25 (9) | 34 (26–43) |
| Combined end points that included death | | |
| Death and myocardial infarction | 274 (90) | NA |
| Death and reintervention | 162 (53) | NA |
| Death and stroke | 94 (31) | NA |
| Death and angina | 28 (9) | NA |
| Death and hospitalization | 36 (12) | NA |
| Death and cardiac failure | 23 (8) | NA |

NA = not applicable.

* The composite outcome may have included additional end points.

Figure 1. P values of the composite outcome.



A symmetrical plot of the *P* values of the composite outcome, truncated between <0.001 and 0.100, illustrates the asymmetrical distribution of the points that lie predominantly below the dotted line, where $P = 0.050$.

meaningful difference of 20% in the relative risk estimate between the composite outcome and the composite outcome without death or revascularization, and we illustrate this 20% threshold with 2 dashed lines on either side of the diagonal line. Circles outside this boundary represent a clinically meaningful difference in risk with and without the individual outcome (death or revascularization), and larger circles outside this boundary represent clinically meaningful differences with greater certainty. Larger circles toward the top left and bottom right quadrants signify a greater contribution of the individual outcome (death or revascularization) to a trial's conclusions based on the composite outcome.

We used Stata, version 9.0 (StataCorp, College Station, Texas), and R, version 2.4.1 (R Foundation for Statistical Computing, Vienna, Austria) for statistical analyses.

Role of the Funding Source

No funding was received for this study.

RESULTS

We identified 1231 randomized cardiovascular trials published between 1 January 2000 and 1 January 2007, of which 454 (37%) used composite end points. We excluded 150 trials that did not have a 2-group, parallel design or did not specify individual outcomes, which left 304 trials (28%): 221 (73%) that used a composite outcome as a primary outcome and 83 (27%) that used a composite outcome for its secondary outcome only.

A median of 3 (interquartile range, 3 to 4) individual outcomes were used by researchers to create their composite outcome. The Table shows the distribution and contribution of individual outcomes to the trials' composite outcome. Death was the most common end point in the composite outcome in 298 (98%) trials.

Consistent with the argument that the smaller the sample size of a trial, the more end points are included in the composite outcome to inflate the number of events,

our linear regression analysis suggested that the inclusion of each additional individual end point to the composite was associated with 721 fewer trial participants ($P = 0.071$).

The total number of events for the individual outcomes was more than the total number of events represented by the composite outcome in 222 (73%) trials and was fewer than the total number of events represented by the composite outcome in 17 (6%) trials. We could not determine whether the discrepancy in the former instance was attributable to the use of a hierarchy (that is, in which only 1 event in a patient who had 2 or more events was selected for inclusion in the composite outcome) because most trials did not report whether they used a hierarchy.

A plot of the P values of the composite outcome revealed marked asymmetry around the midpoint of the plot (at $P = 0.050$) (Figure 1), suggesting possible publication bias or that individual outcomes were selected for inclusion in the composite outcome to ensure statistical significance for the composite.

Death and revascularization differed in their contribution to the composite outcomes. In Figure 2, the circles scattered away from the dashed parallel lines toward the top left and bottom right quadrants are few and small, suggesting that differences in estimates of relative risk based on composite outcomes with and without death may be attributable to the random variability that comes from trials with few deaths, and that death as an individual end point makes a relatively small contribution to the overall effect summarized by trials' composite end point. In contrast, in Figure 2, the circles scattered away from the dashed parallel lines toward the top left and bottom right quadrants are larger, indicating that revascularization as an individual end point is common and makes a much larger contribution to estimates of effect summarized by the trials' composite end point.

DISCUSSION

A recent survey found that cardiovascular trials use composite outcomes more than trials in any other field (7). We found that of 1231 such trials published over 7 years, 37% reported a composite outcome—most using it as their primary outcome measure—that comprised 3 to 4 individual end points. Seventy-nine percent of trials had a discrepancy in the total number of individual events and the total number of events represented by the composite outcome. Death had little influence on the estimate of effect based on the composite outcome, whereas revascularization had far greater influence.

Our results substantiate previously published impressions that estimates of effect based on composite outcomes are sometimes inconsistent with those for the outcomes' individual components and that composite outcomes are frequently used when the assumption that each individual event is equally important is not met (4–7). The findings

also demonstrate that commonly occurring but less clinically important end points contribute more to composite estimates of effect and that less commonly occurring but more clinically serious end points contribute less, albeit with substantial statistical uncertainty. We propose that it may be better to avoid using a composite outcome under these circumstances (4), but if it cannot be avoided, caution should be used when interpreting the significance of the overall composite in relation to the individual end points.

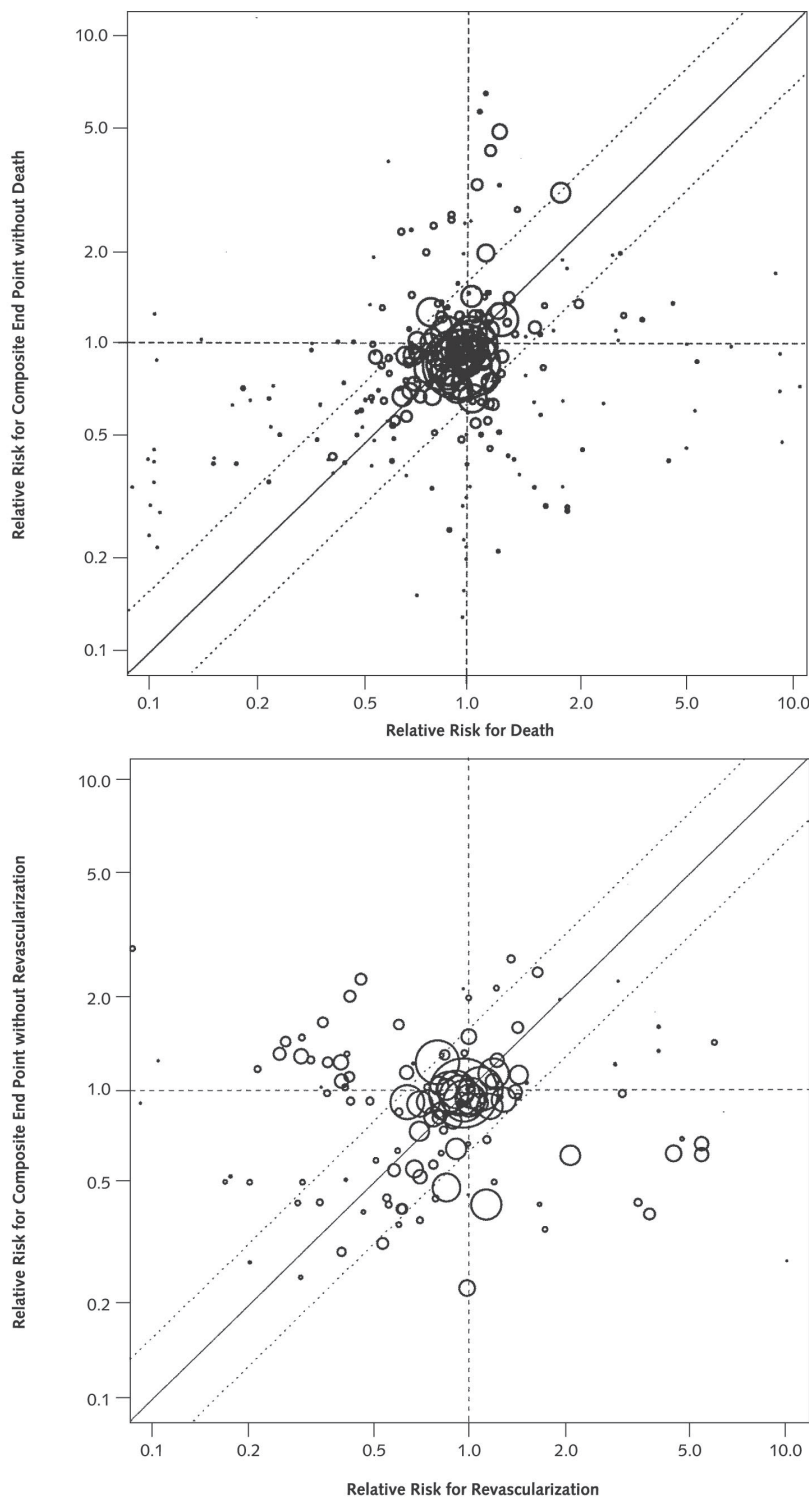
Some would argue that death should always be included in a composite outcome so that event-free survival is reported (8), but this requirement conflicts with the premise that each individual end point should have equal importance (4). Few would agree that death is as important as revascularization, a combination used in 53% of cardiovascular trials (9). When death is part of a composite outcome (98% of surveyed trials), competing risks can influence the results because patients who die cannot have any further end points. The extent to which the results are affected increases as the magnitude of imbalance in the number of deaths in the 2 groups increases. One can account for death with time-to-event analysis (survival), but we recommend the simultaneous reporting of each individual end point so that readers can ascertain its individual contribution.

Regardless of whether time-to-event analysis is used, the problem of several events occurring in individual patients can be addressed by counting only 1 event per patient. We acknowledge that this introduces its own difficulties. For example, such strategies as reporting the first event only emphasize early outcomes compared with later outcomes, and selecting an event by perceived magnitude of importance conflicts with the premise that each event is equally important.

Our study has several limitations. We did not use statistical methods to formally compare the risk estimates of composite outcomes, and no consensus exists about the presentation, classification, and interpretation of magnitude of difference in those estimates. Clearly, further work is required in this area. Also, we did not attempt to distinguish all-cause from cardiovascular mortality in the analysis, and one would expect heterogeneity in relative risk ratios obtained by using different (between-study) definitions of death. Finally, our review and analysis was restricted to cardiovascular trials with cardiovascular outcomes, and the relationship between composite and individual outcomes might substantially differ in other fields.

In summary, composite outcomes in cardiovascular trials are common and comprise individual outcomes that vary in clinical significance. More serious end points, such as death, that occur less frequently have relatively little influence on estimates of effect represented by composite measures, whereas less serious end points, such as revascularization, are more frequent and have greater influence on those measures. Therefore, readers are cautioned against

Figure 2. A log–log, inverse variance, weighted circle plot of the relative risk for death against the composite end point without death (*top*) and the relative risk for revascularization against the composite end point without revascularization (*bottom*).



Top. Each circle represents the relative risk for death plotted against the relative risk for the composite outcome without death. The area of the circles is inversely proportional to the variability of the relative risk for death (larger circles represent smaller variances). The solid diagonal line represents perfect agreement, and the dotted, parallel lines on either side of it represent a 20% difference in the relative risk. The large scatter of small circles indicates the uncertainty of the contribution of death to the overall composite outcome. **Bottom.** The scatter of moderately large circles represents effects with greater certainty. The increasing distances and larger circles toward the top left and bottom right indicate the greater propensity to influence the direction of the overall composite outcome.

assuming that the overall estimate of effect for the composite outcome can be interpreted to be the same for each individual outcome. Clearer reporting of hierarchies, in which individual outcomes were selected for inclusion in the composite outcome, and a clear listing of the individual end points and the number of participants experiencing them would make the interpretation of composite outcomes more meaningful.

From Papworth Hospital and Addenbrooke's Hospital, Cambridge, and the Centre for Statistics in Medicine, Oxford, United Kingdom.

Potential Financial Conflicts of Interest: None disclosed.

Reproducible Research Statement: *Study protocol:* Not available. *Statistical code:* Available from Dr. Lim (e-mail, e.lim@rbht.nhs.uk). *Data set:* Available subject to approval by the study committees by contacting Dr. Lim (e-mail, e.lim@rbht.nhs.uk).

Requests for Single Reprints: Eric Lim, MBChB, MSc, MD, Imperial College and Academic Division of Thoracic Surgery, Royal Brompton Hospital, Sydney Street, London SW3 6NP, United Kingdom; e-mail, e.lim@rbht.nhs.uk.

Current author addresses and author contributions are available at www.annals.org.

References

1. Braunwald E, Cannon CP, McCabe CH. Use of composite endpoints in thrombolysis trials of acute myocardial infarction. *Am J Cardiol.* 1993;72:3G-12G. [PMID: 8279357]
2. Cannon CP. Clinical perspectives on the use of composite endpoints. *Control Clin Trials.* 1997;18:517-29; discussion 546-9. [PMID: 9408715]
3. Friedman LM, Furberg CD, DeMets DL. *Fundamentals of Clinical Trials.* 3rd ed. New York: Springer; 1998.
4. Montori VM, Permyer-Miralda G, Ferreira-González I, Busse JW, Pacheco-Huergo V, Bryant D, et al. Validity of composite end points in clinical trials. *BMJ.* 2005;330:594-6. [PMID: 15761002]
5. Lauer MS, Topol EJ. Clinical trials—multiple treatments, multiple end points, and multiple lessons [Editorial]. *JAMA.* 2003;289:2575-7. [PMID: 12759328]
6. DeMets DL, Califf RM. Lessons learned from recent cardiovascular clinical trials: Part I. *Circulation.* 2002;106:746-51. [PMID: 12163438]
7. 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. [PMID: 12759327]
8. Lubsen J, Kirwan BA. Combined endpoints: can we use them? *Stat Med.* 2002;21:2959-70. [PMID: 12325112]
9. Massel D. Composite confusion [Letter]. *J Am Coll Cardiol.* 2004;43:1926-7; author reply 1927. [PMID: 15145125]

DOWNLOAD IMPORTANT REFERENCES TO CITATION MANAGERS

At www.annals.org, article citations may be directly downloaded to any of the following formats: EndNote, Reference Manager, ProCite, BibTeX, or Medlar.

Current Author Addresses: Dr. Lim: Imperial College and Academic Division of Thoracic Surgery, Royal Brompton Hospital, Sydney Street, London SW3 6NP, United Kingdom.

Drs. Brown and Helmy and Mr. Mussa: Department of Cardiothoracic Surgery, Papworth Hospital, Papworth Everard, Cambridge CB23 8RE, United Kingdom.

Dr. Altman: Centre for Statistics in Medicine, Wolfson College Annexe, Linton Road, Oxford OX2 6UD, United Kingdom.

Author Contributions: Conception and design: E. Lim.

Analysis and interpretation of the data: E. Lim, D.G. Altman.

Drafting of the article: E. Lim, A. Helmy, S. Mussa, D.G. Altman.

Critical revision of the article for important intellectual content: E. Lim, A. Helmy, S. Mussa, D.G. Altman.

Final approval of the article: E. Lim, A. Helmy, S. Mussa, D.G. Altman.

Statistical expertise: E. Lim, D.G. Altman.

Administrative, technical, or logistic support: A. Helmy, S. Mussa.

Collection and assembly of data: E. Lim, A. Brown, A. Helmy, S. Mussa.