BJA

British Journal of Anaesthesia, 126 (1): 56–66 (2021)

doi: 10.1016/j.bja.2020.09.023
Advance Access Publication Date: 20 October 2020
Review Article

CARDIOVASCULAR

Systematic review and consensus definitions for the Standardized Endpoints in Perioperative Medicine (StEP) initiative: cardiovascular outcomes

W. Scott Beattie[1,*], Manoj Lalu[2], Matthew Bocock[2], Simon Feng[2], Duminda N. Wijeysundera[3], Peter Nagele[4], Lee A. Fleisher[5], Andrea Kurz[6], Bruce Biccard[7], Kate Leslie[8], Simon Howell[9], Giovanni Landoni[10], Hilary Grocott[11], Andre Lamy[12], Toby Richards[13], Paul Myles[14] on behalf of the StEP COMPAC Group[†]

[1]Cardiovascular Anesthesia, University Health Network, University of Toronto, Toronto, ON, Canada, [2]Department of Anesthesia, University of Ottawa, Ottawa, ON, Canada, [3]Li Ka Shing Knowledge Institute, St Michaels Hospital, University of Toronto, Toronto, ON, Canada, [4]Department of Anesthesia and Critical Care, University of Chicago, Chicago, IL, USA, [5]Department of Anesthesiology and Critical Care, University of Pennsylvania Health System, Philadelphia, PA, USA, [6]Outcomes Research, Cleveland Clinic, Cleveland, OH, USA, [7]Department of Anesthesia and Perioperative Medicine, University of Cape Town, Cape Town, South Africa, [8]Department of Anaesthesia and Pain Management, Royal Melbourne Hospital, Melbourne, Vic., Australia, [9]University of Leeds School of Medicine, Leeds, UK, [10]Center for Intensive Care and Anesthesiology, IRCCS San Raffaele Scientific Institute, Milan, Italy, [11]Department of Anesthesia and Perioperative Medicine, University of Manitoba, Winnipeg, MB, Canada, [12]Department of Surgery, McMaster University, Hamilton, ON, Canada, [13]University of Western Australia, Perth, Australia and [14]Alfred Health and Monash University Department of Anaesthesia and Perioperative Medicine, Melbourne, Vic., Australia

*Corresponding author. E-mail: scott.beattie@uhn.ca

[†]The full list of the Standardized Endpoints in Perioperative Medicine Core Outcome Measures in Perioperative and Anaesthetic Care members (StEP COMPAC) is found in .Supplementary Appendix 1

Abstract

Background: Adverse cardiovascular events are a leading cause of perioperative morbidity and mortality. The definitions of perioperative cardiovascular adverse events are heterogeneous. As part of the international Standardized Endpoints in Perioperative Medicine initiative, this study aimed to find consensus amongst clinical trialists on a set of standardised and valid cardiovascular outcomes for use in future perioperative clinical trials.

Methods: We identified currently used perioperative cardiovascular outcomes by a systematic review of the anaesthesia and perioperative medicine literature (PubMed/Ovid, Embase, and Cochrane Library). We performed a three-stage Delphi consensus-gaining process that involved 55 clinician researchers worldwide. Cardiovascular outcomes were first shortlisted and the most suitable definitions determined. These cardiovascular outcomes were then assessed for validity, reliability, feasibility, and clarity.

Results: We identified 18 cardiovascular outcomes. Participation in the three Delphi rounds was 100% (n=19), 71% (n=55), and 89% (n=17), respectively. A final list of nine cardiovascular outcomes was elicited from the consensus: myocardial infarction, myocardial injury, cardiovascular death, non-fatal cardiac arrest, coronary revascularisation, major adverse cardiac events, pulmonary embolism, deep vein thrombosis, and atrial fibrillation. These nine cardiovascular outcomes were rated by the majority of experts as valid, reliable, feasible, and clearly defined.

Received: 08 June 2020; **Accepted:** 9 September 2020

Conclusions: These nine consensus cardiovascular outcomes can be confidently used as endpoints in clinical trials designed to evaluate perioperative interventions with the goal of improving perioperative outcomes.

Keywords: cardiovascular events; clinical trials; MACE; myocardial infarction; outcome measures; perioperative medicine; standardised endpoint

Editor's key points

- Definitions of perioperative cardiovascular adverse events are heterogeneous, but are critical for use in future perioperative clinical trials.
- As part of the international Standardized Endpoints in Perioperative Medicine initiative, a systematic review and consensus process by clinical trialists was used to develop a set of standardised and valid cardiovascular outcomes.
- Nine cardiovascular outcomes were rated by the majority of experts as valid, reliable, feasible, and clearly defined.
- These outcomes can be confidently used as endpoints in clinical trials designed to evaluate perioperative interventions.

Major noncardiac surgery is undertaken to cure or treat debilitating diseases, or is performed as a palliative measure with the overarching goal to improve a patient's quality of life. Surgery is now well recognised to be frequently associated with adverse events or complications that mitigate the planned positive effects from surgery.[1] Cardiovascular events are particularly common, causing increased disability, costs, and mortality.[2–4]

Clinical trials are required to study interventions that reduce the incidence and magnitude of cardiovascular events after surgery. Perioperative clinical trials require clearly defined cardiovascular outcomes for transparent reporting to improve interventions and patient-centred care.

The Standardized Endpoints in Perioperative Medicine (StEP) initiative is an international collaboration with the aim of identifying a set of endpoints supported by literature, expert guidance, and international consensus for use in perioperative medicine trials.[5] The current study describes the results of a systematic literature review and an anonymous Delphi process with the goal of standardising cardiovascular outcomes for use in future perioperative clinical trials.

Methods

The overall methodology was similar to previously published StEP projects.[6,7] We undertook a systematic review of noncardiac perioperative clinical trials (2005–17). The StEP Steering Committee had made the decision in 2015 to review the last 10 yr of perioperative literature, such that the first year reviewed was 2005. We conducted the search in 2017. A Delphi process then followed, which is a validated method for establishing consensus of diagnostic criteria[8,9] to refine existing cardiovascular outcomes definitions.

Inclusion/exclusion criteria and definitions

Our systematic review of relevant literature followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines as outlined in detail in the Supplementary material. The online supplement includes an *a priori* protocol, search strategy, and results. The review was not registered. To summarise the process in brief, we included RCTs and experimental and observational studies that reported interventions to improve cardiovascular outcomes within the perioperative setting in adults ≥18 yr old. Included trials were required to have a sample size of >150 patients. We did not consider studies that used outcomes related exclusively to critical care.

Literature search and data extraction

We performed a systematic search on MEDLINE, Embase, and the Cochrane database for studies published between January 1, 2005 and December 31, 2017 in core clinical journals, as defined by the National Library of Medicine. To identify potentially eligible studies according to title and abstract content, two authors (ML and MB) independently performed the literature search (Supplementary Appendix 2). The reference lists of retrieved articles were also searched for additional studies. We did not apply any language restriction. An *a priori* (September 2016) detailed scoping protocol description was provided (Supplementary Appendix 3). Selected articles were independently analysed by the two teams of authors and extracted according to a standardised extraction and coding template (Supplementary Appendix 4) using commercially available systematic review software (DistillerSR; Evidence Partners, Kanata ON, Canada). Extractors were trailed using five predefined articles. The overall process was performed by eight authors. Discrepancies were resolved by consensus between three extractors (ML, MB, and WSB). Risk of bias in studies was not assessed, as the purpose of the review was to identify the scope, definitions, and validity of cardiovascular outcomes in common use, not the efficacy of specific interventions.

Major adverse cardiovascular event (MACE) is a composite outcome that has been utilised as the primary outcome in all the major noncardiac perioperative clinical trials. During the systematic review, it was noted that the definition of MACE lacked uniformity; we therefore decided to ask members to rate each definition used in trials and individual components of the varied MACE definitions. Likewise, the definition of cardiovascular death was heterogeneous. We asked members to rate individual components of the varied definitions.

Delphi process

We used a Delphi method to gain consensus around the clarity, reliability, and validity of each cardiovascular outcome. The initial list of retrieved definitions was then provided to members of the StEP cardiovascular subgroup.

Delphi Round 1

After a discussion with the subgroup members, the theme subgroup chair (WSB) prepared the initial list of endpoints and associated definitions retrieved from the literature according to a predefined format prepared by the StEP Steering Committee. All members of the clinical cardiovascular subgroup (n=12) and the overall StEP Steering Committee (n=7) were invited to participate. The participants were asked to score each of the listed indicators for clinical importance using a scale of 1–9. Scores of 1–3 indicated 'not that important or invalid', 4–6 indicated 'important but require revision', and 7–9 'critical for inclusion'. The participants were offered the option to select 'not applicable/not sure' if they were unable to form an opinion about the importance or not of the clinical indicator. The participants had 2 weeks to answer before reminder e-mails (up to three) were sent to prompt completion of the survey. For each indicator, the participants were also invited to add any comments and suggestions for modifications of existing definitions that they believed were important. Individual indicator scores were then calculated using mean, median, and range of scores. The comments and suggestions provided by the participants were collated to be integrated to the second Delphi round.

Delphi Round 2

The theme subgroup chair (WSB) selected indicators that had been rated as 'critical' (score 7–9) by at least 70% of participants to prepare the first list of indicators for Delphi round 2. Outcomes rated as 'not that important or invalid' (a score of 1–3) or as 'important but requiring revision' (score 4–6) by at least 70% of participants were also included in the second round, but clearly identified as such on the Delphi Round 2 data collection form.

The participants were asked to score the cardiovascular outcomes using the same questionnaire format and rating procedure as the one used during Delphi Round 1. For Round 2, the participants were provided with the mean scores of each clinical indicator after Round 1. Comments after Delphi Round 1 were also added. The second Delphi round was circulated via e-mail to the entire StEP Working Group (n=76).

Delphi Round 3

Before Round 3, the results of Round 2 were sent to all members of the cardiovascular subgroup for their input. The theme subgroup chair (WSB) then selected the cardiovascular endpoints that had been rated as 'critical' (score >7) by at least 70% of participants during the second round for the third Delphi round. Outcomes rated as 'not that important or invalid' (score 1–3) or 'important but requiring revision' (score 4–6) were not included. If responses to the second stage Delphi process comments section suggested that modification to endpoint definitions or rating had to be made, this was discussed within the theme subgroup of that indicator via e-mail. For this third round, the participants were provided with the short list of selected indicators and attached definitions and all comments provided after Rounds 1 and 2. They were asked to score the item using a second questionnaire. The questionnaire included four rating criteria per indicator:

(i) Validity: the degree to which the indicator measures what it purports to measure
(ii) Reliability: the degree of stability of the indicator when measurement is repeated under identical conditions
(iii) Feasibility: practicability/ease of use in the clinical setting
(iv) Clarity of the definition: the degree to which the clinical indicator meaning can be easily understood

For each question, the participants were again asked to rate each cardiovascular outcome on a 1–9 scale. Scores of 1–3 indicate 'no', 4–6 indicate 'unsure', and 7–9 indicate 'yes'; meanwhile, a score of 10 meant 'not assessable'. At the end of the third Delphi round, cardiovascular outcomes that had a score of 7–9 ('yes') in more than 70% of responses for each question were automatically included. Cardiovascular outcomes rated 4–6 ('unsure') for one or several of the four rating criteria were discussed by e-mail within the indicator's subgroup. Those that had a score of 1–3 ('no') for any of the rating criteria were considered only as optional, but not recommended.

Each Delphi round was coordinated by the Department of Anaesthesia and Perioperative Medicine at the Alfred Hospital in Melbourne, Australia. The participants' answers to the different Delphi rounds were anonymised, recorded, transformed, and analysed using the STATA v14 (StataCorp, College Station, TX, USA). Data are reported as mean, median, number, and proportion of respondents. Comparisons between myocardial injury and myocardial injury after noncardiac surgery (MINS) in Round 2 were analysed using the Wilcoxon signed-rank test. In Round 3, validity, reliability, feasibility, and clarity were compared using paired t-tests.

Results

A total of 6342 studies were initially identified, of which 158 were selected for further analyses (Supplementary Appendix 5). After full content assessment and exclusion of duplicates, 21 reports were eliminated. Another 16 publications were excluded, as there was no information or definition of a cardiovascular outcome. The subject of the remaining 121 studies is shown in Table 1. A final list of 18 outcomes was carried forward to the Delphi process (Supplementary Appendix 6).

Participation to the different Delphi rounds was 100% (n=19), 71% (n=55), and 89% (n=17), respectively. The results of Delphi Rounds 1 and 2 are shown in Table 2.

None of the cardiovascular outcomes selected in the first Delphi round were removed for the second Delphi round. All cardiovascular outcomes with a median score of ≥7, and rated important by more than 70% of responders, in Round 2 were carried forward to the third round (Supplementary Appendix 6).

There are two notable exceptions to this. First, MINS (median score: 7; rated ≥7 by 51%) could not have been carried forward to Round 3. Myocardial injury after noncardiac surgery was retained in Round 3, as it is currently being widely used as an outcome[10] (and a clinical decision point)[11] in studies conducted by the Population Health Research Institute. In Delphi Round 3, MINS was compared with myocardial injury, as defined in the Fourth Universal Definition of Myocardial Infarction.[12] Second, congestive heart failure is an important cardiovascular outcome, but the definition used in

Table 1 Type and number of studies extracted. A complete list of these studies is shown in Supplementary Appendix 5. BNP, B-natriuretic peptide; ENIGMA, Evaluation of Nitrous Oxide in the Gas Mixture for Anaesthesia; METS, Measurement of Exercise Tolerance Before Surgery; POISE, Perioperative Ischemic Evaluation Study; VISION, Vascular Events In Noncardiac Surgery Patients Cohort Evaluation Study.

Type of study	Number of articles
Statins	8
Carotid stenting *vs* carotid endarterectomy	21
Anaemia/transfusion	8
Biomarkers (troponin, BNP, etc.)	6
Drug intervention (POISE, ENIGMA, etc.)	21
Regional *vs* general anaesthesia	10
Volatile/ischaemic preconditioning	6
Thrombin inhibitors (prevention of deep venous thrombosis and *pulmonary embolism*)	11
Prophylactic revascularisation	4
Prospective cohort studies (VISION and METS)	15
Miscellaneous	11
Total	121

the major perioperative trials is not consistent with state-of-the-art definitions used in cardiology. No trials have to date utilised the newer definitions being promoted by the American Heart Association (AHA)[13] or the European Society of Cardiology (ESC).[14]

During Delphi Round 2, the Fourth Universal Definition of Myocardial Infarction was published supplanting the Third Universal Definition,[15] which had been widely used in perioperative trials. We therefore changed the myocardial infarction definition based on this global consensus document.

The final list included 11 cardiovascular outcomes rated for validity, reliability, usability, and clarity of definitions, and is provided in Table 3. All cardiovascular outcomes, except MINS (54%) and congestive heart failure (23%), were rated as valid by more than 70% of the 17 evaluators. The same rating (score ≥7) for reliability was provided for all cardiovascular outcomes, except MINS, cardiovascular death, and congestive heart failure. Likewise, when estimating usability, all cardiovascular outcomes were rated as ≥7 by more than 70% of evaluators, except for MINS, cardiovascular death, and congestive heart failure. Congestive heart failure and MINS also failed to achieve a score of ≥7 by 70% of evaluators for the clarity of the definition.

In Delphi Round 2, myocardial injury had a median score 8 (6–9) and 79% of respondents scored myocardial injury ≥7, whereas for MINS the median score was 7 (3–9) and was scored ≥7 by 54%. Myocardial injury was scored higher in 33/55, whilst 6/55 rated MINS higher (Wilcoxon $z=3.494$; $P<0.0005$); the remaining 16 evaluations gave the same score for injury and MINS. In Delphi Round 3, paired analysis showed myocardial injury had higher rating for validity ($d=2.4$ [0.58]; $P<0.001$), reliability ($d=2.5$ [0.61]; $P<0.001$), feasibility ($d=2.6$ [0.66]; $P<0.001$), and clarity ($d=2.6$ [0.63]; $P<0.001$) than MINS.

The consensus opinion was that when reporting myocardial infarction, it was imperative to report the number of patients with postoperative troponin measurements (median: 8 [range: 2–9] score ≥7; 76%). When reporting myocardial injury, the panel recommends reporting the incidence of myocardial infarction (median: 9 [5–9] score ≥7; 74%), that the incidence of acute and chronic renal failure be reported (median: 7 [3–9] score ≥7; 61%), and that postoperative ECG be performed on all patients with a troponin elevation (median: 8 [1–9] score ≥7; 69%).

The definition of cardiac death was reached by evaluating individual components. In Delphi Round 2, the consensus suggested cardiac death should exclude deaths attributable to haemorrhage (median: 2 [1–8] score ≥7; 7%), pulmonary embolism (median: 5 [1–9] score ≥7; 34%), stroke (median: 6 [1–9] score ≥7; 34%), and unknown causes (median: 3 [1–8] score ≥7; 9%).

MACE was evaluated in Delphi Round 2. None of the major trials used the same definition. The highest-rated definition was that used in Perioperative Ischemic Evaluation Study 1: cardiac death, myocardial infarction, and non-fatal cardiac arrest (median: 7 [3–9] score ≥7; 43%). The consensus recommended the exclusion of pulmonary embolism (median: 2 [1–9] score ≥7; 2%) and haemorrhage (median: 2 [1–9] score ≥7; 2%) from future definitions of MACE.

Discussion

As a result of a protocolised systematic review and three anonymous Delphi consensus-seeking rounds with feedback from 55 international perioperative clinical trialists, the StEP group has identified nine cardiovascular outcomes that can be used in future trials to measure the effectiveness of perioperative interventions (Table 4).

At the outset of this process, we surmised that a set of standardised cardiovascular outcomes would be critically important for future clinical investigations. Standardisation would allow for transparent result reporting and homogeneous meta-analyses. We asked the panel to consider cardiovascular outcomes used in a perioperative setting and to agree on cogent definitions. The scientific basis of the methods used has been widely validated and accepted.

The literature search found that the definitions of several cardiovascular definitions conflicting. Further, we found heterogeneity for many of these outcome definitions. Heterogeneity is a major issue, as it can significantly impact on the reliability of clinical trials, which in turn can limit the development of effective interventions. Heterogeneity also limits comparison between studies and further combination of studies in meta-analyses. As an example, MACE is the composite outcome used in nearly all major clinical trials and had as many definitions as trials conducted. The definition of cardiac death had similar variability.

As a result of three Delphi rounds, nine cardiovascular clinical outcomes were defined and clarified, and their 'face' validity was confirmed. These outcomes refer to clinically important, patient-centred cardiac or vascular events. Most of the cardiac outcomes were evaluated by the expert panel as reliable, clearly defined, and usable as endpoints in clinical trials. Importantly, this consensus process found that two of the outcomes evaluated (MINS and congestive heart failure) scored below the *a priori* threshold set for validity, reliability, feasibility, and clarity, and therefore, are not recommended for use in clinical trials until further studies are completed.

Our goal was to identify cardiovascular outcomes that were feasible for use in clinical trials, observational trials, and health services research. Feasibility, however, is a trait that is often influenced by time and financial restraints. For instance, the accurate incidence of myocardial infarction requires protocolised biomarker surveillance and regular postoperative

Table 2 Results of Delphi Rounds 1 and 2. CABG, coronary artery bypass grafting; ENIGMA, Evaluation of Nitrous Oxide in the Gas Mixture for Anaesthesia; hs cTn, highly sensitive cardiac troponin; MACE, major adverse cardiac event; MANAGE, Management of Myocardial Injury After Noncardiac Surgery Trial; N/A, not available; NSQIP, National Surgical Quality Improvement Program; PCI, percutaneous coronary intervention; POISE, Perioperative Ischemic Evaluation Study; VISION, Vascular Events In Noncardiac Surgery Patients Cohort Evaluation Study.

	Delphi Round 1			Delphi Round 2		
	Response rate: 19/19			Response rate: 55/67		
	Median	Range	More than 7 (%)	Median	Range	More than 7 (%)
Myocardial infarction						
Third Universal Definition	8	6–9	74	9	4–9	82
NSQIP definition	5	1–9	26	4	1–9	14
(i) Report the number of patients assayed	8	5–9	74	8	2–9	76
(ii) Requirement to utilise hs cTn assays	8	2–9	63	8	3–9	76
(iii) Requirement to report assay type	6	2–9	32	6	3–9	45
Myocardial injury						
Myocardial injury (from Third Universal Definition)	8	6–9	79	8	1–9	80
Myocardial injury after noncardiac surgery	7	3–9	54	7	3–9	51
(i) Incidence of myocardial infarction should be reported	9	8–9	100	9	5–9	74
(ii) Postoperative ECGs should be performed to assess injury	8	3–9	58	8	1–9	69
(iii) Number of postoperative heart failure should be reported	6	2–9	52	6	1–9	23
(iv) Incidence of chronic kidney disease should be reported	7	5–9	66	7	3–9	61
(v) Atrial fibrillation or other arrhythmias should be reported	7	3–9	45	6	1–9	45
Cardiovascular death						
Death from a cardiovascular cause and including: (i) Myocardial infarction (ii) Cardiac arrest (iii) Cardiac revascularisation procedure (iv) Pulmonary embolism (v) Haemorrhage (vi) Unknown cause (vii) Stroke	7	3–9	58	7	1–9	51
(a) Deaths attributable to pulmonary embolism should be included	7	2–9	47	5	1–9	34
(b) Deaths attributable to haemorrhage should be included	7	2–9	21	2	1–8	7
(c) Deaths where the cause is unknown should be included	3	1–8	21	3	1–8	9
(d) Deaths attributable to stroke should be included	5	1–8	50	6	1–9	34
Non-fatal cardiac arrest						
Successful resuscitation from either documented or presumed ventricular fibrillation, sustained ventricular tachycardia, asystole, or pulseless electrical activity	8	6–9	94	8	4–9	91
Coronary revascularisation						
Cardiac revascularisation procedure was defined as PCI or CABG surgery	8	6–8	94	8	5–9	87
Should this be time limited to within 30 days of surgery?	8	6–8	N/A	8	6–8	N/A
MACE						
POISE (i) Cardiovascular death (ii) Non-fatal myocardial infarction (iii) Non-fatal cardiac arrest at 30 days	7	5–9	63	7	3–9	43
POISE ll (i) Mortality (ii) Non-fatal myocardial infarction (iii) Cardiac revascularisation procedure (iv) Non-fatal pulmonary embolism (v) Non-fatal deep venous thrombosis	5	3–8	21	4	1–8	34
ENIGMA ll (i) Death (ii) Non-fatal myocardial infarction (iii) Cardiac arrest (iv) Pulmonary embolism (v) Stroke during the initial 30 postoperative days	7	3–8	52	6	1–8	38
MANAGE (i) Vascular mortality (ii) Myocardial infarction (iii) Cardiac revascularisation procedure (iv) Non-haemorrhagic stroke	5	1–8	16	4	1–8	7

Continued

Table 2 Continued

	Delphi Round 1			Delphi Round 2		
	Response rate: 19/19			Response rate: 55/67		
	Median	Range	More than 7 (%)	Median	Range	More than 7 (%)
(v) Peripheral arterial thrombosis (vi) Amputation (vii) Symptomatic venous thromboembolism (viii) Rehospitalisation for vascular reasons						
VISION (i) Myocardial infarction (ii) Cardiac arrest, stroke, cardiac revascularisation procedure (iii) Pulmonary embolus (iv) Haemorrhage	4	1–8	16	3	1–8	7
(a) Should thromboembolism be considered as a MACE outcome?	3	1–8	6	2	1–9	2
(b) Should haemorrhage be considered a MACE outcome?	3	1–8	16	2	1–9	2
(c) Should stroke be considered a MACE outcome?	8	1–9	53	7	1–9	53
Pulmonary embolism Required any one of the following: (i) A high probability ventilation/perfusion lung scan (ii) An intraluminal filling defect of segmental or larger artery on a helical CT scan (iii) An intraluminal filling defect on pulmonary angiography (iv) A positive diagnostic test for deep venous thrombosis (e.g. positive compression ultrasound) and one of the following: (a) Non-diagnostic (i.e. low or intermediate probability) ventilation/perfusion lung scan (b) Non-diagnostic (i.e. sub-segmental defects or technically inadequate study) helical CT scan	7	5–9	58	7	5–9	63
Should the lack of routine surveillance for pulmonary embolism be stated as a major limitation?	7	4–9	43	7	5–9	43
Deep venous thrombosis Requires any one of the following: (i) A persistent intraluminal filling defect on contrast venography (ii) Non-compressibility of one or more venous segments on B-mode compression ultrasonography (iii) Clearly defined intraluminal filling defect on contrast enhanced CT	8	5–9	89	8	5–9	80
(a) Should the lack of routine surveillance for postoperative deep venous thrombosis be listed as a major limitation)	8	3–9	55	8	5–9	71
Congestive heart failure Requires at least one of the following clinical signs: (i) An elevated jugular venous pressure (ii) Respiratory rales/crackles and crepitations (iii) Presence of S3 and at least one of the following radiographic findings: (a) Vascular redistribution (b) Interstitial pulmonary oedema (c) Frank alveolar pulmonary oedema	7	4–9	47	7	3–9	40
Atrial fibrillation The occurrence of atrial fibrillation (irregularly irregular heart rate in the absence of P waves) lasting at least 30 s or for the duration of the ECG recording (if <30 s)	7	3–8	58	7	4–9	71
New atrial fibrillation that results in angina, congestive heart failure, or symptomatic hypotension, or that requires treatment with a rate-controlling drug, anti-arrhythmic drug, or electrical cardioversion	7	5–8	58	7	2–9	48
(i) Continuous surveillance should be used to determine this outcome.	6	3–9	43	6	1–9	28
(ii) Lack of continuous ECG monitoring for atrial fibrillation surveillance should be stated as a major limitation	7	3–9	63	7	3–9	

ECG, which results in added costs. Regular biomarker surveillance is not currently a standard of care in most institutions,[16] and ECG was not performed in a third of patients with postoperative troponin elevations.[10] Thus, the true incidence of myocardial infarction is likely biased in many studies, especially in all health services research and prospective surgical improvement databases.[17] Postoperative pulmonary embolism is often asymptomatic[18] and requires

Table 3 Results of the Delphi third round. Validity: the endpoint measures what it purports to measure; reliability: the endpoint is reproducible and is stable when the measurement is repeated under identical conditions; feasibility: the endpoint can be collected and used easily with adequate training, with minimal effort or missing data; clarity of the definition: the endpoint is easily understood.

Response rate: 17/19	Validity			Reliability			Feasibility			Clarity of definition		
	Median	Range	≥7 (%)	Median	Range	≥7 (%)	Median	Range	≥7 (%)	Median	Range	≥7 (%)
Myocardial infarction fourth universal definition	8	6–9	88	8	6–9	82	7	5–9	76	8	5–9	94
Myocardial injury fourth universal	8	6–9	94	8	7–9	100	8	7–9	100	8	7–9	100
Myocardial injury after noncardiac surgery	6	1–8	56	6	1–8	50	6	1–9	44	6	1–9	50
Cardiac death	7	6–8	71	7	5–8	59	7	3–8	53	7	5–9	71
Non-fatal cardiac arrest	8	7–9	100	8	6–9	94	8	7–9	100	8	7–9	100
Coronary revascularisation	8	7–9	100	8	7–9	100	8	7–9	100	8	7–9	100
Major acute cardiac event (i) Myocardial infarction (ii) Non-fatal cardiac arrest (iii) Coronary revascularisation (iv) Cardiac death	8	6–9	94	8	6–8	82	8	6–9	94	8	6–9	94
Pulmonary embolism	8	6–9	94	7	5–9	76	7	5–9	76	8	7–9	100
Deep venous thrombosis	8	6–9	94	8	7–9	100	8	6–8	82	8	7–9	100
Congestive heart failure	6	1–9	23	5	1–9	17	6	1–9	23	6	1–9	47
Atrial fibrillation	8	7–9	100	7	4–9	94	7	6–9	94	8	7–9	100

expensive, time-consuming, and scarce imaging resources for diagnosis, making it infeasible in nearly all clinical trials.

The consensus of our expert panel recommends the use of myocardial injury, as defined by the Fourth Universal Definition of Myocardial Infarction[12] over that of MINS, as defined by Botto and colleagues.[19] The definition and diagnostic criteria of myocardial infarction consist of an elevated biomarker (most often cardiac troponin) and ischaemic symptoms or an ischaemic ECG finding. Perioperatively, myocardial infarction frequently occurs asymptomatically and postoperative ECG is not routine, thus making a missed diagnosis a distinct possibility. It was with this rationale that Botto and colleagues[19] proposed MINS. The original description of MINS showed that postoperative high-sensitivity troponin T (hsTnT) >30 ng L^{-1} was associated with increased cardiac death (defined as deaths after myocardial infarction, cardiac arrest, stroke, cardiac revascularisation procedure, pulmonary embolism, haemorrhage, or death attributable to an unknown cause). In addition, a diagnosis of MINS required *a priori* exclusion of known causes of troponin elevation, such as chronic renal failure, pulmonary embolism, and sepsis. In a second publication, the diagnosis of MINS, after the exclusions listed previously, required an elevated hsTnT >30 ng L^{-1} in patients without a preoperative test, or an increase of 5 ng L^{-1} in patients with preoperative elevated hsTnT. However, after these biomarker thresholds are met, the diagnosis of MINS does not require further evidence of ischaemia (ECG, chest pain, etc.).[20] The alternative, and that endorsed by the consensus panel, is to enumerate both the incidence of myocardial infarction and myocardial injury using the criteria set out in the Fourth Universal Definition of Myocardial Infarction. Myocardial infarction exhibits higher mortality, at 30 days[10] and 1 yr,[4] than myocardial injury alone. Furthermore, myocardial injury, occurring without any other complications, is also prognostically important, albeit less so than myocardial infarction, associated with a two-fold increase in mortality at 1 yr over patients with no postoperative complications.[21]

There are several limitations to our approach to obtaining these consensus recommendations.

First, as the outcomes evaluated were gleaned from a systematic review of previously published perioperative trials, only existing and sometimes out-of-date cardiovascular outcomes could be analysed. For example, the definition of congestive heart failure, used in perioperative trials to date, does not reflect the diagnostic criteria advocated by the two large heart associations. The AHA and ESC guidelines advocate for both measurement of left ventricular function and biomarkers (natriuretic peptides). These new heart failure criteria have not been used for perioperative clinical trials before 2018, pointing to the need for these to be evaluated in the perioperative setting. In addition, the Fourth Universal Definition of Myocardial Infarction was published during our consensus process. Whilst this has not been used in a perioperative trial to date, we have adopted it over the Third Universal Definition.

Second, we evaluated a composite endpoint widely used in perioperative studies. Major adverse cardiovascular event has been used as primary endpoint in every major perioperative cardiovascular trial. Over the time period we reviewed, the number of components used to define MACE has increased. As has been observed by others,[22] increasing the number of outcomes to a composite outcome, which is usually of less clinical importance, is done to increase statistical power. We therefore elected to deconstruct MACE and ask each respondent to assess the individual components of the most recognisable trials. Our final 'recommended definition' was scored as valid, feasible, and clear by 94% and reliable by 82%. The definition includes four major outcomes with significant patient-centred effects: myocardial infarction, non-fatal cardiac arrest, coronary revascularisation within 30 days of surgery, and cardiac death within 30 days of surgery. A second limitation of this composite outcome is the inclusion of cardiac death. The panel rated cardiac death as valid and clear (71%). In contrast, only 59% of the panel found the definition reliable and 53% found it feasible. As the definition of cardiac death varied

Table 4 Consensus recommendations for cardiac outcomes in future perioperative trials. cTn, cardiac troponin; URL, upper reference limit.

Outcome	Definition	Reporting requirements	Consensus rating
Myocardial infarction	Acute myocardial injury with clinical evidence of acute myocardial ischaemia and with detection of an increase or decrease in cTn values with at least one value above the 99th percentile URL and at least one of the following: (i) Symptoms of myocardial ischaemia (ii) New ischaemic ECG changes (iii) Development of pathological Q waves (iv) Imaging evidence of new loss of viable myocardium or new regional wall motion abnormality in a pattern consistent with an ischaemic aetiology (v) Identification of a coronary thrombus by angiography or autopsy *Post-mortem* demonstration of acute atherothrombosis in the artery supplying the infarcted myocardium Cardiac death in patients with symptoms suggestive of myocardial ischaemia and presumed new ischaemic ECG changes before cTn values become available	Report: (i) Type of assay (ii) URL (iii) Number of patients tested	Valid=88% Reliable=82% Feasible=76% Clear=94%
Myocardial injury	Detection of an elevated cTn value above the 99th percentile URL is defined as myocardial injury. The injury is considered acute if there is an increase or decrease in cTn values. Note: it is not clinically possible to distinguish which increases of cTn levels are attributable to which mechanisms. A diagnosis of myocardial infarction requires an increase of cTn values and evidence of myocardial ischaemia that may be evident from the peri- and postoperative period (e.g. ST segment changes on telemetry/ECG, repeated episodes of hypoxia, hypotension, tachycardia, or imaging evidence of myocardial injury). In the absence of evidence for acute myocardial ischaemia, the diagnosis is acute myocardial injury.	Report: (i) Type of assay (ii) URL (iii) Number with myocardial infarction (iv) Number of patients with postoperative cTn tested (v) Postoperative EKG to be obtained in all patients with elevated cTn (vi) Number with chronic renal failure	Valid=94% Reliable=100% Feasible=100% Clear=100%
Cardiac death	Death with a vascular cause and included those deaths after a myocardial infarction, cardiac arrest, and cardiac revascularisation procedure.	Excludes (i) Death after pulmonary embolism (ii) Death after haemorrhage (iii) Death after multi-organ failure (iv) Cause of deaths unknown	Valid=71% Reliable=59% Feasible=53% Clear=71%
Non-fatal cardiac arrest	Successful resuscitation from either documented or presumed ventricular fibrillation, sustained ventricular tachycardia, asystole, or pulseless electrical activity requiring cardiopulmonary resuscitation, pharmacological therapy, or cardiac defibrillation		Valid=100 Reliable=94% Feasible=100% Clear=100%

Continued

Table 4 Continued

Outcome	Definition	Reporting requirements	Consensus rating
Coronary revascularisation	Cardiac revascularisation procedure was defined as percutaneous coronary intervention or coronary artery bypass graft surgery within 30 days of the index surgery.		Valid=100% Reliable=100% Feasible=100% Clear=100%
Major adverse cardiac event	Is composite outcome that should include (i) Cardiac death (as defined previously) (ii) Myocardial infarction (as defined previously) (iii) Non-fatal cardiac arrest (as defined previously) (iv) Coronary revascularisation (as defined previously) within 30 days of the index surgery	Excludes: (i) Pulmonary embolism (ii) Haemorrhage (iii) Deep venous thrombosis (iv) All-cause mortality	Valid=94% Reliable=82% Feasible=94% Clear=94%
Pulmonary embolism	Diagnosis of pulmonary embolism requires any one of the following: (i) A high probability ventilation/perfusion lung scan (ii) An intraluminal filling defect of segmental or larger artery on a helical CT scan (iii) An intraluminal filling defect on pulmonary angiography (iv) A positive diagnostic test for deep venous thrombosis (e.g. positive compression ultrasound) and one of the following: (a) Non-diagnostic (i.e. low or intermediate probability) ventilation/perfusion lung scan (b) Non-diagnostic (i.e. sub-segmental defects or technically inadequate study) helical CT scan	As postoperative pulmonary embolism is often asymptomatic, a lack of routine surveillance for pulmonary embolism should be stated as a major limitation.	Valid=94% Reliable=76% Feasible=76% Clear=100%
Deep venous thrombosis	Diagnosis of deep venous thrombosis required any one of the following: (i) A persistent intraluminal filling defect on contrast venography (ii) Non-compressibility of one or more venous segments on B-mode compression ultrasonography (iii) A clearly defined intraluminal filling defect on contrast enhanced CT	As many of deep venous thromboses are asymptomatic, the lack of routine surveillance is a major limitation and should be stated.	Valid=94% Reliable=100% Feasible=82% Clear=100%
Atrial fibrillation	New onset of irregularly irregular heart rate in the absence of P waves lasting at least 30 s or for the duration of the ECG recording (if <30 s)	As atrial fibrillation can occur asymptomatically, the lack of continuous EKG monitoring for atrial fibrillation surveillance should be stated as a major limitation.	Valid=100% Reliable=94% Feasible=94% Clear=100%

between studies, we again deconstructed it and again asked the panel to evaluate each component separately. There was consensus agreement that the definition of cardiac death should not include deaths from (i) unknown causes, (ii) haemorrhage, (iii) pulmonary embolism, or (iv) multi-organ failure.

Third, because of the elapsed time to complete three Delphi consensus-seeking rounds, the search may be considered out of date, last updated December 31, 2017. Therefore, it is possible that new definitions had been advanced that our process had not considered. We therefore have undertaken a *post hoc* search of anaesthesia and general medical journals from 2018 to present. This modified search (seen in Supplementary material) revealed 17 studies that satisfied our inclusion criteria. None of these studies utilised unique or new cardiovascular endpoints.

Finally, as for any diagnostic tool in medicine, the specificity rarely reaches 100%. In our study, only coronary revascularisation achieved 100% consensus for validity, reliability, feasibility, and clarity of definition. However, in most cases,

the cardiovascular outcome definitions outlined here achieved a high degree of agreement amongst our expert panel, and we therefore submit that these outcomes can be used confidently as endpoints in clinical trials. An important caveat is that most cardiovascular outcomes require active, protocolised surveillance for accurate reporting. As routine surveillance for these outcomes is not currently the standard of clinical care, the outcomes in this report will have limited use in health services and observational research.

Despite these limitations, we identified a number of standardised cardiovascular endpoints to be used in future studies assessing effectiveness of perioperative interventions. This study provides guidance to achieve consistency in future perioperative clinical investigation, resulting in improved interpretation of study results and translation into clinical practice.

StEP Steering Committee members

Paul Myles (Australia) and Michael Grocott (UK), Co-Chairs; **Bruce Biccard (South Africa)**, Jane Blazeby (UK), Oliver Boney (UK), Matthew Chan (Hong Kong), Elisabeth Diouf (Senegal), **Lee Fleisher (USA),** Cor Kalkman (The Netherlands), **Andrea Kurz (USA),** Ramani Moonesinghe (UK), and **Duminda Wijeysundera (Canada).** *Members of the Cardiovascular sub-group are in* ***BOLD****.*

Authors' contributions

Study concept: WSB, PSM
Protocol development: WSB, DNW
Systematic review: ML, MB, SF, WSB
Analysis of responses: WSB, ML
Drafting of paper: WSB
Critical review/revisions of paper: KL, PSM, SH, GL, BB, PN
All members of the cardiovascular subgroup participated in all Delphi rounds. All members of the Steering Committee and the cardiovascular subgroup approved the final version of this paper.

Acknowledgements

We would like to acknowledge the work of the data extractors who are listed in the appendix. Finally, we would like to thank Sophia Wallace for her tireless work on coordinating the DELPHI rounds, communications, and collation of the results.

Declarations of interest

The authors declare that they have no conflicts of interest.

Funding

The R. Fraser Chair in Cardiovascular Anesthesiology endowment.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bja.2020.09.023.

References

1. Pearse RM, Fang X, Holt B, Kahan BC, Rhodes A, Wijeysundera DN. Critical care resource provision and patient outcomes after elective surgery: prospective cohort study in 27 countries. *Br J Anaesth* 2016; **117**: 601–9
2. Devereaux PJ, Yang H, Yusuf S, et al. Effects of extended-release metoprolol succinate in patients undergoing non-cardiac surgery (POISE trial): a randomised controlled trial. *Lancet* 2008; **371**: 1839–47
3. Devereaux PJ, Mrkobrada M, Sessler DI, et al. Aspirin in patients undergoing noncardiac surgery. *N Engl J Med* 2014; **370**: 1494–503
4. Beattie WS, Wijeysundera DN, Chan MTV, et al. Implication of major adverse postoperative events and myocardial injury on disability and survival: a planned subanalysis of the ENIGMA-II trial. *Anesth Analg* 2018; **127**: 1118–26
5. Myles PS, Boney O, Botti M, et al. Systematic review and consensus definitions for the Standardised Endpoints in Perioperative Medicine (StEP) initiative: patient comfort. *Br J Anaesth* 2018; **120**: 705–11
6. Haller G, Bampoe S, Cook T, et al. Systematic review and consensus definitions for the Standardised Endpoints in Perioperative Medicine initiative: clinical indicators. *Br J Anaesth* 2019; **123**: 228–37
7. Moonesinghe SR, Jackson AIR, Boney O, et al. Systematic review and consensus definitions for the Standardised Endpoints in Perioperative Medicine initiative: patient-centred outcomes. *Br J Anaesth* 2019; **123**: 664–70
8. Graham B, Regehr G, Wright JG. Delphi as a method to establish consensus for diagnostic criteria. *J Clin Epidemiol* 2003; **56**: 1150–6
9. Jones J, Hunter D. Consensus methods for medical and health services research. *BMJ* 1995; **311**: 376–80
10. Writing Committee for the VISION Study Investigators, Devereaux PJ, Biccard BM, Sigamani A. Association of postoperative high-sensitivity troponin levels with myocardial injury and 30-day mortality among patients undergoing noncardiac surgery. *JAMA* 2017; **317**: 1642–51
11. Devereaux PJ, Duceppe E, Guyatt G, et al. Dabigatran in patients with myocardial injury after non-cardiac surgery (MANAGE): an international, randomised, placebo-controlled trial. *Lancet* 2018; **391**: 2325–34
12. Thygesen K, Alpert JS, Jaffe AS, et al. Fourth universal definition of myocardial infarction. *Circulation* 2018; **2018**(138): e618–51
13. Yancy CW, Jessup M, Bozkurt B, et al. 2017 ACC/AHA/HFSA focused update of the 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Failure Society of America. *J Card Fail* 2017; **23**: 628–51
14. Ponikowski P, Voors AA, Anker SD, et al. 2016 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure. *Rev Esp Cardiol* 2016; **69**: 1167 (Engl Ed)
15. Thygesen K, Alpert JS, Jaffe AS, Simoons ML, Chaitman BR, White HD. Third universal definition of myocardial infarction. *Circulation* 2012; **126**: 2020–35

16. Duceppe E, Parlow J, MacDonald P, et al. Canadian Cardiovascular Society guidelines on perioperative cardiac risk assessment and management for patients who undergo noncardiac surgery. *Can J Cardiol* 2017; **33**: 17–32
17. Gupta PK, Gupta H, Sundaram A, et al. Development and validation of a risk calculator for prediction of cardiac risk after surgery. *Circulation* 2011; **124**: 381–7
18. Grobben RB, van Waes JAR, Leiner T, et al. Unexpected cardiac computed tomography findings in patients with postoperative myocardial injury. *Anesth Analg* 2018; **126**: 1462–8
19. Botto F, Alonso-Coello P, Chan MT, et al. Myocardial injury after noncardiac surgery: a large, international, prospective cohort study establishing diagnostic criteria, characteristics, predictors, and 30-day outcomes. *Anesthesiology* 2014; **120**: 564–78
20. Vascular Events In Noncardiac Surgery Patients Cohort Evaluation Study Investigators, Devereaux PJ, Chan MT, Aloso-Coello A. Association between postoperative troponin levels and 30-day mortality among patients undergoing noncardiac surgery. *JAMA* 2012; **307**: 2295–304
21. Beattie WS, Wijeysundera DN, Chan MTV, et al. Survival after isolated post-operative troponin elevation. *J Am Coll Cardiol* 2017; **70**: 907–8
22. Tan NS, Ali SH, Lebovic G, Mamdani M, Laupacis A, Yan AT. Temporal trends in use of composite end points in major cardiovascular randomized clinical trials in prominent medical journals. *Circ Cardiovasc Qual Outcomes* 2017: 10. e003753

Handling editor: Hugh C Hemmings Jr