Systematic review and consensus definitions for the Standardised Endpoints in Perioperative Medicine (StEP) initiative: infection and sepsis

Barnes, Jonathan; Hunter, Jennifer; Harris, Steve; Shankar-Hari, Manu; Diouf, Elisabeth; Jammer, Ib; Kalkman, Cor; Klein, Andrew A.; Corcoran, Tomas; Dieleman, Stefan; Grocott, Michael P.W.; Mythen, Michael G.

Collaborators: Tramer, Martin; Haller, Guy Serge Antoine

How to cite

BARNES, Jonathan et al. Systematic review and consensus definitions for the Standardised Endpoints in Perioperative Medicine (StEP) initiative: infection and sepsis. In: British journal of anaesthesia, 2019, vol. 122, n° 4, p. 500–508. doi: 10.1016/j.bja.2019.01.009

This publication URL: https://archive-ouverte.unige.ch/unige:158512
Publication DOI: 10.1016/j.bja.2019.01.009

BJA

British Journal of Anaesthesia, 122 (4): 500–508 (2019)

doi: 10.1016/j.bja.2019.01.009
Advance Access Publication Date: 14 February 2019
Review Article

QUALITY AND PATIENT SAFETY

Systematic review and consensus definitions for the Standardised Endpoints in Perioperative Medicine (StEP) initiative: infection and sepsis

Jonathan Barnes[1,#], Jennifer Hunter[2,#], Steve Harris[3,*,#], Manu Shankar-Hari[4,5], Elisabeth Diouf[6], Ib Jammer[7,8], Cor Kalkman[9], Andrew A. Klein[10], Tomas Corcoran[11,12], Stefan Dieleman[13], Michael P. W. Grocott[14,15,16], Michael G. Mythen[2,17], and the StEP-COMPAC group[18]

[1]Department of Anaesthesia, Great Western Hospital, Swindon, UK, [2]Department of Anaesthesia and Perioperative Medicine, University College Hospital, London, UK, [3]Bloomsbury Institute of Intensive Care Medicine, University College London Hospital, London, UK, [4]School of Immunology & Microbial Sciences, Kings College London, UK, [5]Guy's and St Thomas' NHS Foundation Trust, London, UK, [6]Anesthesia and Intensive Care Unit, Le Dantec Hospital, Cheikh Anta DIOP University, Dakar, Senegal, [7]Department of Anaesthesiology and Intensive Care, Haukeland University Hospital, Bergen, Norway, [8]Department of Clinical Medicine, University of Bergen, Bergen, Norway, [9]Division of Anesthesiology, Intensive Care and Emergency Medicine, University Medical Center Utrecht, Utrecht, The Netherlands, [10]Department of Anaesthesia and Intensive Care, Royal Papworth Hospital, Cambridgeshire, UK, [11]Department of Anaesthesia and Pain Medicine, Royal Perth Hospital, Perth, Australia, [12]School of Medicine and Pharmacology, University of Western Australia, Perth, Australia, [13]Department of Anaesthesia & Perioperative Medicine Westmead Hospital, Sydney, Australia, [14]Anaesthesia and Critical Care Research Group, NIHR Biomedical Research Centre, UK, [15]Faculty of Medicine, University of Southampton, Southampton, UK, [16]Department of Anaesthesiology, Duke University School of Medicine, Durham, NC, USA and [17]National Institute of Health Research Biomedical Research Centre, London, UK

*Corresponding author. E-mail: doc@steveharris.me

[18]The StEP-COMPAC group collaborators are listed in Supplementary section.
[#]J. Barnes, J. Hunter, and S. Harris share first authorship of this work.

This article is accompanied by an editorial: Patient-important outcomes and core outcome sets - increased attention needed! by M.H. Møller, *Br J Anaesth* 2019:122:408–410, doi: https://doi.org/10.1016/j.bja.2019.02.007.

Editorial decision: 08 January 2019; **Accepted:** 8 January 2019

Abstract

Background: Perioperative infection and sepsis are of fundamental concern to perioperative clinicians. However, standardised endpoints are either poorly defined or not routinely implemented. The Standardised Endpoints in Perioperative Medicine (StEP) initiative was established to derive a set of standardised endpoints for use in perioperative clinical trials.

Methods: We undertook a systematic review to identify measures of infection and sepsis used in the perioperative literature. A multi-round Delphi consensus process that included more than 60 clinician researchers was then used to refine a recommended list of outcome measures.

Results: A literature search yielded 1857 titles of which 255 met inclusion criteria for endpoint extraction. A long list of endpoints, with definitions and timescales, was generated and those potentially relevant to infection and sepsis circulated to the theme subgroup and then the wider StEP-COMPAC working group, undergoing a three-stage Delphi process. The response rates for Delphi rounds 1, 3, and 3 were 89% (n=8), 67% (n=62), and 80% (n=8), respectively. A set of 13 endpoints including fever, surgical site, and organ-specific infections as defined by the US Centres for Disease Control and Sepsis-3 are proposed for future use.

Conclusions: We defined a consensus list of standardised endpoints related to infection and sepsis for perioperative trials using an established and rigorous approach. Each endpoint was evaluated with respect to validity, reliability, feasibility, and patient centredness. One or more of these should be considered for inclusion in future perioperative clinical trials assessing infection, sepsis, or both, thereby permitting synthesis and comparison of future results.

Keywords: core outcome measures; infection; perioperative medicine; postoperative outcome; sepsis; standardised endpoints; surgical site infection

Editor's key points

- The Standardised Endpoints in Perioperative Medicine (StEP) initiative was established to derive standardised endpoints for use in perioperative clinical trials.
- After a systematic review and Delphi consensus process, a set of 13 outcome measures were identified that should be considered in designing perioperative clinical trials.
- Use and reporting of these endpoints will support improved benchmarking and meta-analysis of future perioperative trials involving infection and sepsis.

The management of infection in the perioperative setting has been of concern since the time of Semmelweis' *Open Letter to all Professors of Obstetrics in 1862*.[1] However, 'varied definitions and inconsistent reporting of outcomes across trials…limit the value of…research' to combat this problem.[2] The Core Outcome Measures in Effectiveness Trials (COMET) Initiative was founded in 2010 to develop an international repository of standardised outcomes known as a 'core outcome measures' (COM[3]) that should represent the minimum required endpoints to be collected and reported.[4] Clinical trialists are now establishing COMs for their specific domains. The process for this has been standardised in perioperative medicine using a defined consensus process that will generate Standardised Endpoints and COMs for Perioperative and Anaesthetic Care (StEP-COMPAC).[2]

The international consensus definition of sepsis has recently been updated in Sepsis-3 to incorporate the Sequential Organ Failure Assessment (SOFA) score, with one aim being providing greater consistency for clinical trials,[5] whereas the US Centers for Disease Control (CDC) definitions of surgical site infections (SSI) have been widely used since their publication in 1992.[5,6] Although robust definitions relating to aspects of perioperative infection exist, their utility in perioperative medicine trials is yet to be evaluated. Aspects of validity, reliability, and practicality need to be considered in assessing the suitability of these and other endpoints for use in this area.

The overall aim of the Standardised Endpoints in Perioperative Medicine (StEP-COMPAC) initiative is to derive a set of standardised outcomes for use in perioperative medicine trials based on current evidence, expert guidance, and international consensus.[2] Here, we describe the results of a systematic review and Delphi process to recommend existing definitions or identify reliable, valid, and feasible outcomes for use in trials around infection and sepsis in the perioperative setting.

Methods

We performed a systematic review to identify outcome measures related to infection or sepsis reported in studies of the perioperative period. We defined the perioperative period as that from surgical planning to full recovery, and broadly defined sepsis and infection using modifications of the search strategies used for Sepsis-3.[7] A Delphi process was then undertaken to refine a consensus list of outcomes to be recommended for use in future work. We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines for systematic reviews[8] (Supplementary Table S3) where possible to ensure that our work meets agreed standards.

Literature search

Systematic searches were undertaken of the Medline, Embase, and Cochrane Databases (see ESM Section 1: Search Strategy) for studies in anaesthetic, surgical, and perioperative care related to systemic infection, sepsis, and the systemic inflammatory response syndrome (SIRS), as well as wound

infection and specific organ/system infection. We included randomised controlled trials, observational studies, consensus statements and guidelines, and meta-analyses from high-quality journals (Abridged Index Medicus)[9] in both adult and paediatric populations published after 2011. We excluded studies with a sample population <200, review articles (as these consistently did not include objective endpoints), studies not related to the perioperative period, studies of non-surgical procedures (e.g. central line insertion), and studies specifically in neonates.

Articles first underwent title review to exclude those not related to the perioperative period, and then abstract review to exclude those not meeting the above criteria. Two authors (JB, JH), independently assessed the studies using the predefined inclusion criteria listed above. Any differences between the two reviewers were settled by a third reviewer (SH) with full agreement of the two primary reviewers. After abstract review and exclusions, the remaining articles underwent full data extraction, where all endpoints (primary and secondary) were listed for each paper, including follow-up periods and definitions used. Fig. 1 shows a PRISMA flowchart for this process.

Definitions of criteria used for judging endpoints

Based on the definitions previously used for judging septic shock outcomes[10] and those suggested in StEP Delphi Methods guidelines,[11] the following definitions for characteristics used to judge outcomes were used throughout the Delphi process.

Validity is the ability to capture what the investigator seeks to measure.

Reliability is the agreement between observers and by the same observer during repeated measurements, that is consistency and reproducibility.

Feasibility is a composite concept that depends on the purpose of the definition—a compromise between validity and reliability.

Patient centredness refers to whether the endpoint has a meaningful impact on a patient's recovery, for example their discomfort or distress, length of hospital stay, need for reoperation, risk of ongoing disability, or increased risk of death.

Delphi process

A Delphi process was used to curate a list of endpoints reflecting the consensus of the StEP Sepsis subgroup, comprising the four stages below. The StEP working group comprised an international group of experienced perioperative trialists (see Supplementary data) and was overseen by a Steering Committee (see Appendix).

The Delphi process was run for three rounds, and each round was coordinated through the Health Services Research Centre of the Royal College of Anaesthetists, UK, and the Department of Anaesthesia and Perioperative Medicine at University College Hospital, London, UK. For each round, respondents' scores and comments were tabulated using Excel (Microsoft Corporation, Redmond, WA, USA). Median scores and centiles were again generated using Excel.

Extracting potential trial endpoints and definitions

This refers to extraction of outcomes from articles, including definitions, time periods, and whether primary or secondary. Frequency of endpoints and these characteristics were tabulated, and suggested endpoints with definitions and time periods were developed by the authors based on those most frequently occurring.

Delphi round 1 (theme subgroup): formal rating of the recommendations

A long list of endpoints from outcome generation were RAG (Red–Amber–Green) rated based on their relevance to infection and inflammation in perioperative medicine by two researchers (JB and JH): green (definitely relevant), amber (possibly relevant), and red (not relevant). The long list, including definitions and time periods, was circulated to all members of the StEP theme subgroup (n=10). Participants were first asked whether they were in agreement with the RAG rating. On the basis that they agreed with this rating, they were then asked to rate each green endpoint using a score from 1 to 9 based on increasing importance for inclusion in a final list of endpoints, with a score of 10 for any outcome that a member of the subgroup was unsure about. The group were also invited to suggest endpoints for inclusion not generated by outcome extraction for articles, and amendments to definitions.

Delphi round 2 (full StEP working group)

The mean, median, and range of scores from round 1 was calculated, and any endpoints with a 70th centile score ≥7, median score of ≥7, or considered important for inclusion by the subgroup, were shortlisted for a second round. Any items with a median score <3 were rejected. The revised shortlist was then sent to the entire StEP working group, along with a summary of round one scores, and the Delphi process repeated. For each endpoint, participants gave an individual score of 1–9 against each of the criteria of validity, reliability, feasibility, and patient centredness (higher scores were better, with a score of 10 used to indicate uncertainty). Participants could again return comments on individual endpoints to suggest amendments to definitions.

Delphi round 3 (theme subgroup): final round and recommendations

Median scores and 70th centile scores from round 2 were generated and circulated to the theme subgroup for further discussion. Based on the feedback from the two previous rounds the subgroup was again asked to rank each endpoint a third time, from 1 to 9 based on increasing importance for inclusion, with 10 indicating unsure. Consensus was defined as a mean score of ≥7 and a 70th centile score ≥7 based on the round 3 scores, and endpoints meeting these criteria were put forward as the recommended endpoints for the infection and sepsis theme.

Results

A total of 1857 articles were retrieved after duplicate removal. After title review, 601 underwent abstract or full text review. After exclusions based on these, 255 underwent endpoint extraction (see Fig. 1). The endpoints identified fell into three principal categories: those related to SSI; those related to sepsis (organ specific or otherwise); and those covered by the other themes of the StEP working group (e.g. mortality, organ dysfunction).

The initial long list of potential trial endpoints is presented in Supplementary Table S1 alongside their frequency as either a primary or secondary endpoint.

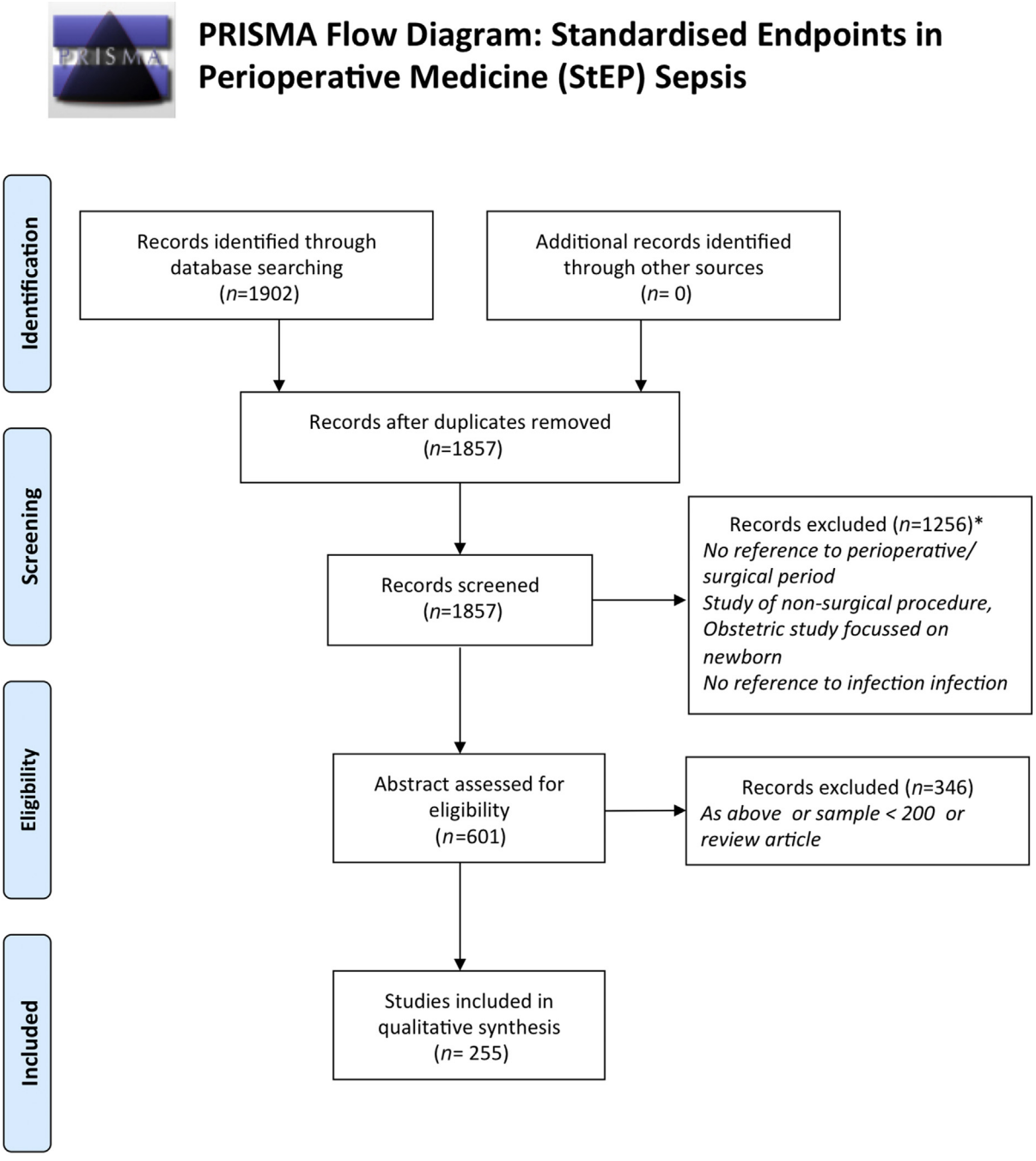

Fig 1. PRISMA diagram showing inclusion and exclusion process for literature search and review. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

The response rates for Delphi rounds 1, 2, and 3 were 89% (n=8), 67% (n=62), and 80% (n=8), respectively.

In Delphi round 1 the subgroup accepted the recommendation that those rated as not definitely relevant to the infection and inflammation theme were immediately discarded. The remaining endpoints were then ranked. Endpoints considered important for inclusion but not present in the initial long list were inserted and ranked.

In Delphi round 2 the endpoints and definitions presented in Supplementary Table S2 were circulated, along with the results of round 1. No endpoints were discarded for Delphi round 2. Table 1 summarises the results of Delphi round 2, with scores indicating the group's assessment of validity, reliability, feasibility, and patient centredness.

In Delphi round 3, the results of round 2 plus detailed comments and suggestions from the full working group were again presented to the theme subgroup. Thirteen endpoints meeting the criteria of a median Delphi round 3 score of ≥7 and a 70th centile score ≥7 are shown in Table 2 with the agreed on definitions, representing the final recommended outcomes regarding infection and sepsis for use in perioperative medicine trials.

Table 3 summarises the results of Delphi rounds 1 and 3, with scores indicating how critical for inclusion each endpoint was considered by the subgroup at each round.

Discussion

We applied the methodology developed by the StEP-COMPAC group[6] to define a core outcome set for infection-related endpoints measured in perioperative trials. This comprised a systematic review of the literature followed by a Delphi process involving perioperative medicine experts to reach consensus. Based on this Delphi process, we recommend the use of the following outcomes (full recommended definitions are given in Table 2).

General markers of infection and inflammation

In the perioperative period, markers of infection overlap significantly with markers of inflammation, thus reducing their diagnostic specificity. Although magnitude of inflammation may be of interest perioperatively, the presence or absence of infection was the focus of the group, and these markers had lower ratings with respect to validity and reliability. All other recommended endpoints relate to proven infection.

Fever was the only general marker of infection and sepsis meeting the criteria for inclusion in the final round. Fever was not clearly or uniformly defined in the literature; in most cases no definition was given, and a range of definitions are used in studies regarding postoperative fever.[12–14] The proposed

Table 1 Summary of Delphi round 2 results, with median and percentage of score ≥7 for each criterion of validity, reliability, feasibility, and patient centredness. CRP, C-reactive protein; SIRS, systemic inflammatory response syndrome; WBC, white blood cell count

Summary of item	Validity			Reliability			Feasibility			Patient centredness		
	Unsure	Scores ≥7 (%)	Median	Unsure	Scores ≥7 (%)	Median	Unsure	Scores ≥7 (%)	Median	Unsure	Scores ≥7 (%)	Median
1a. Fever	3	78	7	3	68	7	3	98	8	3	66	7
1b. CRP measure	3	42	6	3	27	6	3	73	7	3	24	5
1c. Procalcitonin and other biomarkers	5	26	5	5	18	6	5	42	6	5	16	4
1d. WBC/pattern	4	62	7	4	60	7	4	95	8	4	36	6
1e. Antibiotic use	3	53	7	4	43	6	3	86	8	3	49	6
2a. Respiratory infectious complication	4	86	8	4	71	7	4	85	7	4	81	7
2b. Neurological infectious complication	4	90	8	4	85	8	4	72	7	4	79	8
2c. Urological/ genitourinary complication	4	91.	8	4	86	7.5	4	88	8	4	86	8
2d. *Clostridium difficile* colitis or infection	4	88	8	4	86	8	4	83	8	4	79	8
2e. Endometritis	4	78	7	4	66	7	4	74	7	4	83	7
3. Identification of pathogenic organism from culture	5	86	8	5	81	7	5	88	7	5	54	7
4a. Surgical site infection	5	89	8	5	80	7	5	91	8	5	82	8
4b. Superficial incisional surgical site infection	4	88	8	4	77	7	4	88	8	4	84	8
4c. Deep incisional surgical site infection	4	93	8	4	88	8	4	90	8	4	91	8
4d. Organ/space surgical site infection	4	91	8	4	86	8	4	91	8	4	93	8
5a. Sepsis	4	86	8	4	77	7	4	84	7	4	71	7
5b. Septic shock	4	91	8	4	79	8	4	86	8	4	79	8
5c. SIRS	4	65	7	4	63	7	4	84	7	4	56	7

Table 2 Endpoints and definitions proposed for infection and inflammation theme. *Note any deviation from CDC definition is attributable to paraphrasing for ease of presenting in results table. †Note as per the CDC definitions there is a group of operations where a 90 day duration is used for this endpoint, please consult Supplementary Table S4.The exact CDC definition is proposed, and these are given fully in the appendix. ‡Definition as per the Third International Consensus Definitions for Sepsis and Septic Shock (SEPSIS-3). CDC, US Centers for Disease Control; CRP, C-reactive protein; SOFA, Sequential Organ Failure Assessment; WBC, white blood cell count

Endpoint	Proposed definition
Fever	▪ Core body temperature >38.5°C more than 24 h after operation, and two readings in a 12 h period
Respiratory Infectious Complication—CDC definition*	▪ *Signs/Symptoms/Laboratory*: ONE of [fever >38.0°C, WBC <4×10^9 or >12×10^9 L^{-1}, altered mental status in >70 yr old with no other recognised cause] and TWO of [new onset purulent sputum/change in character of sputum/increased respiratory secretions or increased suctioning OR worsening cough/dyspnoea/tachypnoea OR rales or bronchial breath sounds OR worsening gas exchange] ▪ *Imaging*: two or more serial chest imaging results with either [new and persistent OR progressive and persistent] changes of [infiltrate OR consolidation OR cavitation] ▪ *OR one of*: - Organism seen on Gram stain of lung tissue or pleural fluid, or identification of pathogenic organism from fluid or tissue from affected site (see outcome 3) - Abscess or other evidence of infection on gross anatomical or histopathologic examination - Imaging test evidence of abscess or other infection which if equivocal is supported by clinical correlation, specifically, physician documentation of antimicrobial treatment for lung infection
Neurological Infectious Complication—CDC definition*	▪ ONE of: - Identification of pathogenic organism from fluid or tissue from affected site (see outcome 3) - Abscess or evidence of intracranial infection on gross or anatomical or histopathologic examination ▪ TWO of [headache, dizziness, fever >38°C, localising neurological signs, changing level of consciousness, confusion] and ONE of [organism seen on microscopic examination of brain or abscess tissue obtained by needle aspiration or during invasive procedure or autopsy OR imaging suggestive of infection which is equivocal is supported by clinical correlation OR diagnostic single antibody titre or 4-fold increase in paired sera for organism] ▪ TWO of [fever >38°C or headache, meningeal signs, cranial signs] AND ONE OF: [increased WBC elevated protein and low glucose in CSF OR organism on Gram stain of CSF OR organism identified on blood culture, diagnostic antibody titre or 4-fold increase in paired sera for organism] ▪ Within 30 days[2]
Urinary System Infectious Complication—CDC definition[1]*	▪ ONE of: - Identification of pathogenic organism from fluid or tissue from affected site (see outcome 3) - Abscess or other evidence of infection on gross anatomical examination, during invasive procedure, or during histopathologic examination - ONE of [Fever >38°C, localised pain or tenderness with no other recognised cause] AND ONE OF [purulent drainage from affected site OR organism identified in blood by culture or non-culture based biological testing OR imaging suggestive of infection which if equivocal is supported by clinical correlation, specifically physician documented treatment for urinary system infection] ▪ Within 30 days[2]
Clostridium difficile Colitis/Infection—CDC definition*	▪ ONE of: - Positive test for toxin-producing *C. difficile* on an unformed stool sample - Patient has evidence of pseudomembranous colitis on gross anatomical or histopathologic examination ▪ Within 30 days[2]
Endometritis—CDC definition*	▪ ONE of: - Identification of pathogenic organism from fluid or tissue from affected site (see outcome 3) - TWO of [fever >38.0°C, uterine or abdominal pain or tenderness with no other recognised cause, purulent drainage from the uterus] ▪ Within 30 days[2]
Identification of Pathogenic Organism from Tissue or Fluid	▪ Organisms identified from aseptically obtained fluid or tissue in the organ space by a culture or non-culture based testing method which is performed for the purposes of clinical diagnosis and treatment ▪ Within 30 days[2]

Continued

Table 2 Continued

Endpoint	Proposed definition
Surgical Site Infection	▪ Total of superficial and deep SSIs
Superficial Incisional Surgical Site Infection—CDC definition*	▪ Involves only skin and subcutaneous tissue of the incision AND at least one of [purulent discharge from superficial incision or subcutaneous tissue OR organisms identified from specimen from superficial incision OR superficial incision deliberately opened by surgeon/treating physician with one of pain, tenderness, localised swelling, erythema, or heat OR diagnosis of superficial SSI by surgeon/treating physician] ▪ Within 30 days[2]
Deep Incisional Surgical Site Infection—CDC definition*	▪ Involves deep soft tissues of incision (fascial and muscle layers) AND at least one of [purulent discharge from deep incision OR deep incision that spontaneously dehisces or is deliberately opened or aspirated by treating surgeon/treating physician and has microorganism identified on microbiological testing with either fever >38.0°C or localised pain or tenderness OR an abscess or other evidence of infection involving the deep incision is identified on gross anatomical examination or imaging] ▪ Within 30 days†
Organ/Space Surgical Site Infection—CDC definition*	▪ Infection involves a part of the body deeper than the fascial/muscle layers, that is opened or manipulated during the operative procedure AND one of [purulent drainage from drain in organ/space OR organisms identified from aseptically obtained fluid or tissue in the organ space by a culture or non-culture based testing method which is performed for the purposes of clinical diagnosis and treatment OR an abscess or other evidence of infection involving the organ/space that is detected on gross histopathologic examination OR imaging test evidence suggestive of infection] ▪ Within 30 days†
Sepsis‡	▪ Increase in SOFA score of 2 or more, with evidence of infection ▪ Within 30 days†
Septic Shock‡	▪ Sepsis (SOFA score or 2 or more with evidence of infection) with shock ▪ Vasopressor requirement to maintain a mean arterial pressure of 65 mm Hg or greater and serum lactate level greater than 2 mM (>18 mg dl^{-1}) in the absence of hypovolaemia

consensus definition is a core body temperature >38.5°C more than 24 h after operation, with two readings within a 12 h period.[14] Studies of postoperative fever suggest that fever in the first 24 h is rarely infection-related,[14–17] with a wide range of aetiologies.[18] The temperature threshold and time periods included in the definition of fever were deliberately selected to focus on fever related to infection, so use a higher temperature than the CDC definitions which rely on other indicators of infection.

Our review also highlighted C-reactive protein (CRP) and white blood cell count (WBC) as potential endpoints, and antibiotic use and procalcitonin were suggested by the subgroup members for consideration in the first round. However, these outcomes were often rated poorly for validity and reliability by both the wider study group (in round 2) and the theme subgroup (in round 3). Although clinical teams continue to use these inflammatory markers, particularly CRP, to guide antibiotic therapy,[19,20] there is little evidence to

Table 3 Summary of Delphi rounds 1 and 3 results. Scores highlighted in italics indicate both scores for that endpoint meet criteria to be proposed as final endpoints. CRP, C-reactive protein; SIRS, systemic inflammatory response syndrome; WBC, white blood cell count

Summary of item	Delphi round 1 (*n*=8)			Delphi round 3 (*n*=8)		
	Unsure	Median score	Scores ≥7 (%)	Unsure	Median score	Scores ≥7 (%)
1a. Fever	0	7.5	63	0	8.0	88
1b. CRP measure	0	6.0	25	0	6.0	38
1c. Procalcitonin and other biomarkers	0	6.0	0	0	4.0	13
1d. WBC/pattern	0	6.5	50	0	7.0	63
1e. Antibiotic use	0	8.0	88	0	7.5	63
2a. Respiratory infectious complication	0	8.0	100	0	8.0	*100*
2b. Neurological infectious complication	0	8.0	100	0	8.0	*100*
2c. Urological/genitourinary complication	0	8.0	100	0	8.0	*100*
2d. *Clostridium difficile* colitis or infection	1	8.0	63	0	8.0	*100*
2e. Endometritis	0	8.0	88	0	8.0	*100*
3. Identification of pathogenic organism from culture	0	8.0	75	0	8.0	*100*
4a. Surgical site infection	0	8.0	88	0	8.0	*100*
4b. Superficial incisional surgical site infection	0	8.5	100	0	8.0	*100*
4c. Deep incisional surgical site infection	0	8.5	100	0	8.0	*100*
4d. Organ/space surgical site infection	0	8.5	100	0	8.0	*100*
5a. Sepsis	0	8.0	75	0	8.0	*100*
5b. Septic shock	0	8.0	75	0	8.0	*100*
5c. SIRS				0	7.0	*63*

suggest they reliably differentiate inflammation from infection.[21–24]

Specific organ system infection

The systematic review exposed a broad range of definitions in use for organ-specific infection. However, the CDC definitions were most commonly used, and have been recommended as standardised endpoints.[25] This is consistent with the subgroup evaluating pulmonary outcomes, that also recommended the CDC definition of respiratory infection.[26]

Endometritis and neurological system infections are likely to be relevant only to specific surgical specialities. However, respiratory, urological, and *Clostridium difficile* infections will be of wider interest regardless of primary procedure, and their prevalence may be indicative of general quality of perioperative care.[27]

Microbiological markers of infection

Identification of pathogens from tissue or fluid was an outcome extracted from the literature review, and would capture proven infection not falling into one of the organ system infection groups recommended as endpoints. The wording is based on CDC definitions, and 30 days was agreed as the recommended period of follow up.[28]

Surgical site infection

SSI was the infection-related outcome extracted from the literature search with the highest frequency as a primary endpoint. Classification, definition, and follow-up period of SSI were varied, but CDC definitions were used in 61 out of 242 of studies as well as by regulatory agencies (e.g. Public Health England in their surveillance of SSIs in the NHS,[29] NICE guidelines on infection prevention[30]). We therefore recommend using CDC definitions for all types of SSI with a 30 day follow-up period except for deep and organ/space SSIs in a subset of operations, including breast, cardiac, and spinal surgery (see Supplementary Table S4, where 90 days is recommended).[28]

Infection with organ dysfunction (sepsis)

Sepsis occurred moderately frequently as an endpoint in the literature (see Supplementary Table S2) and is hugely important as an infective cause of morbidity and mortality.[31] Sepsis definitions were variable in the literature review, and the group therefore recommends adopting the Sepsis-3 definitions of sepsis and septic shock.[5] Systemic inflammatory response syndrome (SIRS) was evaluated but rejected by the wider study group as its specificity was considered poor in the postoperative patient group.[32]

Limitations

Literature review and the Delphi process allow for development of standardised endpoints that reflect current research practice. However, this process does have limitations. Focusing on endpoints extracted from literature review may mean that more novel measures of infection, which are not yet widely used but have the potential to be useful endpoints, are excluded. The Delphi process both adds a further time lag, but also offers a mitigation by specifically seeking alternative endpoints, not arising from the literature review, through the subgroup members (e.g. procalcitonin in our work). Additionally, when generating our long list of outcomes, we selected only higher impact factor journals from recent years, potentially limiting our view of wider research practice. A further limitation is potential publication bias as outcomes that are routinely monitored but not reported in the final publications captured by the systematic review will be missed.[32] This is mitigated by permitting Delphi participants to propose outcomes. However, the StEP-COMPAC proposals should be considered a living document that would require review and iteration over time. A final, but important, limitation is that although our wider working group was made up of numerous expert perioperative physicians, we had no patient representatives. This is partly remedied by the COMPAC initiative that is running in parallel with the work of the StEP group, and includes patient and carer representatives.[2]

Conclusions

Infection is a significant cause of morbidity and mortality in the surgical population. Use of standardised endpoints in future perioperative trials will bring uniformity to results and should allow collaboration and comparison. The endpoints recommended here do not represent a complete requirement for all future trials, but should be the default starting point.

Authors' contributions

Literature search: JB, SH.
Literature review: JB, JH, SH.
Endpoint longlisting: JB, JH, SH, MM.
Templates for Delphi round: JH, JB, SH.
Collation and analysis of data from Delphi rounds: JH, JB, SH.
Drafting of the manuscript: JH, JB, SH.
Manuscript review: MSH, ED, IJ, AAK, TC, JC, MPWG, MGM.
StEP sepsis subgroup: SH, MSH, ED, IJ, CK, AAK, TC, JC, SD, MPWG, MGM.

Declarations of interests

AK reports grants and personal fees from Pharmacosmos, personal fees from Vifor Pharma, personal fees from Masimo, grants from Hemosonics, grants from Hemonetics, grants and personal fees from Fisher and Paykel, outside the submitted work.
MPWG reports personal fees from Sphere Medical Ltd, outside the submitted work; co-editor-in-chief of *Perioperative Medicine*; elected council member of the Royal College of Anaesthetists; and chair of UK National Institute of Academic Anaesthesia board.
MM reports grants and other from Smiths Medical, personal fees from Edwards Lifesciences, personal fees from Baxter, outside the submitted work; founding editor-in-chief of *TopMedTalk*; founding editor of *Perioperative Medicine*; Editorial Board Member of The *British Journal of Anaesthesia*; editorial board of *Critical Care*; elected council member of the Royal College of Anaesthetists; director of Evidence Based Medicine community interest company; co-director of Duke-UCL Morpheus Consortium.
The other authors declare that they have no conflicts of interest.

Funding

The Standardized Endpoints for Perioperative Medicine (StEP) Collaborative has been supported by an unrestricted grant

from the *British Journal of Anaesthesia* as well as administrative support from the National Institute of Academic Anaesthesia Health Services Research Centre and the Department of Anaesthesia and Perioperative Medicine, Monash University. M.P.W.G. is a National Institute of Health Research Senior Investigator. M.S.H. is supported by the National Institute for Health Research Clinician Scientist Award (NIHR-CS-2016-16-011). The views expressed in this publication are those of the authors and not necessarily those of the National Health Service, the National Institute for Health Research or the Department of Health.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bja.2019.01.009.

References

1. Semmelweis IP. *Offener brief an sämtliche professoren der Geburtshilfe*. Österreichische Nationalbibliothek; 1862
2. Myles PS, Grocott MPW, Boney O, Moonesinghe SR, COMPAC-StEP Group. Standardizing end points in perioperative trials: towards a core and extended outcome set. *Br J Anaesth* 2016; **116**: 586–9
3. Gargon E, Williamson PR, Altman DG, Blazeby JM, Tunis S, Clarke M. The COMET Initiative database: progress and activities update (2015). *Trials* 2017; **18**: 54
4. Williamson P, Altman DG, Bagley H, et al. The COMET handbook: version 1.0. *Trials* 2017; **18**: 280
5. Singer M, Deutschman CS, Seymour CW, et al. The third international consensus definitions for sepsis and septic shock (Sepsis-3). *JAMA* 2016; **315**: 801–10
6. Horan TC, Gaynes RP, Martone WJ, Jarvis WR, Emori TG. CDC definitions of nosocomial surgical site infections, 1992: a modification of CDC definitions of surgical wound infections. *Infect Control Hosp Epidemiol* 1992; **13**: 606–8
7. Shankar-Hari M, Phillips GS, Levy ML, et al. Developing a new definition and assessing new clinical criteria for septic shock: for the third international consensus definitions for sepsis and septic shock (Sepsis-3). *JAMA* 2016; **315**: 775–87
8. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009; **339**: b2535
9. National Library of Medicine (U.S.) Bibliographic Services Division. *Abridged Index Medicus (AIM or core clinical) Journal Titles* 2003
10. Shankar-Hari M, Bertolini G, Brunkhorst FM, et al. Judging quality of current septic shock definitions and criteria. *Crit Care* 2015; **19**: 445
11. Myles P, Grocott M, Boney O, Moonesinghe S. Standardising endpoints in perioperative trials: towards a core and extended outcome set. *Br J Anaesth* 2016; **116**: 586–9
12. House J, Alexandraki I. *Assessment of postoperative fever — differential diagnosis of symptoms*. BMJ Best Practice; 2018. Available at: https://bestpractice.bmj.com/topics/en-gb/898/pdf/898.pdf
13. Jain R, Saxena D. Pyrexia: an update on importance in clinical practice. *Indian J Anaesth* 2015; **59**: 207–11
14. Fanning J, Neuhoff RA, Brewer JE, Castaneda T, Marcotte MP, Jacobson RL. Frequency and yield of postoperative fever evaluation. *Infect Dis Obstet Gynecol* 1998; **6**: 252–5
15. Yao Y-T, Li LH, Lei Q, et al. Noninfectious fever following aortic surgery: incidence, risk factors, and outcomes. *Chin Med Sci J* 2009; **24**: 213–9
16. Galicier C, Richet H. A prospective study of postoperative fever in a general surgery department. *Infect Control* 1985; **6**: 487–90
17. Guinn S, Castro Jr FP, Garcia R, Barrack RL. Fever following total knee arthroplasty. *Am J Knee Surg* 1999; **12**: 161–4
18. Pile JC. Evaluating postoperative fever: a focused approach. *Cleve Clin J Med* 2006; **73**: S62–6
19. Dolan RD, McSorley ST, McMillan DC, Horgan PG. Attitudes of surgeons to the use of postoperative markers of the systemic inflammatory response following elective surgery. *Ann Med Surg (Lond)* 2017; **21**: 14–9
20. Facy O, Pacquette B, Orry D, et al. Inflammatory markers as early predictors of infection after colorectal surgery: the same cut-off values in laparoscopy and laparotomy? *Int J Colorectal Dis* 2017; **32**: 857–63
21. Tang BMP, Eslick GD, Craig JC, McLean AS. Accuracy of procalcitonin for sepsis diagnosis in critically ill patients: systematic review and meta-analysis. *Lancet Infect Dis* 2007; **7**: 210–7
22. Cole DS, Watts A, Scott-Coombes D, Avades T. Clinical utility of peri-operative C-reactive protein testing in general surgery. *Ann R Coll Surg Engl* 2008; **90**: 317–21
23. Deirmengian GK, Zmistowski B, Jacovides C, O'Neil J, Parvizi J. Leukocytosis is common after total hip and knee arthroplasty. *Clin Orthop Relat Res* 2011; **469**: 3031–6
24. Bozkurt IH, Aydogdu O, Yonguc T, et al. Predictive value of leukocytosis for infectious complications after percutaneous nephrolithotomy. *Urology* 2015; **86**: 25–9
25. Surveillance Definitions. c2018. National healthcare safety network, Centers for Disease Control and prevention (accessed August, 2018). https://www.cdc.gov/nhsn/pdfs/pscmanual/17pscnosinfdef_current.pdf.
26. Abbott TEF, Fowler AJ, Pelosi P, et al. A systematic review and consensus definitions for standardised end-points in perioperative medicine: pulmonary complications. *Br J Anaesth* 2018; **120**: 1066–79
27. Conde M, Lawrence V. Postoperative pulmonary infections. *BMJ Clin Evid* 2008
28. Surgical site infection (SSI) event. c2018. National healthcare safety network, Centers for Disease Control and prevention (accessed August 6, 2018) http://www.cdc.gov/nhsn/pdfs/pscmanual/9pscssicurrent.pdf.
29. *Surveillance of surgical site infections in NHS hospitals in England, 2016 to 2017*. Public Health England; 2017. Available from: www.gov.uk/phe
30. NICE Quality Standard, Surgical Site Infection. c2013. National Institute for Health and care excellence (accessed August 6, 2018) https://www.nice.org.uk/guidance/qs49/resources/surgical-site-infection-2098675107781.
31. Vogel TR, Dombrovskiy VY, Lowry SF. Trends in postoperative sepsis: are we improving outcomes? *Surg Infect* 2009; **10**: 71–8
32. Smith V, Clarke M, Williamson P, Gargon E. Survey of new 2007 and 2011 Cochrane reviews found 37% of prespecified outcomes not reported. *J Clin Epidemiol* 2015; **68**: 237–45

Handling editor: H.C. Hemmings Jr