BJA
British Journal of Anaesthesia

Core Outcome Measures for Perioperative and Anaesthetic Care (COMPAC): A modified Delphi process to develop a Core Outcome Set for trials in Perioperative Care and Anaesthesia

Journal:	*British Journal of Anaesthesia*
Manuscript ID	BJA-2021-00806-HH484.R1
Article Type:	Clinical Investigation
Date Submitted by the Author:	20-Aug-2021
Complete List of Authors:	Boney, Oliver; University College London, Surgical Outcomes Research Centre, Centre for Perioperative Medicine; Royal College of Anaesthetists, Health Services Research Centre, National Institute of Academic Anaesthesia Moonesinghe, Suneetha Ramani; Royal College of Anaesthetists, Health Services Research Centre, National Institute of Academic Anaesthesia; University College London, Surgical Outcomes Research Centre, Centre for Perioperative Medicine Myles, Paul; Alfred Hospital, Department of Anaesthesiology & Perioperative Medicine; Monash University, Dept of Anaesthesia & Perioperative Medicine Grocott, Michael; University of Southampton, Integrative Physiology and Critical Illness Group, Clinical and Experimental Sciences, Faculty of Medicine; Southampton NIHR Biomedical Research Centre, Perioperative and Critical Care Theme
Keywords:	core outcome, patient-centred outcome, standardised endpoints, perioperative medicine, anaesthesia, surgery

Core Outcome Measures for Perioperative and Anaesthetic Care (COMPAC): a modified Delphi process to develop a core outcome set for trials in perioperative care and anaesthesia

Oliver Boney[1,2,*], S. Ramani Moonesinghe[1,2] Paul S. Myles[3], Michael P.W. Grocott[4,5] and the StEP-COMPAC group.

[1] Surgical Outcomes Research Centre, University College Hospital, London, UK, [2] Health Services Research Centre, National Institute of Academic Anaesthesia, Royal College of Anaesthetists, London, UK, [3] Department of Anaesthesiology and Perioperative Medicine, Alfred Hospital and Monash University, Melbourne, ~~Victoria~~VIC, Australia, [4] Integrative Physiology and Critical Illness Group, Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK, [5] Perioperative and Critical Care Theme, Southampton NIHR Biomedical Research Centre, University Hospital Southampton, Southampton, UK~~.~~

*Corresponding author. E-mail: oliver.boney1@nhs.net

Running title: Core outcomes in perioperative care

Editor's key points

Selecting outcomes for clinical trials is crucial to trail design, study interpretation and impact, reproducibility and data synthesis, and are currently inconsistent.

A modified Delphi approach was used to develop Core Outcome Measures for Anaesthesia and Perioperative Medicine (COMPAC) involving clinicians, patients and public members through a systematic review, a cross-sectional survey and a modified Delphi process to develop the final core outcome set.

The core outcome set includes the clinical outcome measures of postoperative mortality and morbidity, length of hospital stay and unplanned readmission as well as patient-centred outcomes of short- and long-term patient recovery after surgery.

The COMPAC core outcome set provides a broad stakeholder guide for selecting outcomes for future perioperative clinical trials.

Abstract

Background: Outcome selection underpins clinical trial interpretation. Inconsistency in outcome selection and reporting hinders comparison of different trials' results, reducing the utility of research findings.

Methods: We conducted an iterative consensus process to develop a set of Core Outcome Measures for Anaesthesia and Perioperative Medicine, following the established Core Outcome Measures for Effectiveness Trials (COMET) methodology. First, we undertook a systematic review of randomised controlled trials in high impact journals to describe current outcome reporting trends. We then surveyed patients, carers, researchers and perioperative clinicians about important outcomes after surgery. Finally, a purposive stakeholder sample participated in a modified Delphi process to develop a core outcome set for perioperative and anaesthesia trials.

Results: Our systematic review revealed widespread inconsistency in outcome reporting, with variable or absent definitions, levels of detail and temporal criteria. In the survey, almost all patients, carers and clinicians rated clinical outcome measures critically important, but clinicians rated patient-centred outcomes less highly than patients and carers.

The final core outcome set was: (i) Mortality/survival (postoperative mortality, long-term survival); (ii) Perioperative complications (major postoperative complications/adverse events; complications/adverse events causing permanent harm); (iii) Resource use (length of hospital stay, unplanned readmission within 30 days); (iv) Short term recovery (discharge destination, and/or level of dependence); and (v) Longer-term recovery (overall health-related quality of life).

Conclusions: This core set, incorporating important outcomes for both clinicians and patients, should guide outcome selection in future perioperative medicine or anaesthesia trials. Mapping these alongside standardised endpoint definitions will yield a comprehensive perioperative outcome framework.

Keywords: anaesthesia, core outcome, patient-centred outcome, perioperative medicine, standardised endpoints, ~~perioperative medicine, anaesthesia,~~ surgery

Introduction

Selecting outcome measures, and deciding how they are defined and reported, is a fundamental aspect of trial design. Multiple concerns have been raised about outcome measure selection including the use of surrogate endpoints of little relevance to clinicians, and/or patients or both;[1-3] using variable, inconsistent or poorly defined outcomes;[4, 5] and omissions when reporting trial outcomes.[6, 7] Poorly chosen or ill-defined outcomes lead at best to research that is difficult to interpret in the context of existing evidence, and at worst to ambiguous or misleading findings of uncertain value.

Moreover, combining the results of individual trials in systematic reviews and/or meta-analysis is hindered when trials investigating similar research questions report their outcomes using different metrics or timepoints. Such research may be not only wasteful but actively harmful, since it 'muddies the waters' and creates uncertainty where clear answers may in fact exist.[8, 9] Greater attention to outcome reporting is therefore fundamental to avoiding wasteful 'research waste'.[10]

Finally, the increasing emphasis on patient-centred outcome measures (such as health-related quality of life) alongside traditional clinical outcomes (such as postoperative complications) has led to calls for greater involvement of patients in the selection of trial outcomes.[11] Outcomes reported in perioperative research must be truly meaningful for patients, and should therefore include multidimensional outcome measures that capture all aspects of patient recovery after surgery.[12, 13]

Although several guidelines exist to support clinical trial design and reporting,[14, 15] there are few published recommendations specifically addressing the selection and reporting of trial outcomes. One potential solution is the development of Core Outcome sets (COSs) specific to different clinical research settings.[16] The Core Outcome Measures for Effectiveness Trials (COMET) Initiative defines a core outcome set COS as 'an agreed set of outcomes that should be measured and reported, as a minimum, in all clinical trials' in a specific area of health research.[17] Growing interest in improving outcome standardisation has stimulated several COS development initiatives in a variety of disease areas,[18-21] and the use of core outcome sets COSs (where available) has been endorsed by journal editors, regulatory bodies and research funders.[22, 23]

While core outcome sets COSs relevant to certain surgical populations have been published,[18, 20] no COS has yet been developed for research spanning all areas of perioperative care. The COMPAC-StEP initiative, – an international collaboration seeking to address the challenges associated with perioperative research outcome reporting, – was established to address this deficit in 2015 at a meeting sponsored by the BJA*British Journal of Anaesthesia*. It comprised two parallel work streams seeking consensus on, respectively, *what* to measure (COMPAC, Core Outcome Measures for Perioperative and Anaesthetic Care), and *how* to measure it (StEP, Standardized Endpoints in Perioperative Medicine).[24]

Several of the StEP initiative's expert-led working groups have already published their recommendations.[25-32] Here we report the results of COMPAC, whose aim was to identify outcomes of fundamental importance to end users of health services, namely patients, carers and their families (hereafter termed service users) and perioperative healthcare professionals.

Methods

The COMPAC initiative used mixed methods to develop a core outcome set ~~COS~~ for trials in adult patients in the clinical fields of anaesthesia and perioperative medicine. It involved clinicians, patients and public and comprised three phases (figure 1). Phase One was a systematic review to describe recent outcome measurement and reporting trends in the perioperative randomised controlled trial literature. Phase Two was a cross-sectional survey of stakeholders regarding the importance of the most common outcomes identified in Phase One. The findings of Phases One and Two were then combined to generate a longlist of 'candidate' core outcome measures. In phase Three, a purposive sample of stakeholder representatives took part in a modified Delphi process to develop the final core outcome set~~COS~~.

~~**Figure 1.** COMPAC process flow diagram showing the stages in development of the COMPAC COS.~~

Definitions

'Perioperative care': all aspects of patient care (except surgery itself) around a surgical episode, both inpatient and outpatient, from first consideration of surgery until full recovery or death.

'Major surgery': any non-obstetric surgical procedure for which the patient is admitted to hospital (i.e. non-ambulatory surgery).

'Outcome measure' or 'endpoint': any variable used by trial investigators to evaluate the effect of the intervention under investigation.[33, 34]

The process followed the COMET Initiative~~'s published~~ guidance for core outcome set ~~COS~~ development,[17] and was conducted in accordance with the Core Outcome Set–Standards for Development (COS-STAD) and Reporting (COS-STAR) recommendations.[35] The COMPAC study was prospectively registered on the COMET database;[36] ethical approval for Phase Two was obtained from the UK NHS Health Research Authority (Research Ethics Committee reference 17/EM/0096).

All stages of COMPAC were overseen and agreed by the COMPAC Steering Group, which comprised the four authors~~,~~ and two lay representatives from the Royal College of

Anaesthetists Lay Committee. The Steering Group convened periodically to review progress and resolve any methodological issues arising during the project, such as refining the wording used for outcomes in the Delphi surveys.

Phase One: ~~(~~systematic review~~):~~

The purpose of Phase One was to develop a list of outcome measures for the stakeholders to consider for inclusion in the core outcome set ~~COS~~ in Phases Two and Three. Four online databases (OVID Medline, EMBASE, Web of Science and the Cochrane Library) were searched between February and June 2016 (see Appendix for full search strategy). We included randomised controlled trials (RCTs) published between January 2005 and December 2014 in the four highest impact journals in surgery, anaesthesia and general medicine (Table 1).

We only included trials ~~which~~ that recruited at least 100 patients. Titles and abstracts of retrieved publications were screened by the lead author. The full text of each relevant RCT was then accessed, and details of the trial and its reported primary and secondary outcomes were extracted into a database. Outcomes were transcribed verbatim, with details of the timing of measurement and definition(s) where available. Figure 2 shows a PRISMA flowchart for this process.

We did not conduct formal quality assessments for each included trial, as the review's purpose was not to synthesise evidence but to describe perioperative outcome reporting practices across influential, high impact journals.

Phase Two: ~~(~~stakeholder consultation~~):~~

To understand the views of relevant stakeholders about important outcomes after major surgery, we undertook a multi-centre cross-sectional study, P-COMMaS (Patient-Centred Outcome Measures for Major Surgery). Eligible participants were recruited from NHS hospitals across England and Wales, and included:

- Service users – adult patients with experience of having major surgery (defined as any non-obstetric operation requiring at least one night in hospital) and their carers (i.e. any friend or relative with significant experience of caring for a patient after major surgery)
- Clinicians – healthcare professionals who provide care for adult major surgical patients before, during or after surgery

Participants completed a survey in which they were asked to score the importance of the main outcome measures identified in Phase One, and to suggest (via free text comments) any additional outcomes they considered important for reporting in perioperative research. Scores were divided into three categories: 'critically important' (≥8 out of 10); 'fairly important' (5-7 out of 10); 'not very important' (≤4). Mean scores for each outcome were calculated to determine the proportion of respondents in each group rating the outcome

'critically important'. An analysis exploring free text comments and their relationship to each group's scores will be reported separately.

Phase Three: ~~(~~Delphi process~~):~~

The COMPAC Steering Group generated a longlist of potential core outcomes based on the results of Phases One and Two, categorised into six domains adapted from the COMET Initiative's proposed outcomes taxonomy.[37] The longlist was incorporated into an online Delphi survey, which was piloted among a small number of non-clinicians for ease of understanding and explanation about core outcome sets. A purposive sample of P-COMMaS participants who had indicated their willingness to take part were invited to participate in the Delphi process, aiming for ~~between~~ 50 ~~and~~ to 100 participants, with sufficient diversity to include both service users and clinicians, and representation across age groups, genders, ethnic groupings, and personal or professional experience of different surgery types. Two Delphi rounds were undertaken:

Round one: Participants completed an online survey asking them to rate the importance of each candidate outcome on a five-point scale (from '1= not at all important' to '5= very important'). The five point Likert-type scale has been used in other core outcome set ~~COS~~ Delphi processes and was chosen for its simplicity while retaining adequate discriminatory potential.[38] Mean scores for each outcome were calculated for all respondents, and separately for service users and clinicians. Outcomes were retained for the second round if:

- ≥65% of respondents from **either** stakeholder group scored the outcome ≥4 (i.e. 'important' or 'very important'), AND
- <20% of respondents from that stakeholder group scored the outcome ≤2 (i.e. 'not very important' or 'not at all important')

While there are no agreed cut-off criteria for retaining or discarding options in Delphi studies, these limits were similar to those used in other recent core outcome set ~~COS~~ development initiatives,[18, 39] and were agreed collectively by the COMPAC Steering Group *a priori*, as recommended by the COMET Initiative core outcome set ~~COS~~ development guidance.[17]

Round two: Participants were then invited to complete a second online survey, in which they again scored the importance of each retained candidate outcome. However, ~~this time~~for this round, respondents were shown their own score and the mean scores from Round One for all respondents, and for each stakeholder group. Outcomes where a large discrepancy between clinicians' and service users' mean scores had been observed (defined as a difference of >0.5 out of 5) were highlighted in orange; those with a difference >0.3 but <0.5 were highlighted yellow.

Mean scores in this second survey were again calculated for all respondents and for each group. Outcomes were retained for the final workshop if:

- ≥65% of respondents from **both** stakeholder groups scored the outcome ≥4, AND

- <20% of respondents from **both** stakeholder groups score the outcome ≤2

After analysis of the second survey results, the COMPAC Steering Group refined and/or abbreviated the longest or 'wordiest' outcomes for ease of small group viewing and discussion at the final workshop.

Final workshop: The final core outcome set ~~COS~~ was agreed at a face-to-face consensus workshop in February 2020, led by two independent facilitators from the James Lind Alliance. All respondents to the first Delphi survey were invited; as with the initial COMPAC Delphi group, an even balance of service users and clinicians was sought.

The workshop consisted of three rounds of small group discussion followed by plenary discussions. Groups were pre-allocated to provide within-group balance between service users and clinicians; they discussed the candidate outcome measures and then assigned each as either 'definite' core outcomes (scoring one point), 'equivocal/no consensus' (no points), or 'not' core outcomes (minus one). All participants then reconvened for a plenary session where group scores were combined, and their respective choices discussed among all participants in order to reach 'whole group consensus' regarding each outcome's inclusion in the final core outcome set~~COS~~.

Outcomes with a combined score of ≤0 were excluded, while those scoring ≥2 were included as core outcomes. Outcomes scoring 1 were labelled 'desirable', and therefore underwent further discussion in the plenary sessions before taking a final vote via a show of hands, with a threshold for core outcome set ~~COS~~ inclusion of ≥70% support. A final plenary session was held at the end of the workshop, with a last opportunity for participants to propose further discussion and/or a repeat vote for any of the 'desirable' outcomes. This final session concluded with all delegates ratifying the 'included' outcomes as the agreed COMPAC core outcome set~~COS~~.

Results

Phase One: ~~(~~systematic review~~):~~

Our searches yielded 12,971 trials across the four databases, of which 674 were deemed eligible for inclusion after full text review (see Figure 2). The number of included trials by journal is shown in Supplementary Table S1.

~~**Figure 2.** PRISMA diagram showing identification, screening and inclusion process for database searches. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.~~

Verbatim outcomes were extracted from all included trials, and categorised into seven overarching domains. At least one mortality or survival outcome was reported in 419 RCTs

(62%); among the 317 RCTs reporting short-term mortality (up to one year postoperative), the most common were 30-day and in-hospital mortality. Most (222 RCTs, 72%) did not specify causality when reporting mortality. Meanwhile longer-term survival outcomes (beyond a year postoperative) were reported in 164 trials (24%), with almost all reporting overall survival, and over three-quarters reporting disease-specific survival. However even within a single journal (the Lancet), terminology varied between trials for long-term disease-specific survival, with 'disease-free survival', 'progression-free survival', 'biochemical progression-free survival', 'clinical progression-free survival', 'recurrence-free survival', 'invasive-disease-free-survival', 'relapse-free survival', 'metastasis-free survival', and 'event-free survival' variously reported.

Perioperative morbidity outcomes were reported in 412 RCTs (61%). As with mortality, these outcomes were reported at several different time points; details regarding how perioperative morbidity was sought, defined and reported were also extremely variable. A minority (130 RCTs, 32%) used standardised definitions (e.g. for pancreatic fistula)[40] or classification systems (e.g. the Clavien-Dindo grading system) to define morbidity;[41] however others variously reported incidence of specific complications, total complications, or of patients needing re-intervention (usually without further details). This theme of inconsistency and heterogeneity in outcome measurement and reporting was similarly repeated across the other outcome domains.

The most frequently reported outcomes are summarised in Table 1.

Phase Two: ~~(~~P-COMMaS~~):~~

Between April and August 2017, 4,105 participants (2,582 service users and 1,522 clinicians) from ~~thirty~~ 30 NHS hospitals in England and Wales completed the P-COMMaS survey. ~~One~~ There were 118 incomplete ~~hundred and eighteen~~ responses which were ~~incomplete and therefore~~ excluded; the remaining 3,986 responses were included for analysis. Nearly two-thirds (2,489, 62%) were service users; almost half of these (47%) were aged ≥65 yr, over half (61%) were female, and over half (56%) had had at least two previous operations. The 1,497 clinicians comprised mainly nurses (47%) and doctors (35%) from various specialities; the commonest areas of perioperative expertise were general surgical and trauma and orthopaedic patients.

The vast majority of respondents (3,850, 96.6%) rated 'the operation being successful' as critically important, i.e. scoring at least ~~eight~~ 8 out of ~~ten~~10. 'Avoiding postoperative complications that might cause permanent disability' was almost universally deemed critically important (3,958 respondents, 99.3%), while around 95% rated 'Not dying during or soon after your operation', and 'Avoiding postoperative complications that might delay recovery' critically important. For all these outcomes, the proportion of service users rating the outcome critically important was slightly higher than that of clinicians.

Patient-centred outcomes scored overall slightly less highly. Under ~~ninety per cent~~90% of respondents (3,455, 86.7%) ascribed critical importance to 'Being comfortable after the operation', while 2,834 (71.1%) deemed 'Getting back to work, or normal daily activities, as soon as possible' critically important, and 2,655 (66.6%) considered 'Getting out of hospital as soon as possible' a critically important outcome. Furthermore, the proportion of clinicians rating these outcomes critically important was significantly lower than of service users ($p<0.001$, Chi-Squared test).

~~Slightly over~~Just over a third of respondents (1,389, or 35%) submitted additional free text comments regarding important outcomes after major surgery, with similar proportions of service users (871, 35%) and clinicians (518, 35%). While many comments provided rich perspectives on what matters to patients following major surgery, no new outcomes not already encountered in the systematic review were identified.
Respondent demographics and ratings are summarised in Tables 2 and 3 respectively.

Phase Three: ~~(~~Delphi process~~):~~
A longlist of ~~sixty six~~66 candidate outcomes (Supplementary Table S2) was constructed by the COMPAC Steering Group for the subsequent Delphi process, based on the outcome measures most frequently reported in the systematic review and the P-COMMaS survey ratings.

Delphi Round One:
~~Eighty-two stakeholders completed t~~The first Delphi survey~~, representing slightly over a third~~was completed by 82 (41%) stakeholders of the roughly ~~two hundred~~200 email invitations sent. ~~Fifteen~~ The 15 incomplete responses were ~~incomplete and therefore~~ excluded; the 67 valid responses were included in the analysis.
Forty respondents (63%) identified primarily as service users, and 27 (37%) as clinicians. All but one service user had lived experience as a surgical patient; ~~ten~~ 10 also had experience of caring for a loved one after major surgery, while one had carer experience only. The 27 clinicians included ~~ten~~ 10 anaesthetists and/or intensivists, four surgeons, three geriatricians, one GP, three trainee (junior) doctors, two clinical nurse specialists, one advanced nurse practitioner and two nurses.

Nine (14%) of the 66 longlist outcomes did not meet the retention thresholds for Round Two. Of the 57 outcomes retained, 49 (86%) were rated $\geq$4 by over 65% of both service users and clinicians; the other eight were rated $\geq$4 by >65% of service users but <65% of clinicians. As <20% of both groups rated these eight outcomes $\leq$2, however, they were retained for Round Two.

Delphi Round Two:

Forty-three (64%) of those who completed the first Delphi survey also provided complete responses to the second survey (three additional incomplete responses were excluded). ~~Twenty~~ This included 20 (47%) ~~were~~ clinicians (~~ten~~ 10 anaesthetists and/or intensivists, three surgeons, two geriatricians, three trainee doctors, one clinical nurse specialist and one nurse) ~~while~~ and 23 (53%) ~~were~~ service users (all with experience as surgical patients, and six with carer experience). Only two of the 57 outcomes retained from Round One were rated ≥4 by under 65% of both groups and were discarded; the other 55 all met the criteria for retention.

During the COMPAC Steering Group's review of the wording of outcomes retained for the final workshop, three further pairs of similar/overlapping outcomes were amalgamated into a single outcome (see Supplementary Table S3), leaving 52 candidate outcomes for discussion (Supplementary Table S4).

Final workshop:

Twenty-six participants (13 clinicians and 13 service users) attended the final workshop. Following three iterative rounds of small group discussions, scores were refined and agreed for each of the 52 outcomes under consideration. Combined small group scores (from -3, ~~meaning~~ i.e. all groups considered the outcome insufficiently important for core outcome set ~~COS~~ inclusion, to +3, i.e. universally agreed for inclusion) are presented in Supplementary Table S3. Seven outcomes scored ≥2, and were therefore included; 35 scored ≤0, and were therefore excluded. The remaining ~~ten~~ 10 scored ~~one point~~1, and underwent further discussion and voting in the plenary sessions, during which participants reached the following conclusions:

- Strong support for 'postoperative level of dependence' as a recovery metric that is meaningful to patients, carers and clinicians. Delegates therefore decided to include 'Discharge destination from hospital (e.g., own home/rehab facility/care home), AND/OR level of dependence (need for carers)' as a core outcome.
- Strong support for outcomes regarding postoperative pain, nausea and vomiting, and postoperative mental, emotional and psychological wellbeing. While none reached the ≥70% inclusion threshold, consensus was reached to highlight these as 'Additional important patient-centred outcomes.'
- Patient satisfaction also received broad support: although not reaching the 70% inclusion threshold, >70% voted to highlight it as an 'additional important patient-centred outcome.' The final wording agreed was 'Patient satisfaction with their operation, and/or willingness (with hindsight) to choose the same again'.

The final core outcome set ~~COS~~ is presented in Table 4. Where applicable, the relevant StEP working group and any associated publications are also listed.

Discussion

The COMPAC initiative has developed the first available core outcome set ~~COS~~ for anaesthesia and perioperative care, following the methodology recommended by the COMET Initiative. The core outcome set ~~COS~~ includes the already-widely used 'traditional' clinical outcome measures of postoperative mortality and morbidity, as well as resource use outcomes (length of hospital stay and unplanned readmission). However, patient-centred outcomes, ~~–~~ namely both short- and long-term patient recovery after surgery, ~~–~~ also feature prominently. The emphasis on major or permanently disabling postoperative complications ~~also~~ demonstrates the overriding importance of outcomes with meaningful impact for patients. This should be acknowledged when reporting complications which, though arguably of clinical significance (for example, postoperative troponin rise), may not directly affect patient~~s'~~ quality of life.

~~COMPAC's~~ The main strength of COMPAC was the extent of stakeholder involvement at every stage. The importance of balanced stakeholder representation was illustrated by our finding that clinicians rated patient-centred outcomes less highly than patients: under-representation of patients and carers in the core outcome set ~~COS~~ development process could arguably have led these outcomes being omitted from the final core outcome set~~COS~~. We therefore consider this core outcome set ~~COS~~ an accurate reflection of outcomes that matter most to all three main stakeholder groups of patients, carers and clinicians. For this reason, it should be considered a ~~'~~default starting point~~'~~ for selecting outcomes for future perioperative clinical trials unless the outcome is either demonstrably irrelevant to the trial in question, or unpragmatic because of resource constraints. Our findings also strengthen the case for collaborative trial outcome selection between researchers, clinicians and patient representatives.[13]

COMPAC is to our knowledge the first available guidance for reporting outcomes spanning the breadth of anaesthesia and perioperative medicine research. Where more specific ~~COSs~~ core outcome sets already exist for specific surgical populations, our recommendation is that the COMPAC guidance is considered alongside these. ~~COSs~~ Core outcome sets represent the ~~'~~minimum~~'~~ outcomes that should be reported; researchers are not restricted from collecting additional outcomes they consider relevant. Furthermore, while we contend that researchers should have a clear justification for omitting any of the core outcomes, the precise endpoint chosen (for example, in-hospital or 30-day or 90-day mortality) is left to their discretion. The StEP initiative complements the COMPAC core outcomes by providing recommendations and definitions for specific endpoints such as (for example) postoperative pneumonia or cancer survival.

Several research organisations and funding bodies already endorse the use of formal standards for outcome selection and reporting.[42-45] For example, the UK's National Institute

for Health Research (NIHR) recommends in its Health Technology Assessment application guidelines that 'where established Core Outcomes exist, they should be included amongst the list of outcomes unless there is good reason to do otherwise.'[46] However, the potential of ~~COSs~~ core outcome sets to improve consistency of outcome measurement in trials rests on their adoption by researchers. Developing a core outcome set ~~COS~~ that is not implemented may in fact exacerbate research waste rather than reducing it. While endorsement of core outcome sets ~~COSs~~ by funding bodies and regulatory agencies has been shown to improve their uptake,[47] such endorsement should be based on explicit standards for core outcome set ~~COS~~ development.[35]

Although we consider that COMPAC followed the current 'gold standard' process for core outcome set ~~COS~~ development, and that the need for this COS core outcome set ~~was~~ not in doubt given the widely recognised issues (confirmed by our systematic review) with outcome selection and reporting in our specialty, this core outcome set ~~COS~~ does not provide definitive guidance for outcome measurement. COMPAC recommends, in broad terms, 'what' to measure; meanwhile specific details of precisely 'how' to measure each outcome are recommended by the various StEP subgroups.[24-32]

Some questions also emerged regarding the COMPAC consensus methodology. For example, the initial classification of outcomes from the systematic review, the creation of an outcome longlist from the 'raw data', and the rewording/abbreviation of certain outcomes before the final workshop all introduced an element of subjectivity. While noting that similar issues have been reported in other core outcome set ~~COS~~ development processes,[48] we attempted to minimise these potential sources of bias by using recognised outcome classification systems[37, 49] and through collective decision making by the whole COMPAC Steering Group.[50]

Two issues also call into question the overall scope and generalisability of this core outcome set~~COS~~. Firstly, just as the boundaries between surgery, anaesthesia and perioperative medicine are indistinct, there is overlap between this core outcome set ~~COS~~ and certain surgery-specific ones. Secondly, COMPAC was predominantly UK-based, with almost all contributors recruited from England and Wales (unlike the StEP initiative, which brought together perioperative research experts from around the world). Therefore while we consider that this core outcome set ~~COS~~ fairly reflects 'the outcomes that matter after surgery' for UK service users and clinicians, it may not capture all outcomes that are important (or pragmatic) for every healthcare jurisdiction and/or research context. Further research is required to assess the relevance and utility of this core outcome set ~~COS~~ in other healthcare jurisdictions.

Finally, outcome measurement is continually evolving. Outcomes that are valid and important today may become obsolete tomorrow; thus, any core outcome set ~~COS~~ will in time require revision and/or updating via an accepted consensus process to ensure it serves

the needs of its end users. The successful collaboration between service users, clinicians and researchers demonstrated in COMPAC will hopefully serve as an exemplar for future core outcome set ~~COS~~ development and/or reiterations.

Conclusions

The development of this core outcome set ~~COS~~ for anaesthesia and perioperative research confirmed the heterogeneity of outcome reporting in clinical trials in our field, thereby justifying the process we have now completed. COMPAC has also confirmed that traditional outcome measures used in perioperative trials, such as mortality, morbidity and resource use, are valuable to all stakeholders. However, the life impact of surgery on both patients' short-term recovery and longer-term health-related quality of life are also of fundamental importance. We recommend using baseline trial outcomes from this core outcome set ~~COS~~ in all perioperative trials, alongside appropriate standardised endpoint definitions such as those recommended by the StEP subgroups, to improve consistency in perioperative research. Adopting this comprehensive perioperative outcome framework will facilitate comparison of results between trials.

Authors' contributions~~:~~

Study concept and design: All authors

Systematic review: OB, SRM

Data acquisition and interpretation: OB, SRM, MPWG

Drafting and revision of the manuscript: All authors

Declarations of interests

PM is an editor of the *British Journal of Anaesthesia*. 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. The P-COMMaS study was supported by an Association of Anaesthetists small grant. M.P.W.G. is a National Institute of Health Research Senior Investigator. 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.

Acknowledgements

The authors wish to thank the lay members of the COMPAC Steering Group; all those who took part in P-COMMaS, and the research teams in each study site; and all those who contributed to the Delphi process, which could not have succeeded without their time, energy and expertise. The authors also wish to thank Laura Cortes, James Goodwin, Pam Hines and the research team at the National Institute of Academic Anaesthesia, as well as Sarah Hayden at UCL Centre for Perioperative Medicine, for their help administering the COMPAC initiative.

Appendix

Members of the StEP-COMPAC working group:

Executive: Paul Myles (Australia, Co-Chair) and Michael Grocott (UK, Co-Chair), Bruce Biccard (South Africa), Oliver Boney (PhD Fellow, UK), Matthew Chan (Hong Kong), Lee Fleisher (USA), Cor Kalkman (The Netherlands), Andrea Kurz (USA), Ramani Moonesinghe (UK), Duminda Wijeysundera (Canada).

Group members: J. Bartoszko, W. S. Beattie, R. Bellomo, D. Buggy, L. Cabrini, J. Canet, T. Cook, D. J. Cooper, T. Corcoran, P. J. Devereaux, R. Eckenhoff, L. Evered, T. J. Gan, T. Gin, H. Grocott, G. Haller, S. Howell, M. Jayarajah, C. Kalkman, K. Karkouti, B. Kavanagh, A. Klein, G. Landoni, K. Leslie, D. R. McIlroy, D. Mazer, A. Moller, M. Mythen, ~~M. Neuman,~~ M. Neuman, R. Pearse, P. Peyton, J. Prowle, T. Richards, D. A. Scott, D. Sessler, A. Shaw, T. Short, M. Shulman, B. Silbert, M. Singer, J. R. Sneyd, D. Story, D. van Dijk, and W. van Klei.

~~Supplementary online material~~

~~Phase One: Search strategy for systematic review; PRISMA reporting checklist 2020~~

- ~~Supplementary Table S1: Number of trials retrieved from each journal for the COMPAC systematic review, with total participant numbers~~

~~Phase Two: P-COMMaS survey~~

~~Phase Three:~~

- ~~Supplementary Table S2: Longlist of candidate outcomes for Delphi process~~
- ~~Supplementary Table S3: Candidate outcomes agreed by the COMPAC Steering Group to be amalgamated for the final workshop~~
- ~~Supplementary Table S4: Full list of outcomes considered by delegates at the final workshop, with combined small group scores~~

References

1 Fleming TR, Powers JH. Biomarkers and surrogate endpoints in clinical trials. *Stat Med* 2012; **31**: 2973-84
2 Weintraub WS, Lüscher TF, Pocock S. The perils of surrogate endpoints. *Eur Heart J* 2015; **36**: 2212-8
3 Myles PS. Perioperative outcomes: Are we asking the right questions? *Can J Anaesth* 2016; **63**: 138-41
4 Shulman M, Myles P. Measuring perioperative outcome. *Curr Opin Anaesthesiol* 2016; **29**: 733-8
5 Duncan PW, Jorgensen HS, Wade DT. Outcome measures in acute stroke trials: a systematic review and some recommendations to improve practice. *Stroke* 2000; **31**: 1429-38
6 Bonnot B, Yavchitz A, Mantz J, Paugam-Burtz C, Boutron I. Selective primary outcome reporting in high-impact journals of anaesthesia and pain. *Br J Anaesth* 2016; **117**: 542-3
7 Dwan K, Kirkham JJ, Williamson PR, Gamble C. Selective reporting of outcomes in randomised controlled trials in systematic reviews of cystic fibrosis. *BMJ Open* 2013; **3**
8 Bruce J, Krukowski ZH, Al-Khairy G, Russell EM, Park KG. Systematic review of the definition and measurement of anastomotic leak after gastrointestinal surgery. *Br J Surg* 2001; **88**: 1157-68
9 Clarke M. Standardising outcomes for clinical trials and systematic reviews. *Trials* 2007; **8**: 39
10 Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet* 2009; **374**: 86-9
11 Patient Centred Outcome Measures. 2015. Available from https://www.england.nhs.uk/ourwork/pe/pcoms/ (accessed June 2016)
12 Kalkman CJ, Kappen TH. Patient-centered endpoints for perioperative outcomes research. *Anesthesiology* 2015; **122**: 481-3

13 Heidegger T, Nuebling M, Saal D, Kreienbühl G. Patient-centred outcomes in clinical research: does it really matter? *Br J Anaesth* 2008; **100**: 1-3
14 Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; **340**: c332
15 von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. *Int J Surg* 2014; **12**: 1495-9
16 Gargon E, Gurung B, Medley N, et al. Choosing important health outcomes for comparative effectiveness research: a systematic review. *PLoS One* 2014; **9**: e99111
17 Williamson PR, Altman DG, Bagley H, et al. The COMET Handbook: version 1.0. *Trials* 2017; **18**: 280
18 McNair AG, Whistance RN, Forsythe RO, et al. Core Outcomes for Colorectal Cancer Surgery: A Consensus Study. *PLoS Med* 2016; **13**: e1002071
19 Dworkin RH, Turk DC, Farrar JT, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain* 2005; **113**: 9-19
20 O'Donnell CM, Black N, McCourt KC, et al. Development of a Core Outcome Set for studies evaluating the effects of anaesthesia on perioperative morbidity and mortality following hip fracture surgery. *Br J Anaesth* 2019; **122**: 120-30
21 Schmitt J, Apfelbacher C, Spuls PI, et al. The Harmonizing Outcome Measures for Eczema (HOME) roadmap: a methodological framework to develop core sets of outcome measurements in dermatology. *J Invest Dermatol* 2015; **135**: 24-30
22 Webster L, Groskreutz D, Grinbergs-Saull A, et al. Development of a core outcome set for disease modification trials in mild to moderate dementia: a systematic review, patient and public consultation and consensus recommendations. *Health Technol Assess* 2017; **21**: 1-192
23 Duffy J, Rolph R, Gale C, et al. Core outcome sets in women's and newborn health: a systematic review. *BJOG* 2017; **124**: 1481-9
24 Myles PS, Grocott MP, Boney O, Moonesinghe SR, Group C-S. Standardizing end points in perioperative trials: towards a core and extended outcome set. *Br J Anaesth* 2016; **116**: 586-9
25 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 [update]
26 Barnes J, Hunter J, Harris S, et al. Systematic review and consensus definitions for the Standardised Endpoints in Perioperative Medicine (StEP) initiative: infection and sepsis. *Br J Anaesth* 2019; **122**: 500-8
27 Beattie WS, Lalu M, Bocock M, et al. Systematic review and consensus definitions for the Standardized Endpoints in Perioperative Medicine (StEP) initiative: cardiovascular outcomes. *Br J Anaesth* 2021; **126**: 56-66
28 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
29 Buggy DJ, Freeman J, Johnson MZ, et al. Systematic review and consensus definitions for standardised endpoints in perioperative medicine: postoperative cancer outcomes. *Br J Anaesth* 2018; **121**: 38-44
30 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

31 McIlroy DR, Bellomo R, Billings FT, et al. Systematic review and consensus definitions for the Standardised Endpoints in Perioperative Medicine (StEP) initiative: renal endpoints. *Br J Anaesth* 2018; **121**: 1013-24
32 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
33 WHO.int. Health Impact Assessment: Glossary of terms used. 2017.
34 Velentgas P, Dreyer N, Wu A. Outcome Definition and Measurement. *Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide* 2013
35 Kirkham JJ, Davis K, Altman DG, et al. Core Outcome Set-STAndards for Development: The COS-STAD recommendations. *PLoS Med* 2017; **14**: e1002447
36 Grocott MPW, Myles PS, Moonesinghe SR, Boney OC. Core Outcome Measures in Perioperative and Anaesthetic Care (COMPAC). 2014. Available from www.comet-initiative.org/studies/details/632 2015)
37 Dodd S, Clarke M, Becker L, Mavergames C, Fish R, Williamson PR. A taxonomy has been developed for outcomes in medical research to help improve knowledge discovery. *J Clin Epidemiol* 2018; **96**: 84-92
38 De Meyer D, Kottner J, Beele H, et al. Delphi procedure in core outcome set development: rating scale and consensus criteria determined outcome selection. *J Clin Epidemiol* 2019; **111**: 23-31
39 Remus A, Smith V, Gutke A, et al. A core outcome set for research and clinical practice in women with pelvic girdle pain: PGP-COS. *PLoS One* 2021; **16**: e0247466
40 Bassi C, Marchegiani G, Dervenis C, et al. The 2016 update of the International Study Group (ISGPS) definition and grading of postoperative pancreatic fistula: 11 Years After. *Surgery* 2017; **161**: 584-91
41 Rassweiler JJ, Rassweiler MC, Michel MS. Classification of complications: is the Clavien-Dindo classification the gold standard? *Eur Urol* 2012; **62**: 256-8; discussion 9-60
42 Jammer I, Wickboldt N, Sander M, et al. Standards for definitions and use of outcome measures for clinical effectiveness research in perioperative medicine: European Perioperative Clinical Outcome (EPCO) definitions: a statement from the ESA-ESICM joint taskforce on perioperative outcome measures. *Eur J Anaesthesiol* 2015; **32**: 88-105
43 Cho JY, Baron TH, Carr-Locke DL, et al. Proposed standards for reporting outcomes of treating biliary injuries. *HPB (Oxford)* 2018; **20**: 370-8
44 Toozs-Hobson P, Bach F, Daly JO, Klarskov N. Minimum standards for reporting outcomes of surgery in urogynaecology. *Int Urogynecol J* 2020
45 Burns DJP, Arora J, Okunade O, et al. International Consortium for Health Outcomes Measurement (ICHOM): Standardized Patient-Centered Outcomes Measurement Set for Heart Failure Patients. *JACC Heart Fail* 2020; **8**: 212-22
46 HTA Stage 1 guidance notes. 2020. Available from https://www.nihr.ac.uk/documents/hta-stage-1-guidance-notes/11743 (accessed April 2021)
47 Kirkham JJ, Clarke M, Williamson PR. A methodological approach for assessing the uptake of core outcome sets using ClinicalTrials.gov: findings from a review of randomised controlled trials of rheumatoid arthritis. *BMJ* 2017; **357**: j2262
48 Gargon E, Williamson PR, Young B. Improving core outcome set development: qualitative interviews with developers provided pointers to inform guidance. *J Clin Epidemiol* 2017; **86**: 140-52

49 Macefield RC, Boulind CE, Blazeby JM. Selecting and measuring optimal outcomes for randomised controlled trials in surgery. *Langenbecks Arch Surg* 2014; **399**: 263-72
50 Williamson PR, Blazeby JM, Brookes ST, Clarke M, Terwee CB, Young B. Controversy and Debate Series on Core Outcome Sets. Paper 4: Debate on Paper 1 from the perspective of COMET [Core Outcome Measures in Effectiveness Trials]. *J Clin Epidemiol* 2020; **125**: 222-4

Figure legneds

Figure 1. COMPAC process flow diagram showing the stages in development of the COMPAC ~~CO~~Score outcome set.

Figure 2. PRISMA diagram showing identification, screening and inclusion process for database searches. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

Tables ~~of results~~

Outcome domain	RCTs reporting ≥1 outcome in this domain, n (%)	Main variations observed between RCTs	Most frequently reported (n, %)
Mortality and survival outcomes	419 (62%)	N~~/~~A	N~~/~~A
	317 (47%) – short-term perioperative mortality	Timeframe	30 day (145, 46%) In-hospital/inpatient (120, 38%) 12 month/1 year (53, 17%) 90 day/3 month (31, 9.8%)
		Cause of death	All-cause (69, 22%) Mortality from specific cause (21, 6.6%) Unspecified (227, 72%)
	164 (24%) – long-term survival	Cause of death	Overall survival (157, 98%) Disease-specific survival (122, 76%)
Perioperative morbidity, surgical complications, adverse events	412 (61%)	Timeframe	Specified timeframe for measuring complications (387, 94%)
		Criteria for assessing complications	Provided details of complications sought and reported (359, 87%) Used a standardised definition or criteria for complications (130, 32%)
		Severity grading	Graded severity using recognised system (90, 22%)
Composite outcomes	104 (15%)	Generally ‘bespoke’ for each RCT Outcomes of unequal severity combined into single composite outcome	Included mortality in the composite (89, 86%) ‘Major adverse cardiovascular or cerebrovascular events’ (MACCE) (33, 32%) ‘Major adverse cardiovascular events’ (MACE) (20, 19%)
Health resource use	348 (52%)	Reported mean vs median	Hospital length of stay (250, 72%) Critical Care length of stay (73, 21%)

		Details of how resource use was measured Timeframe Index admission only, or all admissions +/- outpatient attendances	Unplanned re-operation or other re-intervention (81, 23%) Unplanned hospital readmission or reattendance (51, 15%) Costs of inpatient care (32, 9.2%) Total costs of all healthcare attendances, interventions and investigations related to surgery (18, 5.2%)
Short-term postoperative recovery	240 (36%)	Aspect(s) of recovery reported Timeframe Assessment of severity/impact Measurement instrument used	Postoperative pain (124, 52%) Postoperative nausea and/or vomiting (99, 41%) Other physical symptoms (66, 28%) Psychological/mental wellbeing (22, 9.2%) Overall quality of postoperative recovery (13, 5.4%) Reported timeframe and/or frequency of assessment (226, 94%) Included implicit or explicit assessment of clinical severity and/or patient impact (227, 95%) [E.g. for postoperative pain, 124 RCTs]: Visual analogue scale (82, 66%) Verbal rating scale (18, 15%) Numerical rating scale (17, 14%) Other rating tool (6, 4.8%)
Patient-reported outcomes	206 (31%)	Outcome(s) reported	Health-related quality of life (HRQoL) - Generic (105, 51%) - Disease-specific (45, 22%) Impact on usual life roles (54, 26%) - Work (20, 9.7%) - Daily activities (18, 8.7%) - Other disability (16, 7.8%) Psychological/emotional/ mental wellbeing outcomes (28, 14%) Chronic pain (22, 11%) Overall extent of recovery (11, 5.3%) Fatigue, energy levels (7, 3.9%)

		Measurement instrument used	Patient satisfaction/regret (50, 24%) [e~~E~~.g. for generic HRQoL, 105 RCTs]: SF-36 (48, 46%) EQ-5D (34, 32%) QLQ-C30 (19, 18%)
		Timeframe	Reported timeframe and/or frequency of assessment (194, 94%)
Surgical success or failure	257 (38%)	Outcome reported	Disease recurrence (144, 56%) Need for repeat surgery or other re-intervention (48, 19%) Incidence of side effects and/or toxicity (102, 40%)
		Assessment of toxicity impact/ severity	Reported severity/patient impact of side effects and/or toxicity using recognised grading system (58, 57%)
		Timeframe	Reported the duration and/or frequency of postoperative follow-up (232, 90%)

Table 1. Most frequently reported outcome measures, and principal sources of heterogeneity in outcome reporting trends, amongst RCTs included in the COMPAC systematic review

	All respondents (n=3,986)	Service users (n=2,489)	Clinicians (n=1,497)
Age category (yr)			
18-25	130 (3.3%)	49 (2.0%)	81 (5.4%)
25-34	538 (13%)	138 (5.5%)	400 (27%)
35-44	649 (16%)	214 (8.6%)	435 (29%)
45-54	779 (20%)	389 (16%)	390 (26%)
55-64	708 (18%)	528 (21%)	180 (12%)
65-74	684 (17%)	673 (27%)	11 (0.73%)
Over 75 and over	497 (12%)	497 (20%)	0 (0%)
Gender	2,549 female (64%)	1,530 female (61%)	1,019 female (68%)
Ethnicity category			
White	3,555 (89%)	2,389 (96%)	1,166 (78%)
Asian	276 (6.9%)	39 (1.6%)	237 (16%)
Black	54 (1.4%)	24 (0.96%)	30 (2.0%)
Mixed	51 (1.3%)	16 (0.64%)	35 (2.3%)
Other	7 (0.18%)	19 (0.76%)	24 (1.6%)
Employment		Retired = 1,175 (47%) Employed = 851 (34%) Self-employed = 171 (6.9%) Unemployed = Carer role(s) = 89 (3.6%)	Nurse = 702 (47%) Doctor = 517 (35%) Allied Health Professional (AHP) = 135 (9.0%) Clinical Support Worker (CSW) = 125 (8.4%)
Surgery experience		1 previous operation = 2,489 (100%) 2 previous operations = 1,404 (56%) ≥3 previous operations = 806 (32%) Carer for relative after major surgery = 153 (6.2%)	General surgery = 1,040 (69%) Trauma & Orthopaedics = 692 (46%) Any other specialty = 765 (51%)

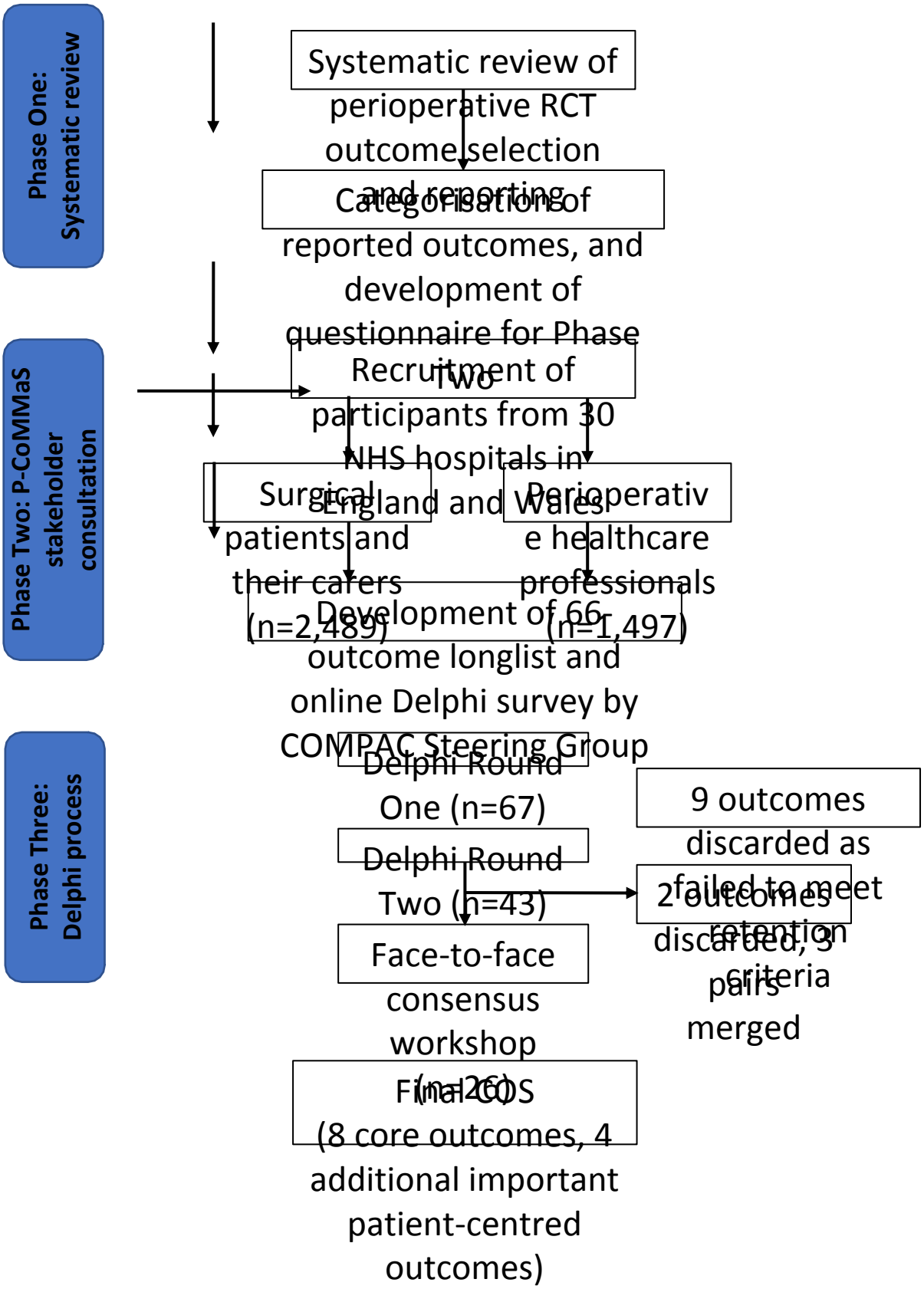
Table 2. Demographic Bbreakdown of respondents to P-COMMaS survey

Outcome domain	Rated 'critically important', all respondents (n=3,986)	Rated 'critically important', service users (n=2,489)	Rated 'critically important', clinicians (n=1,497)	Absolute difference, service users vs clinicians	P value (Chi-square-test)
The operation being successful	96.6%	97.2%	95.7%	1.57%	0.0077
Being comfortable after the operation	86.7%	88.6%	83.6%	4.96%	<0.001
Getting out of hospital ~~ASAP~~as soon as possible	66.6%	70.6%	60.0%	10.6%	<0.001
Returning to normal activities ~~ASAP~~as soon as possible	71.1%	74.4%	65.5%	8.88%	<0.001
Avoiding temporary postoperative complications	94.6%	95.6%	93.1%	2.53%	<0.001
Avoiding permanent postoperative complications	99.3%	99.4%	99.0%	0.44%	0.11
Not dying during or soon after the operation	95.2%	95.5%	94.7%	0.82%	0.24

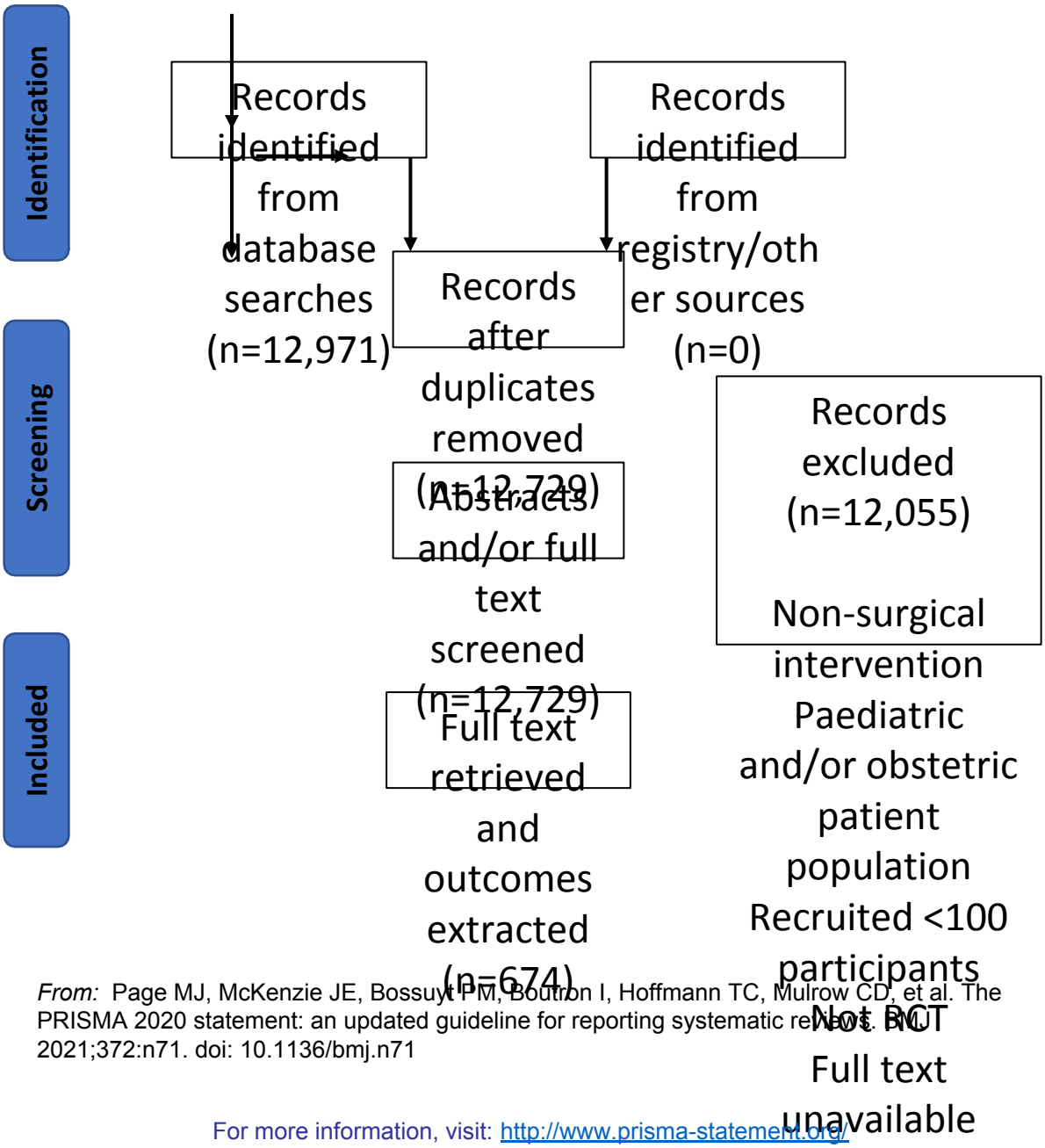
Table 3. Proportion of P-COMMaS survey respondents in each group rating each outcome as critically important (i.e. ≥8/10).

Outcome domain	Core Outcome(s)	Corresponding StEP endpoints
Mortality/survival	Overall mortality (death rate) after an operation	Organ failure & survival
	Overall long-term survival (e.g. after a cancer operation)	Cancer and long-term survival[29]
Peri-operative complications	Major (serious) postoperative complications and adverse events (using accepted, validated definitions of major and minor complications)	Various[26-28, 30, 31]
	Complications and adverse events causing permanent disability or harm	
Resource use	Total number of days spent in hospital for the operation	Health resource use
	Unplanned hospital readmission within 30 days of operation	
Short term recovery after surgery	Discharge destination from hospital (e.g. own home / rehab facility / care home), AND/OR level of dependence (need for carers)	Patient comfort[32]
Longer-term recovery after surgery	Overall health-related quality of life (using a validated scoring tool)	Patient-centred outcomes[25]
Additional important patient-centred outcomes to be considered for inclusion		
Short term recovery after surgery	Pain (incidence/severity/duration)	Patient comfort[32]
	Nausea +/- vomiting (incidence/severity/duration)	
	Mental, emotional and psychological wellbeing	
Overall success/failure of surgery	Patient satisfaction with their operation, and/or willingness (with hindsight) to choose the same again	Patient-centred outcomes[25]

Table 4. Final COMPAC Core Outcome Set, as agreed and ratified at the final workshop


Phase One: Systematic review
Systematic review of perioperative RCT outcome selection and reporting
Categorisation of reported outcomes, and development of questionnaire for Phase Two
Phase Two: P-CoMMaS stakeholder consultation
Recruitment of participants from 30 NHS hospitals in England and Wales
Surgical patients and their carers (n=2,489)
Perioperative healthcare professionals (n=1,497)
Development of 66 outcome longlist and online Delphi survey by COMPAC Steering Group
Phase Three: Delphi process
Delphi Round One (n=67)
9 outcomes discarded as failed to meet retention criteria
Delphi Round Two (n=43)
2 outcomes discarded, 3 pairs merged
Face-to-face consensus workshop (n=26)
Final COS (8 core outcomes, 4 additional important patient-centred outcomes)

PRISMA Flow Diagram:
Core Outcome Measures for Perioperative and Anaesthetic Care (COMPAC) Systematic Review



From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

For more information, visit: http://www.prisma-statement.org/