



**QUEEN'S
UNIVERSITY
BELFAST**

Development of a Core Outcome Set for studies evaluating the effects of anaesthesia on perioperative morbidity and mortality following hip fracture surgery

O'Donnell, C. M., Black, N., McCourt, K. C., McBrien, M. E., Clarke, M., Patterson, C. C., Blackwood, B., McAuley, D. F., & Shields, M. O. (2018). Development of a Core Outcome Set for studies evaluating the effects of anaesthesia on perioperative morbidity and mortality following hip fracture surgery. *British Journal of Anaesthesia*. <https://doi.org/10.1016/j.bja.2018.08.017>

Published in:
British Journal of Anaesthesia

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights

Copyright 2018 Elsevier. This manuscript is distributed under a Creative Commons Attribution-NonCommercial-NoDerivs License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits distribution and reproduction for non-commercial purposes, provided the author and source are cited.

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Development of a Core Outcome Set for studies evaluating the effects of anaesthesia on perioperative morbidity and mortality within the hip fracture population

O'Donnell CM^{1 3}, Black N¹, McCourt KC¹, McBrien ME¹, Clarke M², Patterson CC², Blackwood B³, McAuley DF^{1 3}, Shields MO¹.

¹ Royal Victoria Hospital, Belfast Health and Social Care Trust, Grosvenor Road, Belfast, BT12 6BA

² Centre for Public Health, School of Medicine, Dentistry and Biomedical Sciences, Institute of Clinical Sciences, Block B, Queen's University of Belfast, Belfast BT12 6BA

³ Wellcome-Wolfson Institute for Experimental Medicine, School of Medicine, Dentistry and Biomedical Sciences, Queen's University of Belfast, 97 Lisburn Road, Belfast BT9 7BL

- CM O'Donnell is the corresponding author: codonnell11@qub.ac.uk
- Address for correspondence: Centre for Experimental Medicine, Queen's University Belfast, Wellcome-Wolfson Building, 97 Lisburn Road, Belfast BT9 7BL, Northern Ireland, UK.

Abstract

Rationale: Perioperative studies on hip fractures have large heterogeneity in the outcomes that they report.

Objective: To develop a core outcome set (COS) for use in perioperative studies comparing types of anaesthesia for hip fracture surgery.

Methods: The consensus process consisted of a systematic review of the literature, three rounds of a Delphi survey, two consensus webinars, and face to face patient meetings.

Results: Delphi participants represented 9 stakeholder groups. The number of participants completing rounds 1 to 3 were 242, 186 and 169 respectively. Seventeen outcomes that met predefined consensus criteria were considered at two consensus meetings and a final set of ten core outcomes was agreed: mortality, time from injury to surgery, acute coronary syndrome (ACS), hypotension, acute kidney injury (AKI), delirium, pneumonia, orthogeriatric input, being out of bed at day 1 and pain.

Conclusions: This study generated a consensus-based set of core outcomes recommended for use in all perioperative trials evaluating the effects of anaesthesia for hip fracture surgery. An important next step is developing consensus-based consistency on how they should be measured.

Key Words: perioperative, outcomes, Delphi technique, hip fracture, anaesthesia

There are many outstanding questions regarding the optimal perioperative management of hip fracture patients. One reason why previous clinical studies and subsequent meta-analyses have failed to answer this question may be due to widespread inconsistencies in outcome definitions and reporting.(1) The introduction and adoption of a Core Outcome Set (COS)(2) for all future studies in this field, should provide an opportunity to make more robust comparisons of data from studies.(3) A core outcome set (COS) is defined as “the minimum (number of outcomes) that should be measured and reported in effectiveness trials for a specific condition.”(2) Whilst individual studies may report outcomes above and beyond this minimum set in order to address specific research questions, the use of a COS will allow important baseline outcomes to be consistently described. It will also ensure that outcomes of significance to patients are routinely reported.

The aim of this study was to gain consensus on the outcomes that should be included in a COS for studies comparing types of anaesthesia in the hip fracture population using the Delphi method. A similar approach has been successful for COS development adopted by other fields such as rheumatology(7), cleft palate surgery(8), oncology(9, 10), critical care(9, 11) and most recently upper gastrointestinal surgery.

Methods

The study was developed following guidance from the Core Outcome Measures in Effectiveness Trials (COMET) Handbook(4) and was conducted in accordance with recommendations from Core Outcome Set-Standards for Development (COS-STAD)(5) and Reporting (COS-STAR)(6). The study outline was registered on the COMET database(12). Ethical approval was granted by the research ethics committee, Queen’s University Belfast.

The development of the COS consisted of three stages outlined in Figure 1. For stage one, a systematic review of the literature was performed to identify all perioperative outcomes previously reported in studies comparing modes of anaesthesia for hip fracture surgery.(1) Duplicates were then removed and similar outcomes grouped together into domains by a study advisory group. These domains were: physiological, with outcomes categorized per organ system; mortality; process of care; functional and other. These are similar to domain headings suggested in a recent paper by the COMET group on outcome classification.(13)

An online search of international professional organisations was conducted to identify potential participants from key stakeholder groups that included anaesthetists, orthopaedic surgeons, orthogeriatricians, physiotherapists, occupational therapists and nurses from the field of hip fracture

surgery. Efforts were made to recruit a representative number from each stakeholder group. A standardized email, inviting participation in the Delphi rounds, was sent to individuals or organizational leads for onward distribution. Those wishing to participate were sent a link to the study and once registered, were given a unique identifier to anonymize their responses. Patients were recruited via advertisements on charitable organization websites and posters placed at local amenities.

Stage two commenced with the development of an online Delphi survey involving three sequential rounds using DelphiManager (University of Liverpool). Participants were asked to rate the importance of each outcome from stage one using the Grading of Recommendations Assessment, Development and Evaluations Scale of 1-9; 1 to 3 representing lesser importance, 4 to 6 being important but not critical and 7 to 9 representing critical importance.(14) An option to suggest outcomes in addition to those provided was made available during the first survey round. All outcomes, despite their scores were retained to the end of the study.

The second and third survey rounds provided collective graphical feedback from all stakeholder groups on the distribution of results for each outcome, reminded participants of their previous scores and asked them to rescore the outcomes. Taking part in the previous round was a prerequisite for completing the subsequent rounds, therefore one round could not be completed in isolation. The distribution of scores for each outcome was calculated as a percentage of the total responses. Consensus that an outcome should be considered for inclusion in the COS was defined as a score of 7 to 9 by 70% or more participants and a score of 1 to 3 by 15% or fewer participants.(14) Outcomes that met “consensus in” criteria were presented at the consensus webinars.

Stage three consisted of two consensus webinars. These were held on two separate days to maximize international attendance and to determine concordance between the expert panels. Participants who completed all three rounds of Delphi were listed according to their stakeholder group and continent. The first 5 in each group were invited by email to attend the webinars and if unable to attend, subsequent participants were invited as necessary to produce a consensus panel. An anonymized online voting system was used and results were broadcast immediately. In cases where there was no clear consensus result, a discussion was held and revote taken. Due to the technical barriers for this patient population, face to face meetings were held individually with them at home prior to the consensus meetings and a similar outcome voting process undertaken.

Statistical Analysis

The distribution of scores for each outcome was assessed using the R statistical package (R Foundation for Statistical Computing, Vienna, Austria). Cohen’s Kappa scores were calculated to assess the level of agreement between each Delphi round for all outcomes meeting “consensus”

criteria at the end of round 3 using SPSS for Macintosh, Version 22.0 (IBM Inc., Armonk, NY, USA). Mean and standard deviations were also calculated for these outcomes. To assess for attrition bias, an independent t-test was used to calculate the difference between means for completers and non-completers at each round.

Results

The Delphi study

The systematic literature review yielded 71 outcomes from 32 papers. Twenty seven studies looked at outcomes relating to regional compared to general anaesthesia and 10 studies examined outcomes specifically for spinal compared with general anaesthesia. One study reported outcomes comparing high and low dose spinal anaesthesia, one reported outcomes for deep versus light sedation and another compared spinal versus lumbar and sciatic plexus blocks.(1) Following consolidation of similar outcomes under one heading, a total of 34 discrete outcomes formed round one of the Delphi questionnaire. The Delphi online surveys were conducted between December 2016 and August 2017. Two hundred and forty-two participants completed round one; 189 out of these 242 (78%) participants completed round two; and 169 of the original 242 participants (70%) completed round three. Participants represented nine key stakeholder groups and eight geographical regions (table 1). This included 10 patient participants who completed all three rounds of Delphi; 80% of the patient participants were female, 70% were aged over 80 and ranged from 76 to 102 years, similar to the demographics of patients presenting with hip fractures.(15)

During round one, participants suggested 98 new outcomes (appendix 1), which following review by the study advisory group resulted in 4 additional outcomes being added to round 2. These 4 outcomes encompassed the various permutations suggested. These were as follows; a change in the level of care, application of a Do Not Attempt Cardiopulmonary Resuscitation (DNACPR) decision, the use of palliative care and patient/care giver satisfaction. Table 2 displays all outcomes scored during the Delphi process and those that were rated as critical by at the end of round 3 broken down per stakeholder group. At the end of round 3, 17 outcomes met “consensus” criteria (table 3).

With the exception of pneumonia and sepsis/septic shock, there was an increase in the level of agreement between the Delphi participants regarding which outcomes were important as the rounds progressed (table 3). Between rounds 2 and 3, all of the outcomes meeting “consensus” criteria showed “substantial agreement”, defined by a kappa value of greater than 0.61.(16) 80% of patient participants allocated pain a score of 7-9. Despite narrowly missing “consensus” criteria by the collective group, with just over 65% allocating a score of 7-9 (table 3), pain was retained as an additional outcome for discussion at the webinars due to its importance for patients.

Attrition bias may occur if participants leave the study before all 3 rounds are complete causing the subsequent study sample to no longer resemble the original. These are represented by p values and provide information on the difference between mean scores of those who completed all 3 rounds and those who did not. In this study, there was no attrition bias for round 2; however, there was a small attrition bias demonstrated for one of the outcomes at round 3, cardiac arrest ($p=0.02$) (table 4). The study non-completers scored this outcome slightly higher than the study completers. This did not influence its inclusion in the final list of outcomes that were presented in the consensus meetings.

Consensus meetings

A total of 27 participants took part in the consensus webinars. All stakeholder groups were represented with participants from 5 different countries across 3 continents (appendix 2). The numbers of participants voting outcomes 'in', 'out' and 'uncertain' are presented in table 5. Nine outcomes were chosen for the final COS by both consensus meetings. Pneumonia and sepsis/septic shock were voted for inclusion by participants of only 1 webinar respectively. The study advisory group voted to include only pneumonia. As a complication that may be affected by mode of anaesthesia, pulmonary sepsis was deemed to be the more prevalent source of sepsis to affect this patient population. Individually, other sources of sepsis, urinary tract infections (UTIs) and surgical site infections (SSIs) were voted out during the consensus process.

The final COS consisted of 10 discrete outcomes as follows: mortality, time from injury to surgery, orthogeriatric input, acute coronary syndrome, hypotension, pneumonia, acute kidney injury, delirium, out of bed at day 1 and pain.

Discussion

We suggest that this COS represents the minimum number of outcomes to report in all perioperative studies comparing types of anaesthesia in hip fracture patients. Adoption of this COS should help when developing research protocols and allow better comparison of findings from individual studies.(17, 18) It also widens the scope of studies away from reporting mortality only and should promote more focus on morbidity. The Standard Protocol Items Recommendations for Interventional Trials (SPIRIT) Statement(19) and Consolidated Standards of Reporting Trials (CONSORT) Statement(20), which outline best practice in study design, recommend the use of COSs. Furthermore, the UK National Institute for Health Research Health Technology Assessment (NIHR HTA) funding stream states that where a core outcome set exists, it should be included in the study design.(18)

Previous working groups aiming to achieve standardized endpoints within healthcare,(21) (22) have seldom adopted a formal consensus gathering process such as the one used in this study.(23) A review of healthcare guideline development by Wang and colleagues found that over 30% did not report how consensus was gained and the remaining did so poorly.(24) It is vital to ensure that high standards and consistency regarding methodology are implemented.(25, 26) This can be facilitated by the Delphi method, an objective and structured approach to consensus gathering.(27) . Similar work is ongoing within anaesthesia (25) with the Standardized Endpoints in Perioperative Medicine (STEP) group and Core Outcome Measures in Perioperative Anaesthetic Care (COMPAC) Initiative.(26)

The use of the Delphi method in healthcare related COS development has increased from 12% in 2014 to 46% in 2016.(28) A minimum of two survey rounds, good panel heterogeneity and adequate feedback are key components of a robust Delphi. Rarely have COSs and consensus statements reported using a full Delphi protocol in their methods.(29) Furthermore, 82% of medical COSs to date only represent Europe and North America(28), whereas this study had representation from 6 continents.

The strengths of this study include the use of a heterogeneous, international panel with patient involvement to ensure that the COS did not neglect their opinions. The context of the COS was also well defined in terms of the patient population to whom it applies and the fundamentals of perioperative care involved. Furthermore, this study is novel in its use of statistics to confirm consistency between successive Delphi rounds.(16) (30)

Some controversy arose throughout this study regarding the distinction between processes of care and outcomes. Items within this COS such as time from injury to surgery and the provision of orthogeriatric care were consistently voted in throughout the consensus process due to their potential impact on outcomes. They were ultimately categorized under a *domain* ‘process of care’. As groups have developed COSs,(18) other groups have developed the term “Core Domain Set”.(31) This is a wider definition to adequately measure all relevant concepts and outcomes relating to a specific health condition within a particular setting.(7, 31) This concept has recently been expanded by COMET in a paper proposing a framework of domains to aid COS study design and allow categorisation of outcomes in a standardized way.(13) As such, the list of core outcomes which have been chosen in this study are actually more representative of a Core Domain Set. A COS that is predominantly outcome specific may omit important concepts, yet a Core Domain Set encompasses the complete content of what is measurable in a trial. Therefore, in the future, development of Core Domain Sets to include the relevant outcomes may be a better approach to adopt within anaesthesia and peri-operative medicine.

General limitations of the Delphi method relate to the risks of reduced accountability due to anonymity of views and of volunteer bias. Furthermore, there can be various interpretations of what constitutes a Delphi study. Failure to adhere to robust methodology affects the validity of a study and its ability to reveal true consensus.⁽²⁹⁾ Important factors such as the participant panel, the number of survey rounds, the adequacy of feedback between rounds and the ability of participants to put forward their own views must be clearly considered according to COS-STAR.⁽⁶⁾ The optimal stakeholder panel size and representation is not as yet known, however, guidance recommends maximising the response rate from a range of groups.⁽⁶⁾ In addition, it is important to get the right balance between feasibility of use of the COS, as well as adequate depth of information to discriminate between outcomes of interest. Each of the 10 outcomes in this COS was deemed necessary for inclusion and the majority are not cumbersome to collect. Outcomes which are more difficult to measure and report, such as delirium, are nonetheless vital to include.

Selecting a COS is only part of the process of standardizing outcomes. How each outcome should be defined and measured will play a fundamental role in ensuring that future pooling of data is possible.⁽¹⁷⁾ ⁽³²⁾ ⁽¹⁰⁾ Defining perioperative outcomes is an ongoing focus for the STEP collaboration with two papers in a series of work having been published recently.⁽³³⁾ ⁽³⁴⁾ The majority of our outcomes will be defined by this work with the aims of avoiding duplication and striving for consistency. STEP has already provided a definition for pneumonia.⁽³³⁾ There are four outstanding outcomes within this particular COS which will require definition; mortality, time from injury to surgery, out of bed at day one and hypotension. Mortality and time from injury to surgery will require consensus in terms of the timepoints at which they are measured. Out of bed at day 1 will require clarification with regard to its exact meaning and hypotension will require a robust definition regarding thresholds and factors such as duration. These will be the focus of our future work. This will be done via the same robust consensus process used for this study. It is important to stress that the use of a COS does not preclude the measurement of additional outcomes specific to a particular study. Furthermore, it is important that COSs are regularly updated ⁽³¹⁾ since the validity of an outcome measure, related to its clinical relevance and inter-rater reliability, ⁽¹⁷⁾ is not static over time.

In conclusion, we have used an established Delphi model to develop a COS for future studies examining the effects of anaesthesia in patients having surgery for hip fracture. In the emerging field of COS and Core Domain Set development within anaesthesia and perioperative medicine, it is important that developers adhere to rigorous methodology to provide consensus on clinical outcomes.

Author Contributions

Study concept and design: all authors.

Acquisition and interpretation of data: CO'D, NB.

Statistical analysis: CO'D, CP.

Drafting and revision of manuscript: all authors.

Funding

Department of Anaesthesia, Royal Victoria Hospital, Belfast, UK.

Acknowledgements

The authors thank all those who took part in the Delphi study and the consensus meetings. Their time and expertise has allowed this study to be conducted. They also thank Richard Crew and Paula Williamson from the University of Liverpool for the use of DelphiManager.

Conflicts of interest

Mike Clarke is a member of the COMET Management Group, and has been a co-applicant on funding applications to support COMET.

References

1. O'Donnell CM, McLoughlin L, Patterson CC et al. Perioperative outcomes in the context of mode of anaesthesia for patients undergoing hip fracture surgery: a systematic review and meta-analysis. *Br J Anaes* 2018; 120(1): 37-50.
2. Clarke M, Williamson PR. Core outcome sets and systematic reviews. *Syst Rev* 2016; 5:11.
3. Glasziou P, Altman DG, Bossuyt P et al. Reducing waste from incomplete or unusable reports of biomedical research. *Lancet* 2014; 383:267-76.
4. Williamson PR, Altman DG, Bagley H et al. The COMET Handbook: version 1.0. *Trials* 2017; 18:280.
5. Kirkham JJ, Davis K, Altman DG et al. Core Outcome Set-STAndards for Development: The COS-STAD recommendations. *PLoS Med* 2017;14: e1002447.
6. Kirkham JJ, Gorst S, Altman DG et al. Core Outcome Set-STAndards for Reporting: The COS-STAR Statement. *PLoS Med* 2016; 13:e1002148.
7. Simon LS, Strand V, Bingham CO et al. OMERACT 11: International consensus conference on outcome measures in rheumatology. *J Rheumatol* 2014; 41:145-9.
8. Bruce I, Harman N, Williamson P et al. The management of Otitis Media with Effusion in children with cleft palate (mOMEnt): a feasibility study and economic evaluation. *Health Technol Assess* 2015; 19:1-374.
9. Blackwood B, Ringrow S, Clarke M et al. Core Outcomes in Ventilation Trials (COVenT): protocol for a core outcome set using a Delphi survey with a nested randomised trial and observational cohort study. *Trials* 2015; 16:368.
10. Gourgou-Bourgade S, Cameron D, Poortmans P et al. Guidelines for time-to-event end point definitions in breast cancer trials: results of the DATECAN initiative (Definition for the Assessment of Time-to-event Endpoints in CANcer trials). *Ann Oncol* 2015; 26:2505-6.
11. Turnbull AE, Sepulveda KA, Dinglas VD et al. Core Domains for Clinical Research in Acute Respiratory Failure Survivors: An International Modified Delphi Consensus Study. *Crit Care Med* 2017; 45:1001-10.
12. McAuley D, Shields M, O'Donnell C. Identifying a core outcome set for evaluating perioperative morbidity in the hip fracture population. COMET Initiative. Available from <http://www.comet-initiative.org/studies/details/757>
13. 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.
14. Harman NL, Bruce IA, Callery P et al. MOMENT--Management of Otitis Media with Effusion in Cleft Palate: protocol for a systematic review of the literature and identification of a core outcome set using a Delphi survey. *Trials* 2013; 14:70.
15. White SM, Moppett IK, Griffiths R et al. Secondary analysis of outcomes after 11,085 hip fracture operations from the prospective UK Anaesthesia Sprint Audit of Practice (ASAP-2). *Anaesthesia* 2016; 71:506-14.
16. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med* 2012; 22:276-82.
17. Koroshetz W. A core set of trial outcomes for every medical discipline? *BMJ* 2015; 350:85.

18. Gargon E. The COMET (Core Outcome Measures in Effectiveness Trials) Initiative. *Maturitas* 2016; 91:91-2.
19. Chan AW, Tetzlaff JM, Altman DG et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Ann Intern Med* 2013; 158:200-7.
20. Moher D, Hopewell S, Schulz KF et al. CONSORT 2010 Explanation and Elaboration: Updated guidelines for reporting parallel group randomised trials. *J Clin Epidemiol* 2010; 63:e1-37.
21. Hicks KA, Tcheng JE, Bozkurt B et al. 2014 ACC/AHA Key Data Elements and Definitions for Cardiovascular Endpoint Events in Clinical Trials: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Data Standards. *Circulation* 2015; 132:302-61.
22. Grocott MP. Standardisation of perioperative outcome measures. *Eur J Anaesthesiol* 2015; 32:83-4.
23. 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.
24. Wang X, Chen Y, Yang N et al. Methodology and reporting quality of reporting guidelines: systematic review. *BMC Med Res Methodol* 2015; 15:74.
25. Plumb J, Clissold E, Goddard N, Gill R, Grocott M. Call for consensus on terminology and consistent clinical pathways in relation to anaemia in perioperative patients. *Br J Anaesth* 2016; 116(5):722.
26. Myles PS, Grocott MP, Boney O, Moonesinghe SR, Group CO-S. Standardizing end points in perioperative trials: towards a core and extended outcome set. *Br J Anaesth* 2016; 116:586-9.
27. Hsu. The Delphi technique; making sense of consensus. *Practical Assessment, Research & Evaluation (PARE)* 2007; 12:1-8.
28. Gorst SL, Gargon E, Clarke M, Smith V, Williamson PR. Choosing Important Health Outcomes for Comparative Effectiveness Research: An Updated Review and Identification of Gaps. *PLoS One*. 2016; 11:e0168403.
29. Boukdedid R, Abdoul H, Loustau M, Sibony O, Alberti C. Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. *PLoS One* 2011; 6:e20476.
30. Holeý EA, Feeley JL, Dixon J, Whittaker VJ. An exploration of the use of simple statistics to measure consensus and stability in Delphi studies. *BMC Med Res Methodol* 2007; 7:52.
31. Boers M, Kirwan JR, Gossec L et al. How to choose core outcome measurement sets for clinical trials: OMERACT 11 approves filter 2.0. *J Rheumatol* 2014; 41:1025-30.
32. Altman DG, Bland JM. Assessing Agreement between Methods of Measurement. *Clin Chem* 2017; 63:1653-4.
33. Abbott T.E.F, 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 - 1079

34. 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 - 711