

# BMJ Open Development of core outcome sets for effectiveness trials of interventions to prevent and/or treat delirium (Del-CORs): study protocol

Louise Rose,<sup>1,2,3</sup> Meera Agar,<sup>4</sup> Lisa D Burry,<sup>5</sup> Noll Campbell,<sup>6</sup> Mike Clarke,<sup>7</sup> Jacques Lee,<sup>8</sup> Najma Siddiqi,<sup>9</sup> Valerie J Page,<sup>10</sup> For the Del-CORs group

**To cite:** Rose L, Agar M, Burry LD, *et al.* Development of core outcome sets for effectiveness trials of interventions to prevent and/or treat delirium (Del-CORs): study protocol. *BMJ Open* 2017;**7**:e016371. doi:10.1136/bmjopen-2017-016371

► Prepublication history and additional material for this paper are available online. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2017-016371>).

Received 9 February 2017  
Revised 14 June 2017  
Accepted 24 July 2017

## ABSTRACT

**Introduction** Delirium is a common, serious and potentially preventable condition with devastating impact on the quality of life prompting a proliferation of interventional trials. Core outcome sets aim to standardise outcome reporting by identifying outcomes perceived fundamental for measurement in trials of a specific interest area. Our aim is to develop international consensus on two core outcome sets for trials of interventions to prevent and/or treat delirium, irrespective of study population. We aim to identify additional core outcomes specific to the critically ill, acutely hospitalised patients, palliative care and older adults.

**Methods and analysis** We will conduct a systematic review of published and ongoing delirium trials (1980 onwards) and one-on-one interviews of patients who have experienced delirium and family members. These data will inform Delphi round 1 of a two-stage consensus process. In round 2, we will provide participants their own response, summarised group responses and those of patient/family participants for rescoring. We will randomise participants to receive feedback as proportion scoring the outcome as critical or as group mean responses. We will hold a consensus meeting using nominal group technique to finalise outcomes for inclusion. We will repeat the Delphi process and consensus meeting to select measures for each core outcome. We will recruit 240 Delphi participants giving us 80% power to detect a 1.0–1.5 point (9-point scale) difference by feedback method between rounds. We will analyse differences for subsequent scores, magnitude of opinion change, items retained and level of agreement.

**Ethics and dissemination** We are obtaining research ethics approvals according to local governance. Participation will be voluntary and data deidentified. Support from three international delirium organisations will be instrumental in dissemination and core outcome set uptake. We will disseminate through peer-reviewed open access publications and present at conferences selected to reach a wide range of knowledge users.

## INTRODUCTION

Delirium is a complex syndrome characterised by an acute confusional state with rapid onset, a fluctuating course, circadian

## Strengths and limitations of this study

- Rigorous systematic review and core outcome set development methods that adhere to Cochrane and COMET guidelines.
- Engagement with survivors of delirium during the development of the protocol.
- Support of three international Delirium Societies (American, Australasian and European) will facilitate participant recruitment, dissemination and uptake of our core outcome sets.
- Ability to recruit and retain participants, particularly delirium survivors and their family members. We are using multimodal recruitment strategies and seeking advice from organisations experienced in recruitment and retention of patient/family participants.
- Ability to recruit participants with broad geographical representation.
- Inability to come to consensus on the core outcomes or the measures for these outcomes.
- Important outcomes are identified that are difficult to measure due to the absence of valid and reliable measures.

disturbances and reduced or increased motor activity, as well as changes in cognition, notably in the domains of attention and higher-level thought processing.<sup>1 2</sup> Delirium is a common, serious and potentially preventable source of morbidity, with devastating impacts on the quality of life, and mortality. Delirium impacts all age groups, from infants to the very elderly. This includes patients, including those accessing primary care, resident in nursing homes,<sup>3</sup> those receiving palliative care services<sup>4</sup> and a significant number of hospitalised patients, including the acute and critically ill. Prevalence rates in hospitalised patients range from 25% to 80%.<sup>5–8</sup>

Delirium is not a benign, self-limiting condition. In addition to increased mortality, delirium is associated with prolonged length



CrossMark

For numbered affiliations see end of article.

### Correspondence to

Dr Louise Rose;  
[louise.rose@utoronto.ca](mailto:louise.rose@utoronto.ca)

of stay; higher rates of unintentional device removal; falls and incontinence in the elderly; significant emotional distress for patients, families, caregivers and health-care professionals<sup>9-12</sup> and escalating public healthcare costs.<sup>13-15</sup> Delirium also carries long-term consequences including impaired physical functioning<sup>16 17</sup> and loss of independence resulting in long-term care placement<sup>5</sup> caregiver burden,<sup>18 19</sup> decreased quality of life,<sup>20</sup> cognitive decline and increased risk of dementia and Alzheimer's disease.<sup>11</sup>

With increased recognition of delirium as a common, costly and potentially preventable condition associated with adverse outcomes, encouragingly studies examining interventions to prevent and/or treat delirium continue to proliferate. Currently, there is no systematic approach to the selection and reporting of outcomes and their measures in these studies resulting in reporting of numerous and varied study outcomes and measures for these outcomes. This hinders progress towards improvements in care, as to best inform the evidence base, outcomes must be selected, defined and measured consistently across studies of similar interventions in similar populations. Core outcome sets (COS), developed using rigorous consensus processes involving key stakeholders including patients and carers, comprise outcomes perceived as fundamental to measure in all trials related to a specific and defined area of interest (such as a disease, condition or intervention).<sup>21 22</sup> Although the importance and value of COS for standardising outcomes and measurement across trials is increasingly recognised, in general, they are still in their infancy and as yet have not been developed for trials of interventions to prevent or treat delirium. Therefore, we aim to develop international consensus on two COS appropriate for trials of interventions designed to (1) prevent or (2) treat delirium, irrespective of study population. We also aim to identify additional core outcomes specific to four patient groups: the critically ill, patients requiring hospitalisation in an acute care setting, palliative care and older adults living in residential care or the community.

### Scope of core outcome set development

The scope of our COS will include our four patient populations of interest, considered at high risk of developing delirium.<sup>3 5 23 24</sup> These include (1) critically ill adults and children (medical, surgical and trauma) receiving care in high-acuity settings, including intensive care and high dependency units; (2) non-critically ill adults and children hospitalised in acute care settings, including surgical (all surgeries) and medical patients and patients presenting to an emergency department; (3) adults and children receiving palliative care, either in a hospital, hospice or community setting and (4) older adults ( $\geq 65$  years) living in nursing or residential care homes or living in their own homes and defined as at risk of delirium by study authors. We recognise that certain subpopulations such as children and older adults with dementia spanning these patient populations may need a distinct COS

or outcomes for substitution within a COS. This decision will be made following identification and mapping of outcomes during our systematic review and from interviews with patients/family.

## METHODS

We will use methods outlined in the Outcome Measures in Rheumatology (OMERACT) handbook<sup>25</sup> and those endorsed by the Core Outcome Measures in Effectiveness Trials (COMET) initiative.<sup>26</sup> Our study steering group will comprise two experts with clinical and/or research expertise in delirium and a patient/family representative for each of our four patient populations. We will use the COMET Checklist for Public Research Partners and the COS Study Developers Involved in Designing a COS study checklist<sup>27</sup> to guide and optimise our engagement with patients/family around the COS design and conduct.

### Information sources

We will conduct (1) a systematic review of outcomes and measures reported in published and ongoing trials of interventions to prevent or treat delirium (1980 onwards) and (2) a qualitative study comprising one-on-one interviews with patient survivors, family members and patient advocacy groups to identify outcomes important to patients and families that have experienced delirium.

### Systematic review

#### Search strategy and data sources

We will develop an electronic search strategy through an iterative process informed by an experienced medical information specialist. We will search the following electronic databases, adjusting vocabulary and syntax for each, from 1980 to present: Ovid MEDLINE, Ovid MEDLINE In-Process & Other Non-Indexed Citations, CINAHL, Embase Classic+Embase, PsychINFO and Web of Science (online supplementary file 1). To avoid limiting the scope of outcomes identified, we will not apply a study design filter. We will limit inclusion to studies published in English. A second librarian will review the search strategy prior to execution using the Peer Review for Electronic Search Strategies template.<sup>28 29</sup> We will search for relevant systematic reviews in the Cochrane Library, PROSPERO, and Joanna Briggs and unpublished studies and ongoing trials on the International Clinical Trials Registry Platform (<http://apps.who.int/trialsearch>).

### Study selection

Two investigators will independently screen titles and abstracts for eligible studies. Inclusion criteria include (1) one of the four patient groups of interest; (2) pharmacological and non-pharmacological interventions for delirium prevention, treatment or both; (3) compared with usual care, other pharmacological agents or other non-pharmacological interventions and (4) randomised (individual, cluster and cross-over randomisation), quasi-randomised and non-randomised intervention

studies. If we identify <5 intervention studies in any of the four patient groups, we will expand our inclusion criteria to include observational studies with a control group. We will examine full-text publications of potentially relevant articles for eligibility. We will screen the reference lists of eligible studies and systematic reviews for additional eligible studies for inclusion. We will resolve disagreements through discussion; if unable to achieve consensus, we will refer to an independent arbiter from among the study team.

#### Data extraction

Two investigators will independently extract data from eligible studies on publication date, design, participant characteristics, study objectives, intervention, comparator, outcomes, their definition and measures used to document outcomes.

#### Quality assessment

Two investigators will assess independently the risk of bias using the Cochrane Risk of Bias tool for randomised and quasi-randomised studies and the Scottish Intercollegiate Guidelines Network checklists for non-randomised studies.<sup>30</sup> Two investigators will assess independently the quality of describing and reporting outcomes using the 6-point MOMENT scoring system with a score of  $\geq 4$  representing high-quality outcome reporting.<sup>31</sup> The six elements (each scored as 1 point) include (1) Was the primary outcome stated? (2) Was the primary outcome clearly defined so that another researcher would be able to reproduce its measurement? (3) Were the secondary outcomes clearly stated? (4) Were the secondary outcomes clearly defined? (5) Do the authors explain the choice of outcomes they have selected? and (6) Were methods used to enhance quality of outcome measurement, if

appropriate? We will resolve disagreements through discussion; if unable to achieve consensus, we will refer to an independent arbiter.

#### Data synthesis

We will generate tables of outcomes, their descriptions and measures. We will tabulate the proportion of included studies that report on each outcome and rank order the outcomes accordingly. We will calculate the frequency of the following scenarios: (1) outcomes reported with the same title and definition, (2) outcomes reported with the same title but different definition and (3) outcomes reported with different titles but the same definition. We will then map outcomes to the OMERACT domains.<sup>25</sup> We will use the outcome matrix as recommended by the Outcome Reporting for Brief Intervention Trials (ORBIT) project to organise outcomes.<sup>32</sup> Steering group members will review the outcome list to identify those with similar wording or meaning to be reduced to a single outcome for the purposes of the Delphi round 1 questionnaire.

#### Qualitative study

We will conduct patient and family member interviews as evidence indicates that they may hold different views about which outcomes are of relevance compared with healthcare professionals.<sup>33</sup>

#### Study sample

We will use purposive<sup>34</sup> and maximum variation sampling<sup>35</sup> to identify patient and family participants with the characteristics shown in [table 1](#) (minimum of one representative of each characteristic) for each patient group. For the patient groups representing high-acuity settings, acute care settings and palliative care, we will also target parents and where possible children who have experienced

**Table 1** Stakeholder sampling characteristics

Stakeholder group	Characteristic
Patients/family members*	Age, years ( $\leq 65$ and $> 65$ )
	Sex (male and female)
	Partner status (has partner and no partner)
	Country of residence (North America, Europe/UK, Australasia and other)
Expert clinicians†	Profession (physician, nurse and allied health professionals)
	Years of relevant clinical experience ( $< 5$ , 5–10 and $> 10$ )
	Country of residence (North America, Europe/UK, Australasia and other)
Trialists/researchers‡	Stage of research career (early: $< 5$ years, mid: 5–15 years and senior $> 15$ years)
	Country of residence (North America, Europe/UK, Australasia and other)

\*Patients who survived delirium within the last 18 months and family members who had direct contact with patients while experiencing delirium within the last 18 months irrespective of survival, that is, we will interview family members of patients who did and did not survive the intensive care unit.

†Physicians, nurses and allied health professionals who do not meet the criteria of a trialist.

‡Authors of published (over the last 10 years) or ongoing clinical trials evaluating interventions aimed at preventing or treating delirium.

delirium. For pragmatic reasons related to resource availability, we will only be able to recruit participants fluent in English. We will recruit a sample of 15–20 participants for each patient group which should be sufficient to achieve saturation.<sup>36</sup> We will adjust our sample size using a stopping criterion of three consecutive interviews with no additional material to terminate data collection.

#### Data collection

An experienced qualitative researcher will conduct semi-structured telephone interviews enabling representation across a wide geographical area. Following clarification of what a study outcome is and the importance of COS, patient/family members will be asked to suggest the outcomes of relevance to them when considering their experience of delirium, why these outcomes are important and to identify which outcomes they would consider core and why. All interviews will be audio-recorded and transcribed for analysis.

#### Data analysis

The experienced qualitative researcher and study investigator will independently examine interview transcripts using content analysis methods.<sup>37</sup> Outcomes that do not duplicate those identified from the systematic review will be categorised into domains and noted as only being identified by patients/family. Discussion with another investigator and the patient/family representative on the steering committee will confirm outcomes are of relevance, not duplicative and allocated to the appropriate domain.<sup>38</sup>

#### Delphi consensus building exercise

##### Participants, recruitment and sample size

We will use the eligibility criteria and sampling strategy shown in [table 1](#) ensuring a minimum of two participants from each stakeholder group (patients/family members, expert clinicians and trialists/researchers) representing each of the demographic variables and categories within those variables. If required, we will modify our recruitment advertising to target individuals meeting our demographic targets. For the patient groups representing high-acuity settings, acute care settings and palliative care, we will also aim to have a minimum of five participants representing paediatrics in each group if deemed appropriate to combine in the same COS development process following our systematic review work. We will aim to maintain a minimum of 20 participants representing each stakeholder group (total of 60 participants) for each patient population (total of 240 participants) throughout Delphi rounds (R). Based on an estimated attrition of 30% across rounds, we will target recruitment of 310 participants. A sample size of 240 participants will give us 80% power with two-sided test at  $\alpha=5\%$  to detect a difference of 1.0–1.5 points between rounds on the Grading of Recommendations Assessment, Development and Evaluations (GRADE) scale<sup>39</sup> (range 1–9) by feedback groups when SDs for change vary between 1.4 and 4.1. A

priori we anticipate there may be differences in responses provided by patients and family members compared with those provided by clinicians and researchers. We will test for interaction and if significant, examine each group separately.

We will recruit expert clinicians using recruitment flyers sent through membership lists of the European, American and Australasian Delirium Associations/Societies as well as professional societies of clinicians treating our patient groups. We will continue to enrol participants until our sample size and maximum variation targets are met. We will send personalised recruitment emails to all trialists/researchers identified via our systematic review. As patient/family recruitment may possess challenges, we will use a multimodal strategy including contact with relevant patient/family support/advocacy groups/charities and generic organisations such as the James Lind Alliance and COMET, use of social media including twitter and patient-focused Facebook pages, advertisements placed on public and patient involvement websites, hospital patient engagement and patient and public involvement groups, snowballing techniques and personal contacts.

#### Round 1

We will include all outcomes identified through our systematic review and patient/family interviews. We will describe outcomes in lay terms, with medical terms in brackets, to improve comprehensibility by all. We will seek advice from our patient/family steering group members for lay descriptions. To introduce the Delphi, we will provide plain language summaries developed by COMET. We will programme the Delphi using the online e-management system such as the one developed by COMET. Prior to execution, we will pilot the questionnaire with eight individuals (patients, family members, healthcare professionals and trialists) to assess face validity, understanding and acceptability.

We will provide participants with outcomes identified through systematic review and interviews common to all four patient groups. In addition, we will provide those outcomes specific to one of our four patient groups *only* to participant representatives of that group. We will ask participants to score each outcome using the GRADE scale<sup>39</sup> which ranges from 1 to 9 (1 to 3=not important for inclusion, 4 to 6=important but not critical and 7 to 9=critical for inclusion). We selected this scoring system to facilitate maximum discrimination between questionnaire items as noted by COMET<sup>40 41</sup> and to enable testing of our methodological hypotheses. To avoid presentation bias, we will randomise outcome presentation for each participant. We will provide the opportunity to add additional outcomes. We will send three email completion reminders at 2-week intervals. We will collect demographic information to describe our study sample and to provide each respondent with a unique identifier, enabling personalised reminders for completion of subsequent rounds.

We will examine data distribution of importance scores attributed to each outcome and calculate the mean and SD. We will determine the proportion of participants rating each outcome as 7–9, 4–6 and 1–3. To reduce participant burden, we will retain for R2 those items scored between 7 and 9 (critical importance) by  $\geq 50\%$  and between 1 and 3 (not important) by  $< 15\%$  of respondents. We will apply these criteria separately for patient group.

### Round 2

The steering group will review any additional outcomes provided in R1 to determine whether they represent new outcomes for inclusion and to ensure wording is understandable by all participants. We will provide participants with their own R1 response, summarised responses according to their patient population group, summarised responses of patient/family member participants (also according to patient group), and ask them to rescore the importance of each outcome. We will provide any new outcomes from R1 for scoring on the 1–9 importance scale. As with R1, we will send three email completion reminders at 2-week intervals.

If new outcomes are identified in R1, we will conduct a third round comprising *only* these items to enable two rounds of importance scoring. Items to be brought forward to the consensus meeting will be those scored between 7 and 9 by  $\geq 70\%$  of participants and between 1 and 3 by  $< 15\%$  of participants. We will identify items separately for (1) patients/family and (2) healthcare professionals and researchers combined.

In the event of significant attrition (defined as loss of more than 30% of participants within a stakeholder group) between rounds 1 and 2, we will engage in additional recruitment for Round 2. A priori we anticipate this may be particularly problematic for patients in the palliative group.

### Nested methodological studies

We will conduct nested methodological studies to:

1. Determine whether Delphi feedback provided as the proportion of participants scoring the outcome between 7 and 9 (indicating critical for inclusion) as opposed to mean scores influences subsequent scores, magnitude of change, items retained and level of agreement (overall and by patient population group).
2. Qualitatively explore the process of patient/family engagement and participation throughout COS development to determine barriers and facilitators as well as modification of our processes if needed.
3. Determine whether Delphi versus nominal group technique influences which measures are retained for outcomes included in the COS.

### Delphi feedback

#### Randomisation and allocation

We will randomise (1:1 stratified by patient population group) using a computer-generated schedule developed

by the study statistician. We will generate a questionnaire for each Delphi participant using this allocation schedule. Participants will be randomised to receive feedback as either the proportion of participants scoring the outcome as critical (for their patient population group) or patient population group mean response (figure 1).

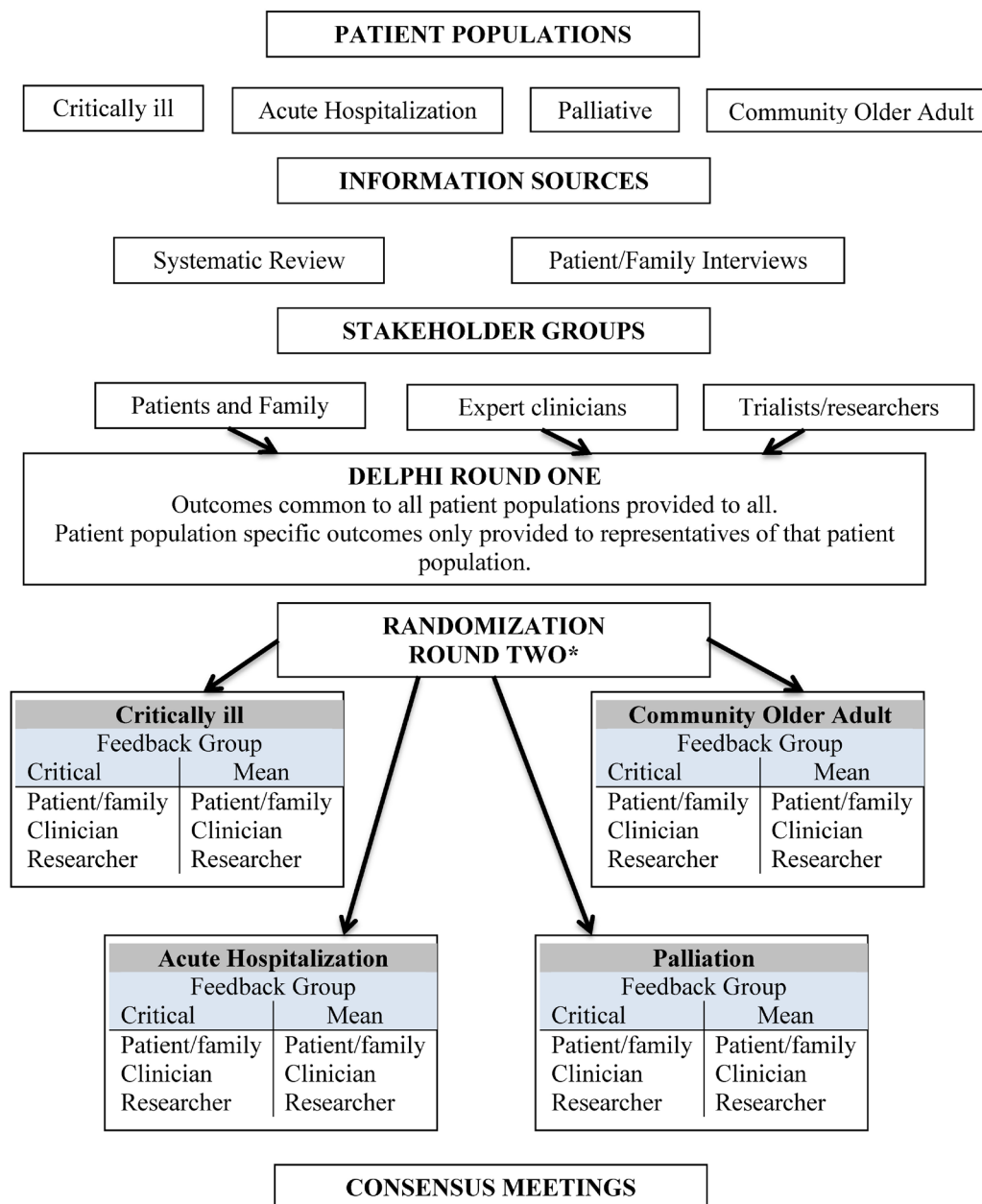
### Statistical analysis

We will analyse differences between feedback groups in terms of: (1) subsequent scores and magnitude of opinion change, (2) items retained at Delphi end and (3) level of agreement between patient population groups. We will calculate the percentage of items for which a participant changed their score between rounds and the mean absolute change in score (ignoring direction of change). We will compare results according to randomisation group using an independent t-test overall and by patient population group. For each outcome, we will use linear regression to compare R2 scores between feedback groups and among patient population groups, adjusting for R1 scores and testing for the interaction between feedback groups and patient populations.

To ascertain feedback group differences for items retained, we will create contingency tables, for each feedback group, to categorise the number of items retained by (1) both feedback groups, (2) critical response feedback group only, (3) mean feedback group only and (4) neither. We will determine the percentage of items for which there was agreement to retain and percentage of discordant items retained by only one feedback group. To ascertain differences between feedback groups in terms of the level of agreement of items retained across patient population groups, we will generate contingency tables categorising number of items retained by (1) all, (2) one group only, (3) two groups only, (4) three groups only and (5) none. We will calculate the percentage agreement and percentage of discordant items.

To explore the impact of feedback on consensus between patient population groups, we will transform the unit of analysis to be the questionnaire item (outcome), with each observation an aggregate statistic. We will use linear regression to determine, for each outcome and for each feedback group, the absolute difference (ignoring direction) in mean R2 scores, adjusting for the participant's R1 score. We will compare absolute mean differences between patient population groups across outcomes between the two feedback groups using a paired t-test. Finally, we will compare responses irrespective of patient population groups within each randomisation arm, calculate the SD for each outcome separately for R1 and R2 and calculate the reduction in each outcome's variability between rounds. We will compare mean reductions in SD across all outcomes between feedback groups using a paired t-test.

Given the anticipated number of statistical tests, we expect 5% to result in a p-value  $\leq 0.05$  by chance; therefore, we will examine the percentage of tests with  $p \leq 0.05$  in relation to this expected percentage.



**Figure 1** Flow of core outcome set development. \*We will conduct a third Delphi round if additional outcomes are identified by participants in round 1.

### Outcome consensus meeting

We will hold consensus meetings to determine the outcomes for inclusion in the two COS's; prevention and/or treatment of delirium irrespective of patient population. We will identify additional outcomes for inclusion in COS specific to our four patient populations. We will aim to be as representative of all stakeholders as possible<sup>42</sup> as we anticipate there may be differences between stakeholder groups in the priority given to outcomes. To ensure we have meaningful input across participant groups, we will invite Delphi participants to attend the meeting. Owing to the large size of our Delphi panel, we will randomly select eight participants to represent each of the stakeholder groups; two representing each patient population group.

We will provide the consensus panel with outcomes established as critical using either method of feedback for inclusion via the Delphi across all four patient groups. We will use a modified nominal group technique to work towards consensus that includes small and whole group discussion and ranking. Ranking will be discussed with the aim of agreeing on the top four or five outcomes across all patients and the top one to two specific to each patient population. To ensure there is no duplication in the final proposed set, each outcome will be discussed to ensure it relates to a distinct construct. If required, we may hold an additional or stand-alone consensus meeting for patients and family members to enable facilitation of their understanding and thus informed voting on outcomes for the COS.

**Table 2** Study timeline

Key project milestones	Start date	End date
Systematic review	May 2017	February 2018
Patient and family member interviews	October 2017	February 2018
Delphi consensus and nested methodological study	April 2018	March 2019
Consensus meeting	June 2019	June 2019
Process evaluation patient/family interviews	February 2018	March 2020
Consensus on measures	September 2019	March 2020
Knowledge translation/dissemination	June 2019 onwards	

### Process evaluation of patient/family participant engagement

We will conduct a process evaluation of patient/family participant engagement and participation throughout COS development.

#### Participants

We will recruit participants to take part in semistructured interviews to determine barriers and facilitators to participation as well as recommendations for improvement strategies to inform future COS development. We will recruit 15–20 participants. Interviews will be audio-recorded and transcribed verbatim.

#### Analysis

We will analyse interview transcripts using content analysis<sup>43–46</sup> employing an inductive, four-step content analysis process.<sup>47</sup> An experienced qualitative researcher and an investigator will independently identify, code and categorise important meanings and predominant themes from the text. Following an immersive reading of the transcripts, initial patterns and recurring categories will be identified by relevant highlighting sections. The second step will seek similarities and differences between participant accounts. Third and fourth steps involve creation of codes and their application over the volume of interviews, respectively. The larger team will be involved in in-depth reading of the coding to ensure credibility. NVivo 10 software will be used for all facets of the analysis.

#### Instruments to measure outcomes

During our systematic review, we will also extract measures for outcomes reported in studies meeting our inclusion criteria. We will assess the measurement and psychometric properties of measures of the outcomes selected for our two COS (prevention and treatment) using the CONsensus-based Standards for the selection of health Measurement INSTRUMENTS (COSMIN) checklist.<sup>48</sup>

#### Participants

We will invite all Delphi participants involved in the establishment of the COS to participate in a second Delphi to establish measures for these outcomes. We will recruit additional participants if required due to attrition. We will recruit an additional 24 participants to take part in a separate consensus building exercise using only a modified nominal group technique to address the following

hypothesis: measures selected for the COS are influenced by the method used for consensus building (Delphi versus nominal group technique).

#### Procedures

We will use the same Delphi methods as described above to establish one set of measures each for the two COSs (prevention and treatment) including the same nested study design of randomisation to two feedback methods (Delphi group). We will provide ‘measure cards’ providing standardised descriptions of the measures, psychometric properties and feasibility of use (ie, time to complete and number of items) in language understandable to all participants. We will use the same nominal group technique methods as described for the COS consensus meetings to establish a second set of measures (nominal group technique group). We will provide to the same description of the measures and their psychometric properties as provided to the Delphi method group. In addition, we will invite a psychometrician, clinicians and/or researchers with familiarity with the measures to the nominal group technique group, thus enabling informed discussion.

#### Statistical analysis

We will perform the same statistical analyses as described for the COS Delphi to determine differences related to consensus group method. To ascertain differences in terms of measures retained between consensus group methods, we will create contingency tables to categorise number of items retained on completion by (1) both Delphi and nominal group technique groups, (2) Delphi group only, (3) nominal group technique only and (4) neither. We will determine the percentage of items for which there was agreement along with the percentage of discordant items, retained by one consensus group method but not the other.

#### Final consensus

We will hold a final steering group meeting to review the findings of the consensus building exercises. Depending on the number of measures rated as critical to include, we will hold a second consensus meeting using the methods described above to guide final decisions (see [table 2](#) for study timeline).

## ETHICS AND DISSEMINATION

We have received research ethics board approval from the University of Toronto (no 34296) and will seek others as required by local governance. We will obtain written consent from participants in interviews and consensus meetings. Participation in Delphi rounds will be considered indicative of consent. Consent will emphasise the voluntary nature of participation and anonymity.

Knowledge users within our investigator team and the support of three international Delirium Societies (American, Australasian and European) will be instrumental in dissemination of the COSs and subsequent uptake. We will provide a one page summary (clinicians/researchers and in lay language for patients and families) to these Societies for distribution among their networks and engage with them to seek additional opportunities to present our findings (educational seminars/workshops). We will disseminate our findings through peer-reviewed and open access publications and presentations at international conferences purposefully selected to reach a wide range of knowledge users taking into account geographical locations. We will engage with journal editors and funding agencies to promote awareness of our COSs.

### Author affiliations

<sup>1</sup>Department of Critical Care Medicine, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada

<sup>2</sup>Lawrence S. Bloomberg Faculty of Nursing, University of Toronto, Toronto, Ontario, Canada

<sup>3</sup>Interdepartmental Division of Critical Care Medicine, University of Toronto, Toronto, Ontario, Canada

<sup>4</sup>Faculty of Health Sciences, University of Technology, Sydney, New South Wales, Australia

<sup>5</sup>Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, Ontario, Canada

<sup>6</sup>College of Pharmacy, Indiana University-Purdue University, Indianapolis, Indiana, USA

<sup>7</sup>School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, Belfast, UK

<sup>8</sup>Sunnybrook Research Institute, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada

<sup>9</sup>School of Medicine, York University, York, UK

<sup>10</sup>Department of Anaesthetics, Watford General Hospital, Watford, UK

**Collaborators** Ansell J, Blackwood B, Jones R, Featherstone I, Fornear L, Hosie A, Johnson M, Needham D, Schofield N, Taylor J, Lawlor P

**Contributors** LR and VJP conceived the study. LR, MA, NS and VJP developed the study protocol. MC advised on the design of the nested methodological studies as well as contributed to the design of the protocol. All authors edited the manuscript and read and approved the final version.

**Funding** This Core Outcome Set is registered on the COMET website (<http://www.comet-initiative.org/studies/details/796>). The systematic review is registered on PROSPERO-ID: CRD42016052704 ([https://www.crd.york.ac.uk/PROSPERO/display\\_record.asp?ID=CRD42016052704](https://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42016052704)). The study is funded by the Canadian Institutes of Health Research. The study has received Research Ethics Board approval from the University of Toronto (no 34296).

**Competing interests** None declared.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Open Access** This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

## REFERENCES

1. American Psychiatric Association. *Diagnostic and statistical manual*. 5th edn. Washington DC: American Psychiatric Association, 2013.
2. Trzepacz P, Meagher D. *Oxford Textbook of Psychiatry*. 2nd edn. UK: LondonOxford University Press, 2009.
3. Siddiqi N, Young J, House AO, *et al*. Stop delirium! A complex intervention to prevent delirium in care homes: a mixed-methods feasibility study. *Age Ageing* 2011;40:90–8.
4. Hosie A, Davidson PM, Agar M, *et al*. Delirium prevalence, incidence, and implications for screening in specialist palliative care inpatient settings: a systematic review. *Palliat Med* 2013;27:486–98.
5. Inouye SK, Westendorp RG, Saczynski JS. Delirium in elderly people. *Lancet* 2014;383:911–22.
6. Mistràletti G, Carloni E, Cigada M, *et al*. Sleep and delirium in the intensive care unit. *Minerva Anestesiol* 2008;74:329–33.
7. Daoud A, Duff JP, Joffe AR, *et al*. Diagnostic accuracy of delirium diagnosis in pediatric intensive care: a systematic review. *Crit Care* 2014;18:489.
8. Bellapart J, Boots R. Potential use of melatonin in sleep and delirium in the critically ill. *Br J Anaesth* 2012;108:572–80.
9. Zaal IJ, Slooter AJ. Delirium in critically ill patients: epidemiology, pathophysiology, diagnosis and management. *Drugs* 2012;72:1457–71.
10. Bickel H, Grading R, Kochs E, *et al*. High risk of cognitive and functional decline after postoperative delirium. A three-year prospective study. *Dement Geriatr Cogn Disord* 2008;26:26–31.
11. O'Malley G, Leonard M, Meagher D, *et al*. The delirium experience: a review. *J Psychosom Res* 2008;65:223–8.
12. Corsinovi L, Bo M, Riccauda Aimonino N, *et al*. Predictors of falls and hospitalization outcomes in elderly patients admitted to an acute geriatric unit. *Arch Gerontol Geriatr* 2009;49:142–5.
13. Leslie DL, Marcantonio ER, Zhang Y, *et al*. One-year health care costs associated with delirium in the elderly population. *Arch Intern Med* 2008;168:27–32.
14. Barr J, Fraser GL, Puntillo K, *et al*. Clinical practice guidelines for the management of pain, agitation, and delirium in adult patients in the intensive care unit. *Crit Care Med* 2013;41:263–306.
15. Traube C, Mauer EA, Gerber LM, *et al*. Cost associated with pediatric delirium in the ICU. *Crit Care Med* 2016;44:e1175–9.
16. Buurman BM, Hoogerduijn JG, de Haan RJ, *et al*. Geriatric conditions in acutely hospitalized older patients: prevalence and one-year survival and functional decline. *PLoS One* 2011;6:e26951.
17. Rudolph JL, Inouye SK, Jones RN, *et al*. Delirium: an independent predictor of functional decline after cardiac surgery. *J Am Geriatr Soc* 2010;58:643–9.
18. Carbone MK, Gugliucci MR. Delirium and the family caregiver: the need for evidence-based education interventions. *Gerontologist* 2015;55:345–52.
19. Finucane AM, Lugton J, Kennedy C, *et al*. The experiences of caregivers of patients with delirium, and their role in its management in palliative care settings: an integrative literature review. *Psychooncology* 2017;26:291–300.
20. Crocker E, Beggs T, Hassan A, *et al*. Long-term effects of postoperative delirium in patients Undergoing cardiac operation: a systematic review. *Ann Thorac Surg* 2016;102:1391–9.
21. Clarke M. Standardising outcomes for clinical trials and systematic reviews. *Trials* 2007;8:39.
22. Sanderson T, Morris M, Calnan M, *et al*. What outcomes from pharmacologic treatments are important to people with rheumatoid arthritis? Creating the basis of a patient core set. *Arthritis Care Res* 2010;62:640–6.
23. Gagnon PR. Treatment of delirium in supportive and palliative care. *Curr Opin Support Palliat Care* 2008;2:60–6.
24. LaHue SC, Liu VX. Loud and clear: sensory impairment, delirium, and functional recovery in critical illness. *Am J Respir Crit Care Med* 2016;194:252–3.
25. Tugwell P, Boers M, Brooks P, *et al*. OMERACT: an international initiative to improve outcome measurement in rheumatology. *Trials* 2007;8:38.
26. Williamson PR, Altman DG, Blazeby JM, *et al*. Developing core outcome sets for clinical trials: issues to consider. *Trials* 2012;13:132.
27. COMET PoPIE group. Checklist for public research partners and core outcome set (COS) study developers involved in designing a COS study. Secondary Checklist for public research partners and core outcome set (COS) study developers involved in designing a



- COS study. 2015 <http://www.comet-initiative.org/assets/downloads/COS%20study%20development%20-%20PPI%20checklist%2016-03-16.pdf>
28. McGowan J, Sampson M, Salzwedel DM, *et al.* PRESS peer review of electronic search strategies: 2015 guideline statement. *J Clin Epidemiol* 2016;75:40–6.
  29. Sampson M, McGowan J, Cogo E, *et al.* An evidence-based practice guideline for the peer review of electronic search strategies. *J Clin Epidemiol* 2009;62:944–52.
  30. Scottish Intercollegiate Guidelines Network. Methodology checklists. 2015 <http://www.sign.ac.uk/checklists-and-notes.html>.
  31. 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.
  32. Kirkham JJ, Dwan KM, Altman DG, *et al.* The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ* 2010;340:c365.
  33. Boers M, Kirwan J, Tugwell P, *et al*; *The OMERACT handbook*. Ottawa: OMERACT, 2014.
  34. Pope C, Mays N. Reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. *BMJ* 1995;311:42–5.
  35. Patton M. Purposeful sampling. In: *Qualitative evaluation and research methods*. Beverly Hills, CA: Sage, 1990:169–86.
  36. Francis JJ, Johnston M, Robertson C, *et al.* What is an adequate sample size? Operationalising data saturation for theory-based interview studies. *Psychol Health* 2010;25:1229–45.
  37. Elo S, Kyngäs H. The qualitative content analysis process. *J Adv Nurs* 2008;62:107–15.
  38. Macefield RC, Jacobs M, Korfage IJ, *et al.* Developing core outcomes sets: methods for identifying and including patient-reported outcomes (PROs). *Trials* 2014;15:49.
  39. The GRADE working group. GRADE. 2016 <http://www.gradeworkinggroup.org/>.
  40. Sinha IP, Smyth RL, Williamson PR. Using the Delphi technique to determine which outcomes to measure in clinical trials: recommendations for the future based on a systematic review of existing studies. *PLoS Med* 2011;8:e1000393.
  41. The COMET Initiative. The COMET Initiative. 2011 <http://www.comet-initiative.org>.
  42. Simon LS, Strand V, Bingham CO, *et al.* OMERACT 11: international consensus conference on outcome measures in rheumatology. *J Rheumatol* 2014;41:145–9.
  43. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse Educ Today* 2004;24:105–12.
  44. Onwuegbuzie AJ, Dickinson WB, Leech NL, *et al.* A qualitative framework for collecting and analyzing data in focus group research. *Int J Qual Methods* 2009;8:1–21.
  45. Carlsen B, Glenton C. What about N? A methodological study of sample-size reporting in focus group studies. *BMC Med Res Methodol* 2011;11:26.
  46. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nurs Health Sci* 2013;15:398–405.
  47. Burnard P. Teaching the analysis of textual data: an experiential approach. *Nurse Educ Today* 1996;16:278–81.
  48. Terwee CB, Mokkink LB, Knol DL, *et al.* Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res* 2012;21:651–7.