ORIGINAL CLINICAL REPORT

OPEN

Measures for the Core Outcome Set for Research Evaluating Interventions to Prevent and/or Treat Delirium in Critically Ill Adults: An International Consensus Study (Del-COrS)

Louise Rose, PhD, MN, BN[1]
Bronagh Blackwood, PhD, RN[2]
Dale M. Needham, MD, PhD[3]
John W. Devlin, PharmD[4,5]
Mike Clarke, PhD[6]
Lisa D. Burry, PharmD[7,8]
for the Del-COrS group

OBJECTIVES: To gain consensus on measurement methods for outcomes (delirium occurrence, severity, time to resolution, mortality, health-related quality of life [HrQoL], emotional distress including anxiety, depression, acute stress, and post-traumatic stress disorder, and cognition) of our Core Outcome Set (COS) for trials of interventions to prevent and/or treat delirium in critically ill adults.

DESIGN: International consensus process.

SETTING: Three virtual meetings (April 2021).

PATIENTS/SUBJECTS: Critical illness survivors/family, clinicians, and researchers from six Countries.

INTERVENTIONS: None.

MEASUREMENTS AND MAIN RESULTS: Measures (selected based on instrument validity, existing recommendations, and feasibility) and measurement time horizons were discussed. Participants voted on instruments and measurement timing (a priori consensus threshold ≥ 70%). Eighteen stakeholders (28% ICU survivors/family members) participated. We achieved consensus on the Confusion Assessment Method-ICU or Intensive Care Delirium Screening Checklist to measure delirium occurrence and delirium resolution (100%), Hospital Anxiety and Depression Scale for emotional distress (71%), and Montreal Cognitive Assessment-Blind for cognition (83%). We did not achieve consensus on EQ-5D five-level for HrQoL (69%) or its measurement at 6 months. We also did not achieve consensus on the Impact of Event Scale (IES)–Revised or IES-6 for post-traumatic stress (65%) or on measurement instruments for delirium severity incorporating delirium-related emotional distress. We were unable to gain consensus on when to commence and when to discontinue assessing for delirium occurrence and time to resolution, when to determine mortality. We gained consensus that emotional distress and cognition should be measured up to 12 months from hospital discharge.

CONCLUSIONS: Consensus was reached on measurement instruments for four of seven outcomes in the COS for delirium prevention or treatment trials for critically ill adults. Further work is required to validate instruments for delirium severity that include delirium-related emotional distress.

KEY WORDS: clinical trials; core outcome set; critical care; delirium; intensive care; outcome measure instruments

A core outcome set (COS) is a minimum set of outcomes recommended to measure and report when undertaking research for specific health conditions or interventions (1). Delirium is highly prevalent in critically adults, with a range of negative outcomes both during and after ICU

DOI: 10.1097/CCE.0000000000000884

KEY POINTS

Question: What measurement methods should be used for the outcomes (delirium occurrence, severity, time to resolution, mortality, health-related quality of life, emotional distress including anxiety, depression, acute stress, and post-traumatic stress disorder, and cognition) of our core outcome set (COS) for trials of interventions to prevent and/or treat delirium in critically ill adults?

Findings: Using three consensus meetings with 18 delirium experts including patient and family representatives, we reached consensus on a measurement instrument for four of the six outcomes that require a measurement instrument (mortality does not require an instrument); Confusion Assessment Method-ICU or Intensive Care Delirium Screening Checklist to assess for delirium occurrence and for resolution; Hospital Anxiety and Depression Scale for anxiety and depression; and Montreal Cognitive Assessment-Blind for cognition.

Meaning: Researchers now have guidance on selection of measures for four of the seven core outcomes of the COS when designing trials of interventions to prevent and/or treat delirium in critically ill adults.

warranting further research (2, 3). However, the lack of a COS for delirium was an important research gap identified in the 2017 Intensive Care Delirium Research Agenda (4). Therefore, in 2021, the Del-COrS study team used rigorous methods (5) and included 179 international participants across three key stakeholder groups (patients and family, clinicians, researchers) to develop a COS for future trials of interventions to prevent and/or treat delirium for critically ill adults. The core outcomes were: delirium occurrence, severity, and time to resolution; mortality; health-related quality of life (HrQoL); emotional distress including anxiety, depression, acute stress, and post-traumatic stress disorder; and cognition including memory (6). This COS is endorsed by the American and Australian Delirium Societies and European Delirium Association.

Outcomes selected for a COS should not be restricted to only those outcomes for which validated measures are available and deemed feasible to use. Instead, selection should be driven by those outcomes that matter most to relevant stakeholders. However, how the outcome should be best measured is important to define to avoid measurement heterogeneity and reduce barriers to evidence synthesis. In a complex research field such as delirium in the critically ill, multiple potential assessment instruments frequently exist; indeed, more than 10 different ICU delirium assessment tools have been published (7).

A core outcome measurement set (COMS) describes how to measure the outcomes chosen for a COS. Within critical care, three COMS have been published for COS designed to inform trials of interventions influencing mechanical ventilation duration (8), acute respiratory failure survivors' post-discharge outcomes (9), COVID-19 (10), and post-COVID condition (11). While some outcomes of these COS overlap with those of Del-COrS (e.g., mortality, HrQoL, cognition), core measures recommended for these nondelirium COMS should not automatically be adopted for delirium trials. Instead, they should be carefully considered in the context of delirium and its longer-term consequences.

Therefore, our study aimed to develop consensus-based recommendations on measurement instrument(s) and measurement timing for each outcome in the COS for clinical trials of interventions designed to prevent and/or treat delirium in critically ill adults.

METHODS

Information Sources

To inform briefing documents for the COMS consensus meetings, we developed a list of outcome measurement instruments and measurement time horizons for each out the outcomes of the Del-COrS COS from our previous systematic review of published trials of interventions to prevent and/or treat delirium in the critically ill (12). We identified measurement instruments and measurement time horizons recommended in existing critical care-specific COMS (8–10) for those outcomes that overlap with the Del-COrS COS (i.e., HrQoL, cognition, and mortality [measurement time horizon only]). For the outcome of cognition, the COMS for acute respiratory failure survivors (www.improveLTO.com) had failed to reach consensus on an appropriate measure yet recommended the Montreal Cognitive Assessment (MoCA)-Blind (13) with the caveat that further evaluation of measurement properties in ICU survivors is required. Therefore, we included

this measure and the background for its inclusion in our briefing documents. Discussion was held among the investigator team to identify any additional measurement instruments that should be included in briefing documents for consideration at the consensus meetings.

Consensus meeting briefing documents (**Supplementary Material**, http://links.lww.com/CCX/B159), emailed ahead of the meetings, outlined the background and scope of the Del-COrS COMS including the overarching aim to select instruments that are valid, discriminative, and feasible to use in a critically ill adult population. In the briefing document, we provided background on each of the proposed measurement instruments and a yes/no assessment checklist as to whether each of measurement instrument had been demonstrated to be valid, discriminative, and feasible to use in an adult critical care population. This assessment, made by the research team, was based on the presence of published studies demonstrating reliability and construct validity against a gold standard measure in critically ill adults (14). Instrument feasibility was assessed via discussion with those study investigators who had experience with the instrument's use.

In advance of the online consensus meetings, participants who agreed to join the consensus meetings were invited to submit additional measurement instruments for consideration at the meeting. For additional measurement instruments submitted, the research team assessed their validity, reliability, and validity before making a decision for inclusion in consensus meetings discussions.

Consensus Meeting Participants

To establish the COMS, we used a random number generator to select participants previously involved in the COS development and sent email invitations to one of the three online consensus meetings. Three meeting times were provided to optimize participation across time zones and schedules of the invited participants. Details of the COS recruitment processes have been reported previously (12). If invitations were declined, additional invitations were made at random from the COS participant list and to other known experts in the field of delirium measurement in the critically ill via snowballing methods. Invited ICU survivors and family members had either participated in the Del-COrS COS consensus meeting or had prior research experience as a patient advisor through ICU steps, the U.K. intensive care patient support charity, or as a lay member of an ICU trial steering committee.

Consensus Meetings

We held three virtual (Zoom) consensus meetings. Meetings were moderated by three Del-COrS team members (V.P., L.R., L.B.) and were planned for a 2-hour duration. On meeting commencement, participants were reminded of the COMS scope, the criteria on which to base consideration of each instrument for inclusion, and the importance of ICU survivor and family member opinion. The voting process was also explained. All meetings were electronically recorded.

For each outcome, the measurement instruments provided in the briefing documents were presented, then group discussion was invited. During the discussion, the moderating team reminded participants of its scope—that is, discussion and voting related to the use of these measures in the context of clinical research not clinical care. Participants were asked to suggest additional measures. We subsequently used the polling function of Zoom for voting for inclusion in the COMS (yes/no/uncertain). When consensus was not reached on the first vote, repeated voting occurred after further discussion. Following the measurement instrument voting, participants then discussed potential options for measurement time horizons (with response options carried over from prior meetings) and then voted on these. Participants were only allowed to vote for one option for each poll and were encouraged not to abstain; instead, voting "uncertain" if unable to vote "yes" or "no." The a priori threshold for consensus was 70% of votes as "yes for inclusion." When failure to gain consensus occurred, reasons were documented in the meeting digital recording.

Ethical Considerations

The Del-COrS study was approved by the Research Ethics Boards of the University of Toronto (34296 on June 23, 2017); Sunnybrook Health Sciences Centre, Toronto, ON, Canada (448-2017 on February 1, 2018); King's College London (LRS-17/18-6646 on July 10, 2018); and the U.K. Health Research Authority and Health and Care Research Wales (18/

LO/1321 on November 29, 2018). Informed oral consent was obtained from consensus meeting participants. This project is registered with the Core Outcome Measures in Effectiveness Trials (COMET) initiative (http://www.comet- initiative.org/ studies/ details/ 796). Procedures were followed in accordance with the ethical standards of the responsible committee on human experimentation (institutional or regional) and with the Helsinki Declaration of 1975.

RESULTS

Eighteen participants (/31 invited) representing six countries were able to take part in one of the three consensus meetings (**Table 1**).

Outcome 1: Delirium Occurrence

There was unanimous (100%) agreement to recommend the Confusion Assessment Method (CAM)-ICU (15) or Intensive Care Delirium Screening Checklist (ICDSC) (16) as measurement tools to screen for delirium occurrence, as recommended in the Society of Critical Care Medicine's 2013 Pain Agitation/sedation and Delirium (PAD) (7) and 2018 PAD-Immobility and Sleep Disruption guidelines (17). No additional measures were suggested by participants for voting during the three meetings. In relation to measurement timing, we did not achieve consensus on when screening for delirium occurrence should commence, although either on ICU admission or within the first 24 hours was the most frequent response (11/18, 61%). We did not gain consensus on when screening should stop; however, ICU discharge was the most common response (12/18, 67%) (**Table 2**). Participants advocating for assessment of delirium occurrence after ICU discharge expressed concern about delirium reoccurrence. Advocates for discontinuing screening on ICU discharge were concerned about the feasibility and resource implications of delirium assessment outside the ICU. We did achieve consensus that assessment frequency should be once per nursing shift, either 8 or 12 hours, whichever worked in the participating ICU with additional assessments as indicated (e.g., change in mental status). Those participants advocating for a once daily assessment were from North America. These participants were concerned about the feasibility of once per nursing shift assessment when the usual practice is for trained research staff to undertake delirium screening assessments for clinical trials, whereas researchers from other continents frequently rely on clinical documentation of delirium screening conducted by the bedside nurses (**Table 3**). Another feasibility issue raised was that dedicated research staff are not consistently available to collect data on evenings or weekends.

Outcome 2: Delirium Severity

The CAM-ICU-7 (18) was the only measurement tool considered as a validated measure of delirium severity developed specifically for the critically ill. The Intensive Care Psychological Assessment Tool (IPAT) was proposed as a measure of delirium-related emotional distress (19). Only four of 18 participants (22%) voted for inclusion of the CAM-ICU-7 in the COMS with 10 of 18 (56%) indicating uncertainty (Table 2). The CAM-ICU-7 was considered insufficient in covering the range of delirium severity symptoms experienced in the critically ill. Additionally, participants expressed concern that it has been used in a limited number of studies

TABLE 1.
Consensus Meeting Participants

Meeting	Survivor/Family Member	Physician	Nurse	Pharmacist	Total
1	2	2	1	1	6
2	3	1	2	2	8
3	0	3	0	1	4
Total	5	6	3	4	18

Countries represented: United Kingdom (7), United States (4), Canada (3), Spain (1), Denmark (1), and The Netherlands (2).
All clinician participants were also active academics with experience in clinical trials of delirium interventions.

TABLE 2.
Voting on Delirium Occurrence and Time to Delirium Resolution

Vote	*n* (%)
Confusion Assessment Method-ICU or ICDSC as the measurement tool for delirium occurrence? (*n* =18)	
Yes	18 (100)
When should we commence assessing for delirium occurrence? (*n* =18)	
On ICU admission or within 24 hr of ICU admission	11 (61)
On randomization/study entry	7 (39)
When should we discontinue assessing for delirium occurrence? (*n* =17)	
ICU discharge	12 (67)
Day 7 after ICU admission	3 (17)
Day 5 after ICU admission	1 (6)
Day 14 after ICU admission	1 (6)
Day 21 after ICU admission	1 (6)
Other	1 (6)
How frequently should we look for delirium occurrence? (*n* =18)	
Every shift (eight or 12 depending on nurse shift pattern) and as indicated, i.e., change in mental status	13 (72)
Daily	5 (28)
What is the minimum time to consider delirium resolution has occurred? (*n* = 18)	
Minimum 48 hr negative screen	10 (50)
Uncertain	5 (28)
ICU discharge	1 (6)
Sometime after hospital discharge	1 (6)
Other	1 (6)

Not percentages may not sum to 100% due to rounding.
ICDSC = Intensive Care Delirium Screening Checklist.

(Table 3). Nine of the 17 participants (53%) voted for the IPAT as a measure for delirium-related distress. Survivors and family members participants expressed a need for a measurement tool that captured delirium severity in terms of both the emotional distress and impact on their lives it may cause. These participants identified that a clinician's perception of severity might differ from that of the patient, highlighting the subjective nature of delirium symptom presence and severity characterization. All participants agreed there was need for a greater understanding of what a delirium severity measure in the critically ill should and could measure. Suggestions included the impact on the patient, distress of perceptual disturbances, as well as intensity and duration of delirium symptoms. However, during discussion, participants expressed uncertainty that this could be translated into a measurement tool with meaningful numeric values. Participants recommended further research is needed to develop a valid and reliable multidomain measure of delirium severity in the critically ill.

During consensus meeting one, there was opinion that as severity is a challenging construct in the critically ill with no consensus on a measurement tool, voting on the time horizon was not useful. In meetings two and three, there was interest to vote on aspects of measurement timing. Of these 12 participants, 9 (75%) indicated that severity should be measured whenever delirium is detected, with severity measurement discontinued when delirium is no longer detected. Therefore, we achieved consensus that delirium severity should be assessed when delirium is detected, which will be useful to consider following further research to develop a valid and reliable delirium severity measurement instrument for critically ill adults.

Outcome 3: Time to Delirium Resolution

Given that we achieved consensus on the use of the CAM-ICU or ICDSC to screen for delirium in the ICU, and with a negative screen indicating resolution, the discussion and voting focused on the timing of delirium resolution; however, did not reach consensus. Survivors and family members described resolution of delirium as a process rather than a specific point in time relative to measurable time points like hospital discharge. Nine meeting participants (50%) voted that delirium resolution should be defined as no delirium detected for 48 hours, other participants were uncertain or advocating for ongoing screening for resolution to be discontinued at ICU or hospital discharge.

Outcome 4: Mortality

There was unanimous agreement to use "confirmation of death" and mortality rate as previously described in other COMS for critically ill adults. Consensus was not reached on the timepoint for measuring mortality. Only four of 16 participants (25%) who voted on the timepoint agreed to measurement at 60 days from

TABLE 3.
Quotes From Consensus Meeting Recordings

Outcome	Topic	Example Quote
Delirium occurrence	Frequency	"Doing it per shift makes trial very hard from research staff perspective (night cover)—either daily or rely on bedside nurse assessment." Nurse researcher
	Start and finish of data collection	"Need to assess from admission to pick up when it occurs" Survivor
		"Should be every nurse shift routinely, and more when patient seems delirious" Physician researcher
		"Disconnect between idea of measuring outcome before the start of a trial." Physician researcher
		"Measurement of occurrence depends on the intervention" Physician researcher
Delirium severity	Definition	"Severity is very subjective, very distressing vs mildly distressing" Survivor
		"Severity equates with frightening." Survivor
		"Two ways of measuring, one observing the patient while they are having an episode of delirium and second and more valid way interrogating the patient after they are 'awake' about the severity of the incident and how it has affected them." Survivor
	Confusion Assessment Method-ICU-7	"Cautious about recommending a tool with limited validation. If recommended it will be used to the exclusion of other instruments being described." Physician researcher
	Future research	"Every study should do something on severity so we can learn more as we don't know or understand it" Nurse researcher
		"I think we should suggest exploring delirium severity but without a recommended tool." Nurse researcher
		"What is more important a high severity score for one day with no impact on long term outcomes or patient delirious for 7–8 days with long term impact?" Physician researcher
Time to resolution	Timing of data collection	"I couldn't put my finger on this is the end of my delirium, it was certainly quite a while after I got back home." Survivor
		"It doesn't stop, it fades with time. When I got back home things became real to me and the haze clears." Survivor
		"If you finish time to resolution on discharge from ICU/hospital and patient still delirious then you are not measuring time to resolution." Nurse researcher
	Future research	"Look at ventilation where we use 48 hours at success in ventilation. We're not there yet in delirium, we don't have data that teaches us this or understand it. When has an episode resolved is beyond what we understand at this point." Physician researcher
Mortality	Timing of endpoint data collection	"Most hospital administration collects data for 28 or 30 days" Physician researcher
		"60 days arbitrary – at 3 months recovery important question is 'what it life like? For acute outcome 'did you make it to one month alive?'" Nurse researcher
Health-related quality of life	Timing of data collection	"Research challenges of collecting outcomes push the outcome if you can." Physician researcher
		"Personally took me 6 months before brain switched on, I was mentally sluggish" Survivor
		"Shorter one would be easier if powers of concentration are weak" Survivor

(*Contiunued*)

TABLE 3.
Quotes From Consensus Meeting Recordings

Outcome	Topic	Example Quote
Emotional distress	Acute emotional distress	"IPAT would be useful but would need more questions and time to talk to the patient, has to have time to listen to the answers to get accurate feedback." Survivor
	Depression	"HADs does not follow the DSM criteria for these conditions. It does not even include one of the 2 cardinal symptoms of depression, which is depressed mood. I think it is convenient because it combines anxiety and depression. You are supposed to pay for it as it is not publicly available. I would not agree that it is a valid tool to assess clinical symptoms of anxiety and depression because it does not align with DSM criteria for either of these conditions." Physician researcher
	Post-traumatic stress disorder	"IES-R does not diagnostic of PTSD but ok to use as long as you know what you are getting from it and we are looking at emotional distress." Physician researcher
Cognition	Recommendation to use Montreal Cognitive Assessment blind	"MoCA blind is not enough, not adequate and not sensitive however you do not want to miss impairment at the level the MoCA blind picks up." Physician researcher
		"For repeated measurements I'm less concerned with when how many times and when but more interested in doing it more than once, they are not static problems because of the trajectory, some may plateau, rise or decline." Physician researcher

study randomization as defined in previous COMS (8, 9); 8 (50%) voted for measurement at 90 days, and the remaining four were uncertain. Participants that were uncertain expressed that there might be a need for an earlier measurement timepoint as well as routine collection of 30-day mortality for hospital benchmarking. Others thought a longer time horizon more appropriate.

Outcome 5: Health-Related Quality of Life

Proposed measures included the EQ-5D three-level (3L), the EQ-5D five-level (EQ-5D-5L) (20, 21) (as recommended in other critical care COMS), and the 12-Item Short Form Health Survey (22) (suggested by participant). In the discussion preceding voting participants at each meeting requested to vote only on the EQ-5D-5L to avoid confusion among researchers given this was the measure that gained consensus in other critical care COMS. We did not reach consensus to use the EQ-5D-5L to assess HrQoL with 11 of 16 (69%) voting yes and the remaining five participants voting unsure, no participants voted no. In subsequent discussion, participants who voted unsure indicated they were not overly familiar with the measure and therefore did not feel confident to vote yes or no. We also did not reach consensus on when to measure HrQoL with 11 participants (69%) voting for measurement at 6 months and 9 (56%) at 3 months.

Outcome 6: Emotional Distress Including Anxiety, Depression, Acute Stress, and Post-Traumatic Stress Disorder

Measures proposed were those recommended in existing critical care COMS (i.e., the Hospital Anxiety and Depression Scale [HADS] (22), Impact of Event Scale-Revised [IES-R], and Impact of Event Scale-6 [IES-6] [23]). During the consensus meeting, participants also proposed the Patient Health Questionnaire-9 (24), Generalized Anxiety Disorder 7-item (25), Beck Depression score (26), and Post-Traumatic Stress Syndrome (PTSS) 10-Questions Inventory (27).

Consensus was reached to use the HADS to measure symptoms of anxiety and depression (12/17, 71%). Consensus was not achieved for the IES-R, the IES-6 to measure PTSS (11/17, 65%), or the IPAT (9/17, 53%) as a measure for delirium-related distress. We achieved consensus that emotional distress should be measured up to 12 months with repeated measurement at various intervals. There was no consensus, however, on these measurement intervals: 30 days (one vote), 90 days (three votes), and 6 months

(five votes) over 12 months. However, participants also expressed concerns about the feasibility of measurement beyond 3 months for an ICU clinical trial. Having an understanding of baseline emotional distress was also highlighted, although participants questioned the feasibility and validity of evaluating this in an ICU population.

Outcome 7: Cognition Including Memory

As described earlier, we proposed the MoCA-Blind (13) as it can be administered via the telephone and was recommended in the COMS for acute respiratory failure survivors despite failure to reach consensus. No additional measures of cognition were proposed by meeting participants. We gained consensus to include the MoCA-Blind (13/16, 83%) in this COMS with consensus (11/14, 79%) that measurement should continue up to 12 months with repeated measurement within this period. However, there was no consensus on these measurement intervals (3 mo [one vote], 6 and 12 mo [six votes], and unspecified [seven votes]). However, the feasibility of obtaining these data at 12 months was again highlighted. Participants identified further work is needed to validate the MoCA-Blind in ICU survivors.

DISCUSSION

In conducting this COMS, we reached consensus on a measurement instrument for four of the six outcomes that require a measurement instrument (mortality does not require an instrument); CAM-ICU or ICDSC to assess for delirium occurrence and for resolution; HADS for anxiety and depression; and MoCA-blind for cognition. We were unable to reach consensus to use the EQ-5D-5L for HrQoL measured at 6 months or for the IES-R or IES-6 for PTSS. There was also no consensus on measurement instruments for delirium severity or delirium-related emotional distress as well as when to commence and when to discontinue assessing for delirium occurrence, although we did gain consensus that it should be assessed at least once a nursing shift. Despite not having consensus on a measure, we did gain consensus that delirium severity should be measured every time delirium occurrence is assessed. We were unable to gain consensus on when to discontinue assessing delirium presence to determine time to resolution and when to determine mortality. We did gain consensus that emotional distress and cognition should be measured up to 12 months from discharge.

This COMS should now be included in the design of all future studies of interventions to prevent or treat delirium in critically ill adults. This will help to reduce the substantial heterogeneity and multiplicity of outcome selection and measurement we previously identified in published studies (12). A recent systematic review conducted by members of COMET identified poor uptake of COS across most research areas (28). To promote implementation of the COMS in future research, we will work with the American and Australian Delirium Societies and European Delirium Association who have endorsed this work to identify dissemination and uptake strategies.

Failure to gain consensus along with the associated discussion regarding a measurement instrument for delirium severity established that there is a need for further research to define delirium severity content domains and subdomains for the critically ill prior to developing and validating a multidomain measure specifically for this patient population. Existing high quality multidomain measures of delirium severity were not developed specifically for a critically ill patient population (29) despite use in previous ICU studies (12). Despite being developed for the critically ill, coverage of the domains of delirium severity assessment by the CAM-ICU-7 is limited (29). Work conducted by the Better ASsessment of ILlness study group has recently established seven generic domains and multiple subdomains of delirium severity (30). Domains include cognitive; level of consciousness; inattention; psychiatric-behavioral; emotional dysregulation; psychomotor features; and functional. Additional condition-specific (Alzheimer's) domains and subdomains have also been established (31). Given their lived experience, ICU survivors and family members should be well represented in future work to define delirium domains and subdomains relevant to the critically ill that will inform measure development.

Although we reached consensus to use the HADS to measure emotional distress associated with anxiety and depression, throughout our COMS development work ICU survivors and family members identified the need to measure delirium-related emotional distress. This distress is reflected in the subdomains of the emotional dysregulation domain proposed for generic delirium severity (i.e., anxiety/fear/sense of unease,

depression/apathy/withdrawal, anger/hostility/irritability) (30). Given that we were unable to gain consensus on the IPAT, which is a screening tool designed to detect acute distress in intensive care (19), we anticipate future work to develop a multidomain delirium severity measure should encompass measurement of delirium-related emotional distress.

We did not gain consensus on the use of the EQ-5D-5L at 6 months for HrQoL. However, given that this measure and measurement timepoint are recommended in other ICU COS (8, 9) and that no participant voted to not use it, we would advise researchers also consider this for delirium prevention or treatment trials. Consensus for repeated measurement up to 12 months for both emotional distress and cognition reflects the protracted nature of recovery from critical illness and post-intensive care syndrome (32). However, we note participants also expressed concern as to the feasibility of measurement up to 12 months with reasons including concerns regarding research funding bodies providing the resources to collect data beyond 6 months and engagement of participants to provide such data. Participants also questioned the relevance of anticipating an effect of an intervention delivered in ICU to prevent or treat delirium outcomes measured after 6 months not detected earlier.

Strengths of our study include the application of COMS development methods as recommended by COMET (5), ICU survivor and family engagement throughout the Del-COrS ICU COMS, and the international expertise of participants with representation from medicine, nursing, and pharmacy. Although virtual meetings enabled international engagement, they also introduced limitations. One family member commented that the virtual meeting limited informal chat and exchange of information that helps to promote consensus. Our final meeting did not have ICU survivor or family member representation, which may have influenced voting; however, we did share ICU survivor/family member comments from meetings one and two. Although the same briefing documents were provided to all participants, knowledge and experience of measures varied leading to some participants voting uncertain at times.

In conclusion, we have established consensus on measures for four of six core outcomes requiring an instrument: 1) CAM-ICU or ICDSC for delirium occurrence measured every nursing shift and used for detecting time to delirium resolution, 2) HADS for measuring anxiety or depression, and 3) MoCA-Blind for assessment of cognitive outcomes both measured repeatedly up to 12 months. Although we did not reach consensus on the EQ-5D-5L at 6 months for HrQoL, we strongly recommend researchers consider this measure. We failed to gain consensus on measures for delirium severity and for delirium-related distress, with further research needed in the development and validation of measurement instruments for these outcomes.

1 *Florence Nightingale Faculty of Nursing, Midwifery and Palliative Care, Division of Applied Technologies for Clinical Care, King's College London, London, United Kingdom.*

2 *Wellcome-Wolfson Institute for Experimental Medicine, School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, Belfast, Northern Ireland, United Kingdom.*

3 *Division of Pulmonary and Critical Care Medicine, School of Medicine and School of Nursing, Johns Hopkins University, Baltimore, MD.*

4 *School of Pharmacy, Northeastern University, Boston, MA.*

5 *Division of Pulmonary and Critical Care Medicine, Brigham and Women's Hospital, Boston, MA.*

6 *School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, Belfast, Northern Ireland.*

7 *Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, Canada.*

8 *Mount Sinai Hospital, Sinai Health System, Toronto, Canada.*

The Del-COrS group are as follows: Meera Agar, Jason Boland, Shirley Bush, Noll L. Campbell, Imogen Featherstone, Leslie Fornear, Annmarie Hosie, Miriam Johnson, Richard N. Jones, Peter Lawlor, Jacques Lee, John Marshall, Najma Siddiqi, and Valerie Page.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website (http://journals.lww.com/ccejournal).

Dr. Rose and Dr. Page conceived the ICU delirium core outcome measurement set. Dr. Rose, Dr. Burry, and Dr. Page led the consensus meetings and analyzed the data. Dr. Rose wrote the first article draft, and all authors critically revised the article for important contents and approved the final version.

This study was funded by the Canadian Institutes of Health Research (Project Grant–FRN 15309).

The authors have disclosed that they do not have any potential conflicts of interest.

For information regarding this article, E-mail: louise.rose@kcl.ac.uk

This Core Outcome Set is registered on the Core Outcome Measures in Effectiveness Trials website (http://www. comet- initiative. org/ studies/ details/ 796).

REFERENCES

1. Williamson P, Altman D, Blazeby J, et al: Developing core outcome sets for clinical trials: Issues to consider. *Trials* 2012; 13:132
2. Salluh JI, Wang H, Schneider EB, et al: Outcome of delirium in critically ill patients: Systematic review and meta-analysis. *BMJ* 2015; 350:h2538
3. Pisani MA, Ferrante L: Among patients admitted to ICU, delirium is most common in those with neurological diagnoses, and is associated with adverse health outcomes. *Evid Based Nurs* 2013; 16:81–82
4. Pandharipande P, Ely E, Arora R, et al: The intensive care delirium research agenda: A multinational, interprofessional perspective. *Intensive Care Med* 2017; 43:1329–1339
5. Williamson P, Altman D, Bagley H, et al: The COMET handbook: Version 1.0. *Trials* 2017; 18:280
6. Rose L, Burry L, Agar M, et al: A core outcome set for research evaluating interventions to prevent and/or treat delirium in critically ill adults: An international consensus study (Del-COrS). *Crit Care Med* 2021; 49:1535–1546
7. Barr J, Fraser G, Puntillo K, et al: American College of Critical Care Medicine: Clinical practice guidelines for the management of pain, agitation, and delirium in adult patients in the intensive care unit. *Crit Care Med* 2013; 41:278–280
8. Blackwood B, Ringrow S, Clarke M, et al: A core outcome set for critical care ventilation trials. *Crit Care Med* 2019; 47:1324–1331
9. Needham D, Sepulveda K, Dinglas V, et al: Core outcome measures for clinical research in acute respiratory failure survivors. An international modified Delphi consensus study. *Am J Respir Crit Care Med* 2017; 196:1122–1130
10. Tong A, Baumgart A, Evangelidis N, et al: Core outcome measures for trials in people with coronavirus disease 2019: Respiratory failure, multiorgan failure, shortness of breath, and recovery. *Crit Care Med* 2021; 49:503–516
11. Munblit D, Nicholson T, Akrami A, et al; PC-COS project steering committee: A core outcome set for post-COVID-19 condition in adults for use in clinical practice and research: An international Delphi consensus study. *Lancet Respir Med* 2022; 10:715–724
12. Rose L, Agar M, Burry L, et al; Development of Core Outcome Sets for Effectiveness Trials of Interventions to Prevent and/or Treat Delirium (Del-COrS) Group: Reporting of outcomes and outcome measures in studies of interventions to prevent and/or treat delirium in the critically ill (Del-COrS): Systematic review. *Crit Care Med* 2020; 48:e316–e324
13. Nasreddine ZS: Montreal cognitive assessment (MoCA)-blind, 2019. Available at: www.mocatest.org. Accessed February 14, 2023
14. Gélinas C, Bérubé M, Chevrier A, et al: Delirium assessment tools for use in critically ill adults: A psychometric analysis and systematic review. *Crit Care Nurs* 2018; 38:38–49
15. Ely E, Inouye S, Bernard G, et al: Delirium in mechanically ventilated patients: Validity and reliability of the confusion assessment method for the intensive care unit (CAM-ICU). *JAMA* 2001; 286:2703–2710
16. Bergeron N, Dubois M, Dumont M, et al: Intensive care delirium screening checklist: Evaluation of a new screening tool. *Intensive Care Med* 2001; 27:859–864
17. Devlin J, Skrobik Y, Gélinas C, et al: Executive summary: Clinical practice guidelines for the prevention and management of pain, agitation/sedation, delirium, immobility, and sleep disruption in adult patients in the ICU. *Crit Care Med* 2018; 46:1532–1548
18. Khan B, Perkins A, Gao S, et al: The confusion assessment method for the ICU-7 delirium severity scale: A novel delirium severity instrument for use in the ICU. *Crit Care Med* 2017; 45:851–857
19. Wade D, Hankins M, Smyth D, et al: Detecting acute distress and risk of future psychological morbidity in critically ill patients: Validation of the intensive care psychological assessment tool. *Crit Care* 2014; 18:519
20. Herdman M, Gudex C, Lloyd A, et al: Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011; 20:1727–1736
21. Euroqol G: EuroQol: A new facility for the measurement of health-related quality of life. *Health Policy* 1990; 16:199–208
22. Ware J, Kosinski M, Keller S: A 12-item short-form health survey: Construction of scales and preliminary tests of reliability and validity. *Med Care* 1996; 34:220–233
23. Weiss D, Marmar C: The impact of event scale – revised. In: *Assessing Psychological Trauma and PTSD*. Wilson J, Keane T (Eds). New York, NY, Guilford Press, 1997
24. Kroenke K, Spitzer R, Williams J: The PHQ-9: Validity of a brief depression severity measure. *J Gen Intern Med* 2001; 16:606–613
25. Spitzer R, Kroenke K, Williams J, et al: A brief measure for assessing generalized anxiety disorder: The GAD-7. *Arch Intern Med* 2006; 166:1092–1097
26. Beck A, Ward C, Mendelson M, et al: An inventory for measuring depression. *Arch Gen Psychiatry* 1961; 4:561–571
27. Stoll C, Kapfhammer H, Rothenhäusler H, et al: Sensitivity and specificity of a screening test to document traumatic experiences and to diagnose post-traumatic stress disorder in ARDS patients after intensive care treatment. *Intensive Care Med* 1999; 25:697–704
28. Williamson P, Barrington H, Blazeby J, et al: Review finds core outcome set uptake in new studies and systematic reviews needs improvement. *J Clin Epidemiol* 2022; 159:154
29. Jones R, Cizginer S, Pavlech L, et al: Assessment of instruments for measurement of delirium severity: A systematic review. *JAMA Intern Med* 2019; 179:231–239
30. Schulman-Green D, Schmitt E, Fong T, et al: Use of an expert panel to identify domains and indicators of delirium severity. *Qual Life Res* 2019; 28:2565–2578
31. Schulman-Green D, Hshieh T, Adamis D, et al: Domains of delirium severity in Alzheimer's disease and related dementias. *J Am Geriatr Soc* 2022; 70:1495–1503
32. Rousseau A, Prescott H, Brett S, et al: Long-term outcomes after critical illness: Recent insights. *Crit Care* 2021; 25:108