Nursing Papers and Journal Articles   School of Nursing

2020

Toward best practice methods for delirium biomarker studies: An international modified Delphi study

Ingrid Amgarth-Duff

Annmarie Hosie
The University of Notre Dame Australia, annmarie.hosie1@nd.edu.au

Gideon Caplan

Meera Agar

This is the peer reviewed version of the following article:

Amgarth-Duff, I., Hosie, A., Caplan, G., and Agar, M. (2020). Toward best practice methods for delirium biomarker studies: An international modified Delphi study. *International Journal of Geriatric Psychiatry, Early View Online First,* doi: 10.1002/gps.5292

This article has been published in final form at: -

https://doi.org/10.1002/gps.5292

Toward Best Practice Methods for Delirium Biomarker Studies: An International Modified Delphi Study

Authors

Ingrid Amgarth-Duff[1] (BBiomedSc (Hons), PhD candidate), Annmarie Hosie[1,2] (PhD),

Gideon Caplan (MD)[3,4], Meera Agar (PhD)[1,5,6]

Affiliations

1. University of Technology Sydney, IMPACCT (Improving Palliative, Aged and Chronic

Care through Clinical Research and Translation), Sydney, NSW, Australia

2. University of Notre Dame Australia, School of Nursing Sydney

3. Prince of Wales Clinical School, University of New South Wales, Sydney, NSW, Australia

4. Department of Geriatric Medicine, Prince of Wales Hospital, Sydney, NSW, Australia

5. South West Sydney Clinical School, University of New South Wales, Liverpool, New

South Wales, Australia

6. Ingham Institute of Applied Medical Research, Liverpool, New South Wales, Australia

Corresponding author: Ingrid Amgarth Duff; Ingrid.Amgarth-Duff@uts.edu.au;

Data availability statement: All data generated or analysed in this study are included within
this manuscript.

Main text word count: 2738

Abstract

Background

Delirium is a serious and distressing neurocognitive condition common in people with advanced illness. The understanding of delirium pathophysiology is limited and largely hypothetical. To accelerate empirical understanding of delirium pathophysiology, robust scientific methods for conducting and reporting delirium biomarker studies are urgently needed. The aim of this study was to develop international consensus on the core elements of high quality delirium biomarker studies.

Methods

A three-round modified Delphi survey was conducted from February to August, 2019. Participants were international researchers experienced in conducting delirium studies from a range of settings (hospital, university, research centres). Round one commenced with open-ended questions developed from results from a prior systematic review and the REMARK checklist. Responses were qualitatively analysed and closed statements were developed. Participants then ranked the importance of these statements using a 5-point likert scale in rounds 2 and 3. *A priori* consensus was defined as ≥70% participant agreement. Descriptive statistics for each item were computed including the mean Likert scores, standard deviation (SD), and median participant scores.

Results

Twenty-eight participants completed survey round one, 16 completed round two, and 19 completed the final round. Consensus was achieved for a total of 60 items.

Conclusion

The Delphi survey identified items that expert researchers agreed were important in the conduct of delirium biomarker studies. These reporting items provide a strong platform for improved methodological quality and opportunities to synthesise future delirium biomarker studies.

Key words: Guidelines, Methodology, Consensus**,** Pathophysiology

Key points:

- Despite the prevalence and impact of delirium, knowledge of its pathophysiology is largely hypothetical. Better understanding of the pathophysiology of delirium is crucial to develop more effective ways to prevent and treat delirium.

- To understand the pathophysiology of delirium, more robust scientific methodologies for delirium biomarker research are needed.

- There are currently no guidelines for conducting and reporting delirium biomarker studies, which impacts on the individual and overall quality of this body of research. Reporting guidelines would improve the rigor of its methodology and reporting, and increase the potential for future studies to be synthesised through meta-analyses.

Introduction

Delirium is a serious, acute and complex neurocognitive condition that is often precipitated by an acute medical event such as infection, or surgery. Delirium is characterized by an acute change in attention, awareness and cognition and variously affects memory, language, visuospatial ability, orientation and perception[1]. Delirium is associated with multiple adverse clinical outcomes including high levels of patient and caregiver distress, increased morbidity, mortality and length of hospital stay and significant costs to the healthcare system[2-6]. A systematic review found delirium prevalence in medical in-patients at admission to hospital to range between 10 and 31%, with incidence of new delirium during admission ranging from 3 to 29%. Occurrence rates for delirium per admission ranged between 11 and 42%[7]. Despite the high prevalence and impact of delirium, knowledge of its pathophysiology is largely hypothetical[8]. Hence, biomarker studies are crucial in this field to accelerate our understanding of delirium biology leading to potential therapies. A biomarker is a biological molecule found in blood, other body fluids, or tissues that is a sign of a normal or abnormal process, or of a condition or disease[9].

Reporting guidelines currently exist that are relevant to biomarker studies. These are the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) for reporting observational studies[10], reporting guidelines for body fluid markers in neurologic disorders[11], the STARD (STAndards for the Reporting of Diagnostic accuracy)[12] and the REMARK (REporting recommendations for tumour MARKer prognostic studies)[13]. However, no reporting guidelines currently exist for delirium biomarker studies, and it is not

known how well these existing guidelines may be modified to inform optimal delirium biomarker research.

In the absence of reporting guidelines in delirium biomarker research, we applied the REMARK checklist,[13] a reporting guideline for tumour marker prognostic studies, to assess the quality of studies included in a recent systematic review of the overlap of delirium and advanced cancer biomarkers (PROSPERO CRD42017068662). The review found that most of the 151 included articles were of low quality. Unfortunately, despite the volume of studies, their overall low quality limits the trustworthiness and impact of outcomes, comparability of results and ability to synthesise findings to inform empirical understanding of delirium pathophysiology. The absence of reporting guidelines for delirium biomarker studies has likely contributed to this identified problem.

Therefore, this study aimed to obtain international consensus from leaders in delirium research, on the core elements for high quality delirium biomarker studies, to improve our understanding of delirium pathophysiology.

Methods

Study design

A three-round survey was employed in accordance with the Delphi method[14].

Participants

Those considered eligible were delirium researchers who had investigated delirium in humans, including but not restricted to biomarkers. Researchers with basic science and animal study backgrounds were also eligible if their research focus was on delirium. Expert panel members were required to have delirium research experience in the last ten years (with no minimum number of years pre-specified), and computer and internet access with an email address to access the online survey. Those who met these eligibility criteria were deemed to have adequate knowledge, expertise and opportunity to make a meaningful contribution to the topic area.

Recruitment

A combination of purposive sampling and snowballing was used to recruit the expert panel[15,16]. Purposive recruitment approaches included: 1) email invitation via membership lists of Delirium Societies' (Australasian Delirium Association, American Delirium Society and the European Delirium Association); 2) email invitations through colleagues and professional networks; and 3) researchers identified from journal articles as having experience in delirium biomarker studies. An indirect approach included a Twitter advertisement on the 2019 'World Delirium Awareness Day'[17]. Snowball sampling was achieved by asking eligible participants and presidents of delirium societies to invite any other eligible researchers who may be interested in taking part in the study, by forwarding the invitation via email.

Data collection

Each potential participant was sent an email invitation with a link to the online REDCap survey in three parts: A participant information sheet outlining the study procedures and their involvement in the study, a demographics section, and the survey questions. Non-completion of a round did not prohibit participants from participating in the subsequent rounds. Demographic details were collected at the beginning of each round, only once per participant. A reminder email was sent around14 days following dissemination of each survey round.

Round 1

Round 1 aimed to generate a broad range of opinions. This round was informed by results from the quality assessment of a prior systematic review, and predominantly used an open-ended qualitative method, as in the traditional approach to the Classic Delphi [16]. The initial draft survey of round 1 was piloted by three researchers with sufficient clinical understanding of delirium and knowledge of biomarker research. These researchers were not involved in the Delphi development and were not eligible to be study participants.

In round 1, participants were provided with both open-ended and closed questions about biomarker research in delirium based on each key domain of the REMARK checklist[13]. Participants were also invited to provide comments after each question. The answers from round 1 informed development of a list of statements for round 2 of the Delphi.

Round 2

In round 2, 56 statements were reduced by a rating process whereby participants rated each statement on a 5-point Likert scale from 1 (not important at all) to 5 (very important). Participants were also invited to provide comments and suggest any alternate wording for

each statement. Reasons for excluding comments or items suggested by participants were recorded.

Round 3

This final round aimed to refine the final list of statements pertaining to recommendations for reporting of delirium biomarker studies. In round 3, participants were sent the survey along with: 1) a summary of round 2 statements that reached consensus; 2) a summary of statements that did not reach consensus (which were repeated in this round); and 3) newly suggested statements from participants' comments in round 2. Group ratings were displayed next to each statement, allowing participants to revise the collective response in a blinded way. Participants were asked to provide a new rating on the 5-point Likert scale. Only statements that did not achieve consensus from round 2 were carried into round 3. Round 2 statements that already achieved a consensus were excluded from round 3, but were still presented in the summary for participants to review.

Data analysis

Round 1

Demographic data from each round was collated and inputted into the IBM Statistical Package for Social Science (SPSS), Version 25. Round 1 open-ended responses were compiled from Excel spreadsheets into Microsoft Word and thematically analysed by the lead author (IAD), with two other reviewers (MA and AM) providing guidance and oversight of the themes and codes. Reviewers discussed any uncertainties about the coding or themes until an agreement was met. Reasons recorded for excluding or amending comments or items prior to round 2 were that the item/comment(s) were:

  i. too vague

  ii. a misunderstanding of the question

  iii. not relevant to the topic or study

  iv. repetitious in meaning or intent

  v. already encompassed within another item and/or or better combined with another item

Rounds 2 and 3

A target 70% agreement for the score of 4 or more on the 5-point Likert scale for each statement was chosen a priori. REDCap data were exported to SPSS for statistical analysis. Descriptive data for each item were obtained, including the mean Likert scores, standard deviation (SD) and the median. Round 2 items with the greatest participant agreement in the very low and low importance categories (Likert score 1 and 2) were deemed unlikely to be included in the list of recommendations; items with the participant agreement in the moderate importance category (Likert score 3) were considered for inclusion in the recommendations and items with the greatest participant agreement in the high to very high importance category (Likert scores $\geq$4), were included in the recommendations. Data analysts were blinded to participants' identities.

Ethical considerations

Ethical approval was obtained from the University of Technology Sydney Human Research Ethics Committee (approval no. ETH18-2673).

Results

Participants

Surveys were delivered over three rounds from February to August 2019 via email. Twenty-nine participants completed round 1, however, one participant's data was removed as it was clear to the authors that the questions had not been understood, and therefore the responses were not able to be coded. Nineteen participants completed round 2, and 20 completed round 3, with a total of 32 participants completing at least one round and 10 completing all three rounds. Participants were from 12 countries (Argentina, Australia, Belgium, Germany, Italy, Norway, Portugal, Sweden, Switzerland, The Netherlands, United Kingdom (UK) and United States (US)). Overall, the expert panel were predominantly clinician researchers (n=21; 64%), with 47% of participants having over 10 years' experience in delirium research and 47% having conducted more than 10 delirium studies. Twenty five (78%) of participants had conducted between 0 and 5 biomarker studies, 13% between 5 and 10, and 3 participants (9%) had conducted over 10 biomarker studies. Twenty two (69%) had conducted a delirium biomarker study, and nine (28%) of participants had a research higher degree in delirium and two (6%) in biomarkers (table 1).

Insert table 1 here

Consensus

The 18 open-ended questions and 5 closed questions of round 1 were grouped and reduced to 56 statements for round 2, with statements adjusted or removed if unclear, repetitive or already encompassed in another statement, not relevant to topic, or better combined with another item. An outline of the process of including items in the final delirium biomarker recommendations is shown in figure 1. Following round 2, 51 statements reached consensus for inclusion, and 5 statements did not. Twelve newly-suggested statements arising from round 2 were carried into round 3, along with the 5 statements that did not reach a consensus (n=17 items in total). Following round 3, 60 statements reached a consensus, and 8 did not.

Insert figure 1 here.

Figure 1. Flow chart illustrating the three-stage Delphi process, informed by a prior systematic review

The 60 statements that achieved a priori level of consensus for inclusion in the delirium biomarker study reporting guidelines (i.e $\geq$ 70% agreement with scores 4 or 5) are shown in table 2. Table 3 lists the 8 items that did not achieve consensus after 3 rounds of the Delphi. No item received a score of $\leq$ 2 and hence were not excluded based on this criteria.

Insert Table 2 here

Insert table 3 here

The final list of recommendations is presented in table 4.

__Insert table 4 here__

Discussion

This study presents the first set of delirium-specific recommendations to aid in the conduct and reporting of future delirium biomarker research. Consensus was achieved in 60 items, with a total of 8 items that did not reach a consensus. Based on open-ended findings from round 1 and 2, consensus was not achieved on the more complex methodological aspects of delirium biomarker research, for example, accounting for underlying diseases in patients with delirium.

Despite a large number of emerging delirium biomarker studies, the pathophysiology of delirium is still poorly understood. A concerted effort is required to standardise the methodology used in delirium biomarker studies, in order to progress this fundamental field of research. Inadequate and/or unclear reporting of methodological processes can lead to discrepancies in results, which may be misleading and potentially detrimental [18]. Reporting guidelines are necessary to promote studies that are standardised and reported in a transparent manner to facilitate reliable and consistent interpretation, application and synthesis of study results. A systematic review examining the extent to which journals encourage reporting guidelines found that nearly half of the online instructions to authors mentioned reporting guidelines (19/41 (46%))[19]. Other studies have found that reporting guidelines such as the CONSORT (Consolidated Standards of Reporting Trials) Statement[20] has led to improvements in the reporting rigor, particularly in the method of sequence generation and the allocation concealment, compared to studies that did not adopt the CONSORT [21].

Current guidelines that focus on different aspects of biomarkers include the REMARK, STARD and CONSORT statements, which are used when the focus is on prognostic biomarkers, diagnostic testing, or when conducting randomised controlled trials. However, none of these guidelines are specific to delirium. We therefore utilised the REMARK checklist as a framework to guide in the development of these preliminary recommendations for guidelines. The final items illustrate areas where specific guidance was deemed useful by international delirium experts, to specifically address methodological issues in delirium. Three domains overlap with the REMARK checklist (assay procedures, sample size calculation, and univariate and multivariate results) and the remainder are unique to delirium biomarker studies.

Limitations and strengths

Several limitations of this study are worth noting. Firstly, some participants in round 1 did not understand the questions which relied on some background knowledge in the biomarker field. This resulted in 66 comments (66/224; 29.4%) that were excluded from round 1. Secondly, there was noteworthy attrition between rounds, with only 10 participants completing all three rounds. Thirdly, since delirium is a condition which often occurs in the context of other conditions with similar pathophysiological processes, such as cancer, complex questions with multiple competing issues that need to be considered in methodological design are not suited to be reduced down to simple statements within a Delphi method. This requires a more in-depth qualitative approach to identify the nuanced methodological considerations needed. Hence the guidelines presented in this study may not

be universal and researchers will still need to consider whether there are additional special

considerations to be considered when applying them to specific scenarios and settings.

Lastly, there is no universally agreed definition of 'consensus' for a Delphi. Some argue that

51% agreement on an item is acceptable[22], while others maintain anywhere from 75%[23] to

100% agreement amongst respondents[24]. It should also be noted that although the Delphi

concludes when a consensus has been achieved, the end results aren't necessarily the most

reliable or appropriate end-product[25] but rather, a majority opinion[26].

Key strengths include: the systematic approach to generate the final items, drawing on both

the existing literature from a prior systematic review and expert opinion. Another key

strength of this study was the breadth of expertise within the international expert panel,

though we acknowledge that we may have not encompassed all possible perspectives. Lastly,

although there is no universal agreement of the ideal sample size for Delphi studies, most

Delphi's have included between 15 and 20 participants, and the expertise of the panel is

considered more important that the size of the sample itself[14,27,28]. Considering the small

cohort of expert delirium researchers worldwide, we believe 32 participants was a sufficient

sample[16].

Implications for future research and practice

This Delphi study proposes the first set of recommendations to inform development of

reporting guidelines for delirium biomarker studies, which can be refined after experience of

their utility in practice. The systematic review undertaken by the same authors demonstrated

a number of poor quality studies that were likely affected by a lack of guidelines for delirium

biomarker research. Developing reporting guidelines was therefore an essential step to improving methodological and reporting rigor, which will increase the potential for future studies to be synthesised through meta-analyses. This Delphi study proposes a preliminary list of 60 items to be considered in these reporting guidelines. To supplement these recommendations, the authors have conducted interviews with experts in the field discussing the key methodological issues that were more complex for which a Delphi approach was not suited. Namely, how to account for other co-existing conditions (e.g. cancer or sepsis) that plausibly impact on the pathophysiological and/or biological findings. Likewise, the practicalities of obtaining biomarkers from people with delirium for research was another issue that arose from this study which was explored in depth in a follow-up interview study. Ongoing international collaboration will be needed to achieve a tighter consensus.

Conclusion

This study presents the first step towards development of reporting guidelines for delirium biomarker studies through a rigorously conducted Delphi survey of international experts in delirium research. Results will support the development of greater methodological rigor in future delirium biomarker research, which will ultimately contribute to better understanding of the pathophysiology of delirium.

References

1.   American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5).* Arlington, VA: American Psychiatric Publisher; 2013.
2.   Brajtman S. The impact on the family of terminal restlessness and its management. *Palliative Medicine.* 2003;17(5):454-460.
3.   Bruera E, Bush SH, Willey J, et al. Impact of delirium and recall on the level of distress in patients with advanced cancer and their family caregivers. *Cancer.* 2009;115(9):2004-2012.
4.   Salluh JI, Wang H, Schneider EB, et al. Outcome of delirium in critically ill patients: systematic review and meta-analysis. *British Medical Journal* 2015;350:h2538.
5.   Witlox J, Eurelings LS, de Jonghe JF, Kalisvaart KJ, Eikelenboom P, Van Gool WA. Delirium in elderly patients and the risk of postdischarge mortality, institutionalization, and dementia: a meta-analysis. *JAMA: Journal of the American Medical Association.* 2010;304(4):443-451.
6.   Pezzullo L, Streatfeild J, Hickson J, Teodorczuk A, Agar MR, Caplan GA. Economic impact of delirium in Australia: a cost of illness study. *BMJ open.* 2019;9(9):e027514.
7.   Siddiqi N, House A, Holmes J. Occurrence and outcome of delirium in medical in-patients: a systematic literature review. *Age and Ageing.* 2006;35(4):350-364.
8.   Maldonado JR. Delirium pathophysiology: An updated hypothesis of the etiology of acute brain failure. *International Journal of Geriatric Psychiatry.* 2017;33(11):1428-1457.
9.   National Cancer Institute. NCI Dictionary of Cancer Terms. https://www.cancer.gov/publications/dictionaries/cancer-terms/?CdrID=45618. Accessed 28th November, 2017.
10.   Von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. *International Journal of Surgery.* 2014;12(12):1495-1499.
11.   Gnanapavan S, Hegen H, Khalil M, et al. Guidelines for uniform reporting of body fluid biomarker studies in neurologic disorders. *Neurology.* 2014;83(13):1210-1216.
12.   Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *Radiology.* 2015;277(3):826-832.
13.   Altman DG, McShane LM, Sauerbrei W, Taube SE. Reporting recommendations for tumor marker prognostic studies (REMARK): explanation and elaboration. *BMC Medicine.* 2012;10(1):51.
14.   Hsu C-C, Sandford BA. The Delphi technique: making sense of consensus. *Practical Assessment, Research & Evaluation.* 2007;12(10):1-8.
15.   Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. *J Adv Nurs.* 2000;32(4):1008-1015.
16.   Keeney S, McKenna H, Hasson F. *The Delphi technique in nursing and health research.* John Wiley & Sons; 2011.

17.   iDelirium. The International Federation of Delirium Societies. 2019; http://www.idelirium.org/. Accessed October 2nd, 2019.

18.   Simera I, Moher D, Hirst A, Hoey J, Schulz KF, Altman DG. Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network. *BMC Medicine.* 2010;8(1):24.

19.   Hirst A, Altman DG. Are peer reviewers encouraged to use reporting guidelines? A survey of 116 health research journals. *PloS one.* 2012;7(4):e35621.

20.   Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMC medicine.* 2010;8(1):18.

21.   Plint AC, Moher D, Morrison A, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Medical journal of Australia.* 2006;185(5):263-267.

22.   Loughlin KG, Moore LF. Using Delphi to achieve congruent objectives and activities in a pediatrics department. *Journal of Medical Education.* 1979;54(2):101-106.

23.   Keeney S, Hasson F, McKenna H. Consulting the oracle: ten lessons from using the Delphi technique in nursing research. *J Adv Nurs.* 2006;53(2):205-212.

24.   Williams PL, Webb C. The Delphi technique: a methodological discussion. *J Adv Nurs.* 1994;19(1):180-186.

25.   Keeney S, Hasson F, McKenna HP. A critical review of the Delphi technique as a research methodology for nursing. *Int J Nurs Stud.* 2001;38(2):195-200.

26.   Rauch W. The decision delphi. *Technological Forecasting and Social Change.* 1979;15(3):159-169.

27.   Murphy M, Black N, Lamping D, et al. Consensus development methods, and their use in clinical guideline development. *Health Technology Assessment.* 1998;2(3):i-88.

28.   Okoli C, Pawlowski SD. The Delphi method as a research tool: an example, design considerations and applications. *Information and Management.* 2004;42(1):15-29.

Table 1. Demographic characteristics of Delphi participants (n=32)

	n (%)
Country of residence	
USA	14 (44)
Europe	11 (34)
United Kingdom	4 (13)
Australia	2 (6)
Latin America	1 (3)
Years in delirium research	
10+	15 (47)
5-10	10 (31)
0-5	7 (22)
Current role	
Clinician/researcher	21 (64)
Researcher	6 (19)
Clinician	5 (15)
Place of work	
Hospital	26
University	22
Research centre	8
Other	1
Main delirium research area	
Clinical trials	22
Epidemiology	14
Health services	9
Implementation/knowledge translation/education	9
Qualitative research	6
Other	2
Number of delirium studies conducted	
10+	15 (47)
5-10	9 (28)

0-5	8 (25)
Number of biomarker studies conducted	
10+	3 (9)
5-10	4 (13)
0-5	25 (78)
Conducted a delirium biomarker study	
Yes	22 (69)
No	10 (31)
Research higher degree (Masters or Doctorate)	
In delirium	9 (28)
In biomarkers	2 (6)
Both	6 (19)
No	15 (47)

Table 2. Summary of ratings for items that reached a ≥70% consensus after three Delphi rounds*

Statement	Very important (5)	Moderately important (4)	Not important or unimportant (3)	Slightly important (2)	Not important at all (1)	Mean rating/Median rating	SD	Total % consensus achieved (category)
In delirium biomarker studies, the study objective statement should at a minimum, include the following key elements:								
The biomarker under study (including source)	14 (87.5)	2 (12.5)	0 (0.0)	0 (0.0)	0 (0.0)	4.8/5	.34	87.5% (5)
The time of collection in relation to delirium onset	11 (68.8)	3 (18.8)	2 (12.5)	0 (0.0)	0 (0.0)	4.5/5	.72	87.6% (5,4)
The clinical endpoint(s) including their definition	13 (81.3)	2 (12.5)	1 (6.3)	0 (0.0)	0 (0.0)	4.6/5	.79	81.3% (5)
The clinical covariates	9 (45.0)	8 (40.0)	3 (15.0)	0 (0.0)	0 (0.0)	4.3/4	.73	85% (5,4)
The methods of biomarker collection [1]	9 (45.0)	6 (30.0)	3 (15.0)	1 (5.0)	0 (0.0)	4.2/4	.91	75% (5,4)
Clarify which delirium pathophysiological theory the study will address	6 (30.0)	10 (50.0)	2 (10.0)	1 (5.0)	1 (5.0)	3.9/4	1.05	80% (5,4)
The biomarker in a delirium study should be:								
Chosen a priori	9 (56.3)	7 (43.8)	0 (0.0)	0 (0.0)	0 (0.0)	4.5/5	.51	100% (5,4)
Supported by a biologically plausible rationale	12 (75.0)	3 (18.8)	1 (6.3)	0 (0.0)	0 (0.0)	4.6/5	.60	75% (5)
Supported by a clear hypothesis	10 (62.5)	3 (18.8)	3 (18.8)	0 (0.0)	0 (0.0)	4.4/5	.81	81.3% (5,4)
Putting practical considerations aside, the type of biological specimen chosen should:								
Be based on the capacity to measure the proposed biological process being evaluated	7 (43.8)	9 (56.3)	0 (0.0)	0 (0.0)	0 (0.0)	4.4/4	.51	100% (5,4)
Have high specificity and sensitivity	8 (50.0)	7 (43.8)	1 (6.3)	0 (0.0)	0 (0.0)	4.4/4.5	.62	83.8% (5,4)
In biomarker studies:								
Delirium cases should be diagnosed by a trained assessor or specialist doctor	6 (37.5)	9 (56.3)	1 (6.3)	0 (0.0)	0 (0.0)	4.2/4	.77	93.8% (5,4)

[1] One participant did not respond to this statement

	5	4	3	2	1	Mean/Median	SD	Agreement
Delirium should be assessed using a validated delirium diagnosis tool	13 (81.3)	2 (12.5)	1 (6.3)	0 (0.0)	0 (0.0)	4.6/5	1.02	81.3% (5)
Delirium should be prospectively evaluated	8 (50.0)	6 (37.5)	2 (12.5)	0 (0.0)	0 (0.0)	4.4/4.5	.71	87.5% (5,4)
Adult and paediatric populations should be considered separately	8 (50.0)	5 (31.3)	2 (12.5)	1 (6.3)	0 (0.0)	4.2/4.5	.93	81.3% (5,4)
In biomarker studies, confounding variables need to:								
Be decided a priori	5 (31.3)	8 (50.0)	3 (18.8)	0 (0.0)	0 (0.0)	4.1/4	.71	81.3% (5,4)
Take into account the population being studied/the clinical condition	12 (75.0)	4 (25.0)	0 (0.0)	0 (0.0)	0 (0.0)	4.7/5	.44	75% (5)
Be clearly defined and justified	13 (81.3)	3 (18.8)	0 (0.0)	0 (0.0)	0 (0.0)	4.8/5	.40	81.3% (5)
Be accounted for in the analysis	15 (93.8)	1 (6.3)	0 (0.0)	0 (0.0)	0 (0.0)	4.9/5	.50	93.8% (5)
The minimum clinical covariates that should be taken into account in delirium biomarker studies are:								
Age, gender, concurrent medication, comorbidities, prior cognitive impairment, prior neurological conditions, frailty, delirium risk and delirium precipitants	12 (75.0)	3 (18.8)	1 (6.3)	0 (0.0)	0 (0.0)	4.7/5	.60	75% (5)
Illness severity	14 (70.0)	5 (25.0)	1 (5.0)	0 (0.0)	0 (0.0)	4.6/5	.58	70% (5)
Sepsis	6 (30.0)	9 (45.0)	3 (15.0)	2 (10.0)	0 (0.0)	3.9/4	.94	75% (5,4)
Inflammation	7 (35.0)	10 (50.0)	1 (5.0)	2 (10.0)	0 (0.0)	4.1/4	.91	85% (5,4)
The following control groups are appropriate in a delirium biomarker study:								
Participants without delirium	10 (62.5)	5 (31.3)	1 (6.3)	0 (0.0)	0 (0.0)	4.5/5	.81	93.8% (5,4)
As delirium is a complex clinical condition with many influencing clinical variables several control groups will strengthen the ability to interpret the findings	7 (35.0)	7 (35.0)	3 (15.0)	3 (15.0)	0 (0.0)	3.9/4	1.07	70% (5,4)
Same illness severity with and without delirium	9 (45.0)	8 (40.0)	2 (10.0)	1 (5.0)	0 (0.0)	4.2/4	1.0	85% (5,4)
Delirium superimposed on dementia	6 (30.0)	8 (40.0)	3 (15.0)	2 (10.0)	1 (5.0)	3.7/4	1.2	70% (5,4)
In studies which follow participants longitudinally, appropriate additional comparator groups are:								
Participants with delirium of a shorter duration	4 (25.0)	8 (50.0)	3 (18.8)	1 (6.3)	0 (0.0)	3.9/4	.85	75% (5,4)
Participants who do not develop delirium	10 (62.5)	4 (25.0)	1 (6.3)	1 (6.3)	0 (0.0)	4.4/5	.89	87.5% (5,4)
Delirium biomarker studies should support the person with delirium and their proxy decision maker by:								

Clear participant information that explains the study to the person with delirium and/or their proxy decision maker	11 (68.8)	4 (25.0)	1 (6.3)	0 (0.0)	0 (0.0)	4.6/5	.81	93.8% (5,4)
Clear procedures to assist staff in interacting and supporting the patient during biomarker collection and other data collection	12 (75.0)	2 (12.5)	2 (12.5)	0 (0.0)	0 (0.0)	4.6/5	.71	75% (5)
The value of the research in lay terms and how it can contribute to the understanding of delirium	12 (75.0)	3 (18.8)	1 (6.3)	0 (0.0)	0 (0.0)	4.6/5	.80	75% (5)
Having clear processes for informed consent	12 (75.0)	3 (18.8)	1 (6.3)	0 (0.0)	0 (0.0)	4.6/5	.80	75% (5)
Description of the assay procedure should include the following as a minimum:								
A detailed assay protocol that includes the reagents/kits used	11 (68.8)	2 (12.5)	2 (12.5)	1 (6.3)	0 (0.0)	4.4/5	.96	81.3% (5,4)
An assay validation for assay repeatability and robustness	6 (37.5)	6 (37.5)	3 (18.8)	1 (6.3)	0 (0.0)	4.0/4	.92	75% (5,4)
The inter- and intra- assay coefficients of variation	7 (43.8)	5 (31.3)	2 (12.5)	2 (12.5)	0 (0.0)	4.0/4	1.06	75.6% (5,4)
Methods of preservation, storage and processing of the biological sample	11 (68.8)	3 (18.8)	1 (6.3)	1 (6.3)	0 (0.0)	4.5/5	.89	87.6% (5,4)
The assay validity	8 (50.0)	7 (43.8)	1 (6.3)	0 (0.0)	0 (0.0)	4.4/4.5	.62	93.8% (5,4)
The sensitivity limits of the assay	9 (56.3)	6 (37.5)	1 (6.3)	0 (0.0)	0 (0.0)	4.4/5	.81	93.8% (5,4)
A scoring and reporting protocol	8 (50.0)	6 (37.5)	2 (12.5)	0 (0.0)	0 (0.0)	4.4/4.5	.71	87.5% (5,4)
Please indicate your level of agreement with the following statements								
In biomarker studies:								
Blinding of the assay is essential if the clinical outcome is subjective	12 (75.0)	2 (12.5)	1 (6.3)	1 (6.3)	0 (0.0)	4.6/5	.89	75% (5)
Method of blinding should be explicit	9 (56.3)	4 (25.0)	2 (12.5)	1 (6.3)	0 (0.0)	4.3/5	.94	81.3% (5,4)
Timing of the sample collection should be determined based on the clinical scenario	6 (37.5)	8 (50.0)	2 (12.5)	0 (0.0)	0 (0.0)	4.2/4	.68	87.5% (5,4)
Timing of the sample collection should be determined based on the hypothesis being tested	12 (75.0)	3 (18.8)	1 (6.3)	0 (0.0)	0 (0.0)	4.7/5	.60	75% (5)

Statement								
In longitudinal sampling of populations AT RISK OF DELIRIUM, it is recommended that samples are collected prior to delirium onset, during delirium episode, and after delirium resolution	9 (56.3)	7 (43.8)	0 (0.0)	0 (0.0)	0 (0.0)	4.6/5	.51	100% (5,4)
In longitudinal sampling of populations WITH DELIRIUM, it is recommended that samples are collected at delirium onset and again after delirium resolution	6 (37.5)	8 (50.0)	1 (6.3)	1 (6.3)	0 (0.0)	4.2/4	.83	87.5% (5,4)
Please indicate your level of agreement with the following statements on sample size in a delirium biomarker study.								
Sample size should be decided a priori based on previous studies/pilot data	6 (37.5)	7 (43.8)	2 (12.5)	1 (6.3)	0 (0.0)	4.1/4	.88	81.3% (5,4)
Sample size should be determined based on the estimated effect size of the biomarker in predicting the outcome	8 (50.0)	6 (37.5)	2 (12.5)	0 (0.0)	0 (0.0)	4.4/4.5	.71	87.5% (5,4)
The analysis plan should plan for clinical and biomarker missing data due to:								
Clinical issues such as overall deterioration, worsening cognition, and death	11 (68.8)	5 (31.3)	0 (0.0)	0 (0.0)	0 (0.0)	4.7/5	.47	100% (5,4)
Practical challenges of biomarker collection in people with delirium	12 (75.0)	4 (25.0)	0 (0.0)	0 (0.0)	0 (0.0)	4.8/5	.44	75% (5)
Univariate analyses of biomarker and clinical endpoints of interest should report the following:								
Estimated effect size	6 (37.5)	7 (43.8)	1 (6.3)	2 (12.5)	0 (0.0)	3.9/4	1.2	81.3% (5,4)
Whether biomarker result was dichotomised using a cut-point and/or threshold	11 (68.8)	3 (18.8)	1 (6.3)	1 (6.3)	0 (0.0)	4.4/5	1.09	87.6% (5,4)
How missing data were handled	12 (75.0)	2 (12.5)	1 (6.3)	1 (6.3)	0 (0.0)	4.5/5	1.09	75% (5)
Number of included participants	14 (87.5)	1 (6.3)	0 (0.0)	0 (0.0)	1 (6.3)	4.7/5	1.01	87.5% (5)
Multivariate analyses of biomarker and clinical endpoints of interest should report the following:								
Estimated effect size	8 (50.0)	8 (50.0)	0 (0.0)	0 (0.0)	0 (0.0)	4.5/4.5	.51	100% (5,4)
Whether biomarker result was dichotomised using a cut-point and/or threshold	11 (68.8)	5 (31.3)	0 (0.0)	0 (0.0)	0 (0.0)	4.7/5	.47	100% (5,4)
How model assumptions were verified	10 (62.5)	5 (31.3)	1 (6.3)	0 (0.0)	0 (0.0)	4.6/5	.62	93.8% (5,4)
How missing data were handled	12 (75.0)	3 (18.8)	1 (6.3)	0 (0.0)	0 (0.0)	4.7/5	.60	75% (5)

Number of included participants	15 (93.8)	1 (6.3)	0 (0.0)	0 (0.0)	0 (0.0)	4.9/5	.25	93.8% (5)
Covariates (including how they were defined)	14 (87.5)	2 (12.5)	0 (0.0)	0 (0.0)	0 (0.0)	4.9/5	.34	87.5% (5)

*Red coloured items indicate those that arose from participant suggestions/comments

Table 3. Summary of ratings for items that did NOT reach a consensus after three rounds of Delphi*

Statement	Very important	Moderately important	Not important or unimportant	Slightly important	Not important at all	Mean rating/Median rating	SD
The following control groups are appropriate in a delirium biomarker study:							
Healthy participants matched by baseline characteristics such as age and gender	3 (15.0)	8 (40.0)	3 (15.0)	5 (25.0)	1 (5.0)	3.3/4.0	1.18
Participants with dementia, without delirium	4 (20.0)	9 (45.0)	5 (25.0)	1 (5.0)	1 (5.0)	3.7/4.0	1.03
In studies which follow participants longitudinally, an appropriate additional comparator group is:							
Participants with less severe delirium	3 (15.0)	6 (30.0)	8 (40.0)	3 (15.0)	0 (0.0)	3.4/3.0	.94
Description of the assay procedure should include:							
Information about where the kit was purchased and whether it was commercially available	4 (20.0)	9 (45.0)	4 (20.0)	3 (15.0)	0 (0.0)	3.7/4.0	.97
The minimum clinical covariates that should be taken into account in delirium biomarker studies are:							
Ethnicity/race	3 (15.0)	6 (30.0)	6 (30.0)	3 (15.0)	2 (10.0)	3.2/3.0	1.20
Education [2]	4 (20.0)	9 (45.0)	3 (15.0)	1 (10.0)	1 (5.0)	3.6/4.0	1.10
Psychiatric history	4 (20.0)	8 (40.0)	4 (20.0)	2 (10.0)	2 (10.0)	3.5/4.0	1.23
Injuries	3 (15.0)	10 (50.0)	6 (30.0)	1 (5.0)	0 (0.0)	3.7/4.0	.78

*Round 3 results shown in this table
Red coloured items indicate those that arose from participant suggestions/comments.

[2] One participant did not respond to this statement

Table 4. The final list of recommendations for delirium biomarker studies

| **The study objective should include the following:** |
| The biomarker under study (including source) |
| The time of collection in relation to delirium onset |
| The clinical endpoint(s) including their definition |
| The clinical covariates |
| The methods of biomarker collection |
| A description of which delirium pathophysiological theory the study will address |
| **In defining the population:** |
| Delirium cases should be diagnosed by a trained assessor or specialist doctor |
| Delirium should be assessed using a validated delirium diagnosis tool |
| Delirium should be prospectively evaluated |
| Adult and paediatric populations should be considered separately |
| **Delirium biomarker studies should support the person with delirium and their proxy decision maker by:** |
| Providing a clear participant information that explains the study to the person with delirium and/or their proxy decision maker |
| Providing clear procedures to assist staff in interacting and supporting the patient during biomarker collection and other data collection |
| Explaining the value of the research in lay terms and how it can contribute to the understanding of delirium |
| Clear processes for informed consent |
| **When selecting control(s) group: study:** |
| 1. As delirium is a complex clinical condition with many influencing clinical variables several control groups will strengthen the ability to interpret the findings |
| 2. The following control groups would be appropriate to consider |
| a. Participants without delirium |
| b. Participants with the same illness severity, with and without delirium |

c.	Participants with delirium superimposed onto dementia
3.	In studies which follow participants longitudinally, the following are appropriate additional comparator groups:
a.	Participants with delirium of a shorter duration
b.	Participants who do not develop delirium

The biomarker in a delirium study should be:

Chosen a priori

Supported by a biologically plausible rationale

Supported by a clear hypothesis

The type of biological specimen chosen should:

Be based on the capacity to measure the proposed biological process being evaluated

Have high specificity and sensitivity

Description of the assay procedure should include the following as a minimum:

A detailed assay protocol that includes the reagents/kits used

An assay validation for assay repeatability and robustness

The inter- and intra- assay coefficients of variation

Methods of preservation, storage and processing of the biological sample

The assay validity

The sensitivity limits of the assay

A scoring and reporting protocol

Blinding of the assay is essential if the clinical outcome is subjective

Method of blinding should be explicit

In biomarker studies, confounding variables need to:

Be decided a priori

Take into account the population being studied/the clinical condition

Be clearly defined and justified

Be accounted for in the analysis

The minimum clinical covariates that should be taken into account are:

Age, gender, concurrent medication, comorbidities, prior cognitive impairment, illness severity, sepsis, prior neurological conditions, frailty, inflammation, delirium risk and delirium precipitants

Timing of collection

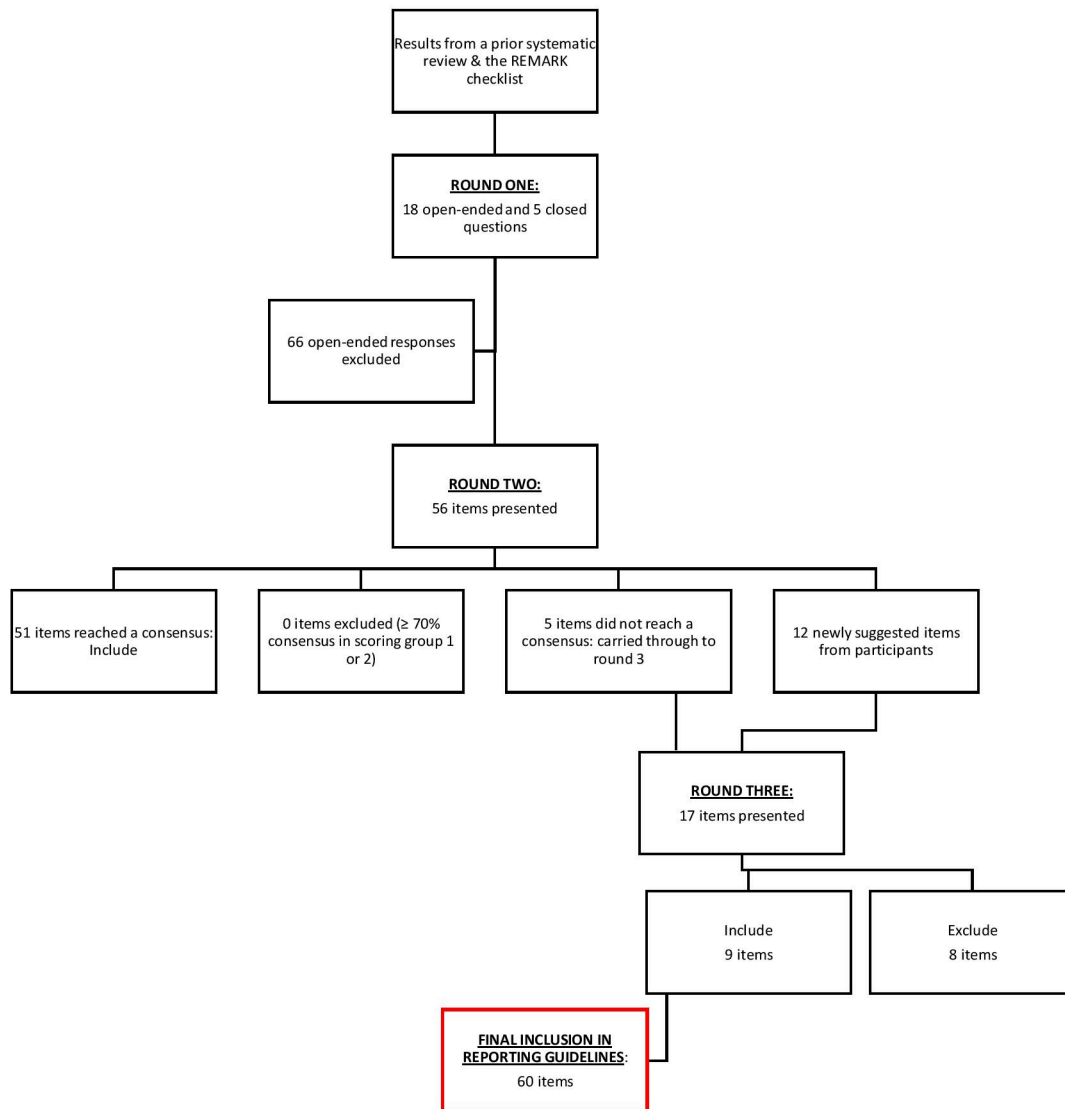| |
|---|
| Timing of the sample collection should be determined based on the clinical scenario and/or the hypothesis being tested |
| In longitudinal sampling of populations AT RISK OF DELIRIUM, it is recommended that samples are collected prior to delirium onset, during delirium episode, and after delirium resolution |
| In longitudinal sampling of populations WITH DELIRIUM, it is recommended that samples are collected at delirium onset and again after delirium resolution |
| **Sample size** |
| Sample size should be decided a priori based on previous studies/pilot data |
| Sample size should be determined based on the estimated effect size of the biomarker in predicting the outcome |
| **The analysis plan should plan for clinical and biomarker missing data due to:** |
| Clinical issues such as overall deterioration, worsening cognition, and death |
| Practical challenges of biomarker collection in people with delirium |
| **Univariate analyses of biomarker and clinical endpoints of interest should report the following:** |
| Estimated effect size |
| Whether biomarker result was dichotomised using a cut-point and/or threshold |
| How missing data were handled |
| Number of included participants |
| **Multivariate analyses of biomarker and clinical endpoints of interest should report the following:** |
| Estimated effect size |
| Whether biomarker result was dichotomised using a cut-point and/or threshold |
| How model assumptions were verified |
| How missing data were handled |
| Number of included participants |
| Covariates (including how they were defined) |

Figure 1. Flow chart illustrating the three-stage Delphi process, informed by a prior systematic review