BMJ Open

Protocol for developing a set of performance measures to monitor and evaluate delirium care quality for older adults in the emergency department using a modified e-Delphi process

Sarah Filiatreault ,[1] Sara Kreindler,[1] Jeremy Grimshaw,[2,3] Alecs Chochinov,[1,4] Malcolm Doupe[1,4]

To cite: Filiatreault S, Kreindler S, Grimshaw J, *et al*. Protocol for developing a set of performance measures to monitor and evaluate delirium care quality for older adults in the emergency department using a modified e-Delphi process. *BMJ Open* 2023;**13**:e074730. doi:10.1136/bmjopen-2023-074730

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2023-074730).

Received 14 April 2023
Accepted 10 August 2023

For numbered affiliations see end of article.

Correspondence to
Sarah Filiatreault;
filiatrs@myumanitoba.ca

ABSTRACT

Introduction Older adults are at high risk of developing delirium in the emergency department (ED). Delirium associated with an ED visit is independently linked to poorer outcomes such as increased length of hospital stay and mortality. Performance measures (PMs) are needed to identify variations in the quality of delirium care to help focus improvement efforts where they are most needed. A preliminary list of 11 quality statements and 24 PMs was developed based on a synthesis of high-quality clinical practice guidelines. The purpose of this study is to gain consensus on a subset of PMs that can be used to evaluate delirium care quality for older ED patients.

Methods and analysis This protocol for a modified e-Delphi study is informed by the Guidance on Conducting and REporting DElphi Studies. Clinical experts from across Canada and internationally will be recruited through peer referral, professional organisations and social media calls for expressions of interest. A minimum of 17 participants will be recruited. The primary survey for each round will consist of closed-ended questions with the opportunity to provide comments to justify decisions and clarify understanding. Using 9-point Likert scales, participants will rate each quality statement according to the concepts of importance and actionability, then its associated PMs according to the concept of necessity. Results will be fed back to participants in subsequent rounds. A priori stopping criteria have been defined in terms of consensus and stability. A minimum of three rounds will be undertaken to allow participants to have feedback, revise previous responses, then stabilise responses.

Ethics and dissemination Ethical approval was provided at the University of Manitoba Health Research Ethics Board (ID HS25728 (H2022:340)). Informed consent will be obtained electronically using the Research Electronic Data Capture secure online platform. Knowledge translation and dissemination will be done through traditional (eg, conference presentations, peer-reviewed publications) and non-traditional (eg, ED Grand Rounds) strategies.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ To ensure rigour and transparency in our research, this protocol is informed by the Guidance on Conducting and REporting DElphi Studies, which provides criteria in a similar format to other reporting checklists and is endorsed by the EQUATOR Network.
⇒ A priori criteria for consensus and stability have been defined for this study to reduce the subjectivity of stopping the Delphi process.
⇒ Recruiting a heterogeneous sample of clinical experts will help enrich the results and increase the credibility and acceptance of Performance measures (PMs) developed.
⇒ This study will use a formal consensus building process to develop a set of guideline-based PMs that are necessary to evaluate the quality of delirium care older emergency department patients receive.
⇒ Situational and personal biases can influence differences in how experts make judgements when using the Delphi method, which is a limitation we have attempted to reduce by planning to limit time between rounds, providing detailed background information and clear definitions for all concepts, providing quantitative and qualitative personalised feedback, as well as clearly defining consensus and stopping criteria.

INTRODUCTION

Older adults (ie, people 65 years of age and older) often use the emergency department (ED) as their first point of contact with the acute healthcare system.[1–4] Older ED patients are at high risk of developing delirium in this care environment[5–8]; however, it often goes undetected or undertreated.[9 10] Delirium is defined as a reversible 'syndrome of abrupt onset, fluctuating course, with prominent cognitive symptoms including decreased attention and awareness, additional deficits such as memory, or disorientation and evidence of an underlying physiological cause'.[11] Delirium associated with an ED visit is independently linked to poorer outcomes

such as loss of independence, increased length of hospital stay, and mortality.[8 12–17] A major barrier in improving care for older ED patients is the underlying knowledge gaps and lack of practice standards for assessing, recognising and managing delirium in the ED.[18] Mechanisms to evaluate ED practice performance are needed first to identify gaps and variations in quality care to focus delirium care improvement strategies where they are most needed.

Much of the research on healthcare quality is based on the seminal work of Donabedian, who defines healthcare provision in terms of structures (ie, conditions under which care is provided), processes (ie, diagnosis, treatment, rehabilitation and prevention of health conditions) and outcomes (ie, changes in an individual or population attributable to healthcare).[19] The Institute of Medicine defines quality of healthcare as 'the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge'.[20] This definition implies a dynamic association between structure, process and outcome.[20–22] The Donabedian structure-process-outcome framework has guided the development of most existing quality measurement systems because it provides a comprehensive, easily understandable model for classifying different types of measures.[23] Performance measures (PMs) are tools to quantify measurable aspects of practice performance (ie, structures and processes).[21 24] The extent that PMs are observed in practice provides an indication of the quality of the care provided and the likelihood of attaining optimal patient outcomes.[20 24 25]

There have been few attempts to establish PMs to evaluate quality of care for older ED patients in relation to delirium. The PMs that exist have been found to be of low methodological quality and predominately based on pre-existing metrics.[26–29] PMs are only as good as the evidence and methods used to develop them.[23 30] Poorly developed PMs can lead to unintended consequences by providing misleading information to guide decision-making, policy development and quality improvement efforts.[23] There is general agreement in the ED quality of care literature that there is a need to rigorously develop new evidence-based PMs instead of basing work on pre-existing metrics.[26 31 32]

Numerous researchers and organisations assert that clinical practice guidelines (CPGs) are an essential first step in developing quality statements (ie, concise statements defining best practice in a specific context), which in turn provide the basis for identifying PMs.[21 24 25 33–35] High-quality CPGs have the power to translate the complexity of scientific evidence into recommendations for practice, and therefore, are essential for developing appropriate metrics to evaluate care quality.[33] In the past decade, work has been done to increase the methodological rigour of developing guideline-based PMs.[21 24 36 37] Nothacker *et al*, as part of the Guidelines International Network, developed a framework for generating guideline-based PMs.[21] A summary of criteria articulated by Nothacker *et al* and how these align with our research is provided in online supplemental file 1.

Prior to the current study, we developed a preliminary set of PMs based on the results of an umbrella review which synthesised recommendations from delirium CPGs appraised to be high quality.[38] These recommendations were grouped into four categories of delirium care: screening, diagnosis, risk reduction and management. The synthesised recommendations provided an evidence-based foundation for the creation of a set of quality statements (N=11), and subsequent PMs (N=24), for delirium care in the ED. The next step in establishing a set of PMs for use, is to conduct a formal consensus process with a diverse panel of experts to finalise a set of PMs from the transformed recommendations.[21]

The purpose of this study is to gain consensus on a set of guideline-based PMs to monitor and evaluate delirium care quality for older ED patients. To achieve this, a modified electronic (e)-Delphi study will be conducted to reach consensus among key experts, including ED clinicians, ED decision-makers and geriatric specialists, on a set of ED PMs.

METHODS AND ANALYSIS

Delphi methods are used to synthesise knowledge and build consensus from a diverse group of experts through an iterative process of surveys and feedback.[39–42] A modified e-Delphi process is the most appropriate method for several reasons: (1) independent and anonymous participation reduces group bias, pressure to conform by dominate personalities and groupthink, (2) controlled feedback between rounds provides participants the opportunity to refine their judgements and promotes consensus, (3) including a heterogeneous sample of experts can improve process and results and (4) using an online platform with electronic surveys increases geographical diversity at an acceptable cost and reduces practical barriers to participating.[41 42]

Study steering group

We convened a steering group to provide study oversight with members consisting of coauthors (with methodological and clinical expertise), two patient/family representatives, and two further clinical experts. The purpose of including individuals beyond the coauthors in this group is to: (1) provide feedback on Delphi survey development, structure and clarity, (2) help identify potential Delphi participants and (3) provide additional important insights throughout the Delphi process as needed. Including patient/family representatives in the steering group helps support the patient-centredness of the quality statements and PMs, such as, by ensuring they contain aspects of care important to patients and families, and suggesting alternative terminology that better reflects patient views. Members of the steering group who are not coauthors were identified through our existing contacts and networks. These members will not have access to raw study data or be able to influence the study process.

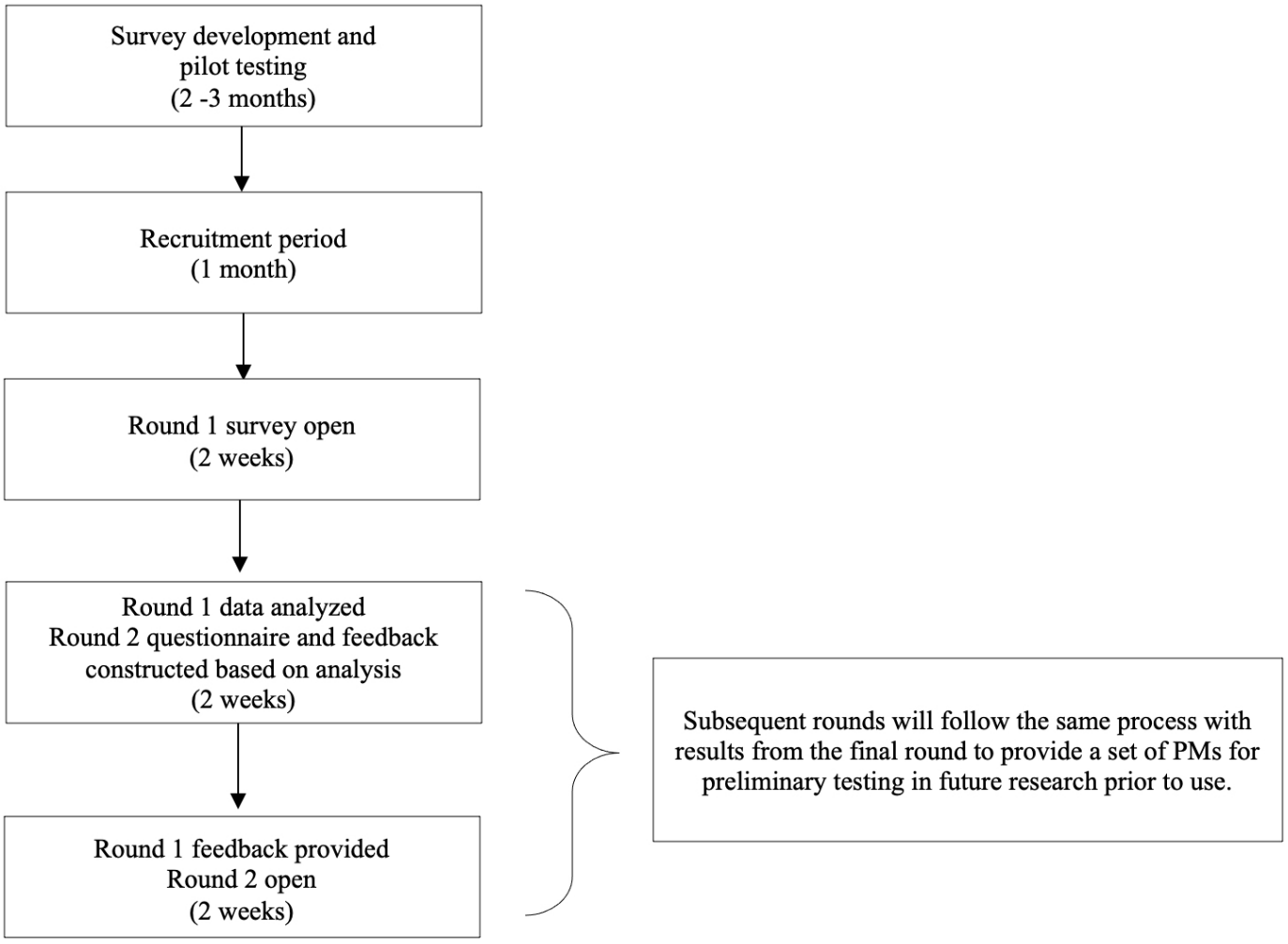

Figure 1 Delphi study flow. PMs, performance measures.

Study design

This e-Delphi study is informed by the Guidance on Conducting and REporting DElphi Studies[43] (online supplemental file 2), which is endorsed by the EQUATOR Network,[44] and other recommended criteria.[39 45] A summary of the proposed e-Delphi approach is presented in figure 1. Data collection is planned between April and July 2023. Participant fatigue and attrition tends to increase after three rounds using the Delphi method,[41] therefore, three Delphi rounds are proposed, however, a fourth round may be considered if a priori stopping criteria are not met. All rounds will be completed anonymously and electronically using the Research Electronic Data Capture (REDCap) secure online platform for building and managing online surveys[46] through the University of Manitoba licensing agreement.[47] In every round, Delphi participants will rate the importance and actionability of the quality statements, and the necessity of the related PMs using a 9-point Likert scale. Each participant will be sent a unique REDcap link to allow tracking of responses and participation in each round.

Study sample

Selection and identification of expert panel

Consistent with the goals of enriching the Delphi consensus results and increasing the credibility and acceptance of PMs developed through this process,[21 30 35–37 39 48] a heterogeneous sample of experts consisting of ED clinicians, clinical decision-makers, as well as geriatric specialists will be selected. System-level decision-makers (eg, regional directors or managers) and patients are not included as experts because the completion of the surveys requires specific clinical knowledge. Therefore, an 'expert' in this study is defined as one with clinical knowledge in the care of older adults in the ED. Participants will be considered for inclusion if they are able to read and write in English, willing to participate and meet one or more of the following criteria: clinical experience in a relevant field to the ED care of older adults for 5 or more years postbasic graduation, hold postgraduate qualifications or credentials relevant to the management of delirium in older adults, or are recognised by peers as an expert in the area (eg, member of a relevant organisation or network). Potential participants will be excluded if they have insufficient experience in a relevant area

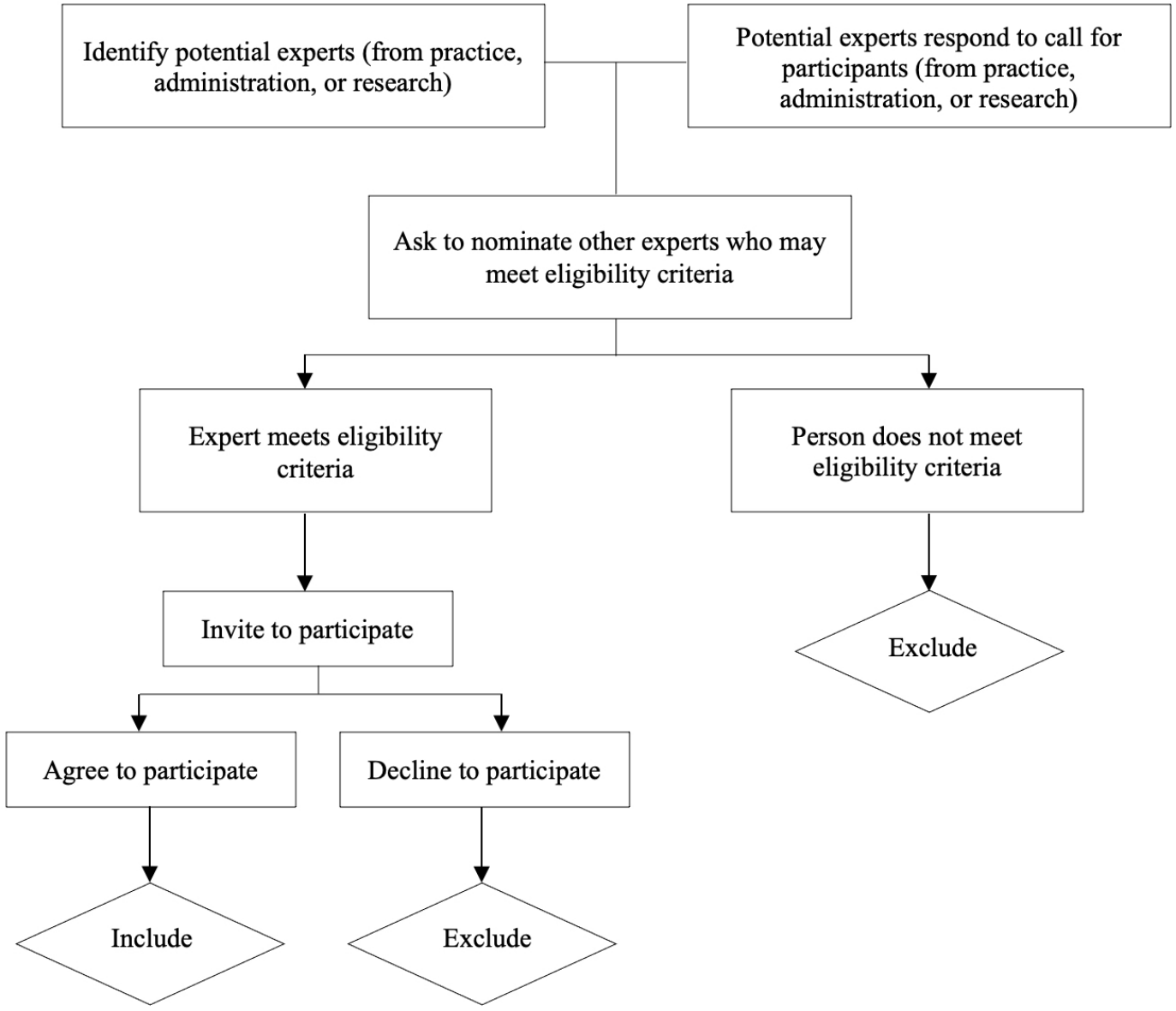

Figure 2 Recruitment process.

or are unable to commit to be available for the entire process.

Panel size and composition

There are no agreed criteria for the ideal sample size for a Delphi study; however, a larger sample size will increase the reliability of the group judgement.[39 41 42] It has been asserted that reliability decreases significantly in sample sizes of approximately six or less, and improvements in reliability are subject to diminishing returns in sample sizes greater than approximately 12.[41 42] Under the assumption that attrition is common in Delphi studies,[41] our minimum sample size (N=17) was calculated to allow for approximately 30% attrition and ensure at least 12 responses in the final round.

Recruitment

Experts will be recruited from across Canada, and internationally, using a multipronged approach (see figure 2), including identification of individuals by coauthors. Recruitment will be augmented by encouraging identified experts to snowball the invitation with other eligible participants. Relevant national professional associations and networks (eg, Canadian Association of Emergency Physicians and National Emergency Nurses Association) and international associations (eg, iDelirium) will be approached to request their support to distribute an advertisement for recruitment to the expert panel via social media (ie, LinkedIn, Twitter and/or Facebook) and email. To prevent overrepresentation from one expert group, recruitment will be monitored to ensure a heterogeneous sample of experts is recruited. Recruitment will be planned over a 4-week period, with a reminder sent after 2 weeks. If there is no response after 4 weeks, the invited expert will not be contacted again. If the minimum target number of potential participants is not recruited after the 4-week period (N=17), the recruitment timeline will be extended for another 4 weeks.

Screening for eligibility

All identified experts will be invited to participate. The potential participants will be sent materials to inform them of the study objectives and design, as well as the commitment required for participation. Potential participants will be screened for eligibility and asked to nominate peers that may be interested in participating at this point. Experts who confirm their interest and eligibility will be provided with a unique link to the REDCap system at the University of Manitoba, where a participant information sheet and consent form will be hosted. The participant information sheet will include information such as the purpose of the study, a clear explanation of the anticipated process, and an explanation that participation will be required over a period of 4 months with three rounds of questioning and feedback.[41] The information sheet will also reiterate the eligibility criteria, as well as explain anonymity and the withdrawal process. Participants will be asked to commit to participate in all rounds of the Delphi process, however, they will be allowed to withdraw at any time during data collection. Due to the nature of the Delphi process, responses will be used up to the point of withdrawal. Participants will be able to withdraw from the study by contacting the principal investigator (SF). Consent to participate will be obtained through REDCap.

Enhancing response rate

We will aim to achieve a 100% response rate in each round, however, recognising attrition and participant fatigue is common in studies using the Delphi method, a round will be accepted as valid if there is a response rate of 70%.[41] We will employ methods to minimise attrition. Participants will be invited to an initial interview with the option for it to be conducted virtually (ie, video) or by telephone to discuss initial questions, as well as to develop rapport.[41] For each round, the online survey will remain active for 2 weeks. After 1 week, an email reminder or thank-you will be sent.

Survey design and development

This study is a modified e-Delphi study, meaning participants are given preselected 'issues' (ie, quality statements and potential PMs) on which to make a judgement based on the current evidence.[41] Therefore, the primary survey will consist of closed-ended questions with the opportunity to provide comments for each quality statement and related PMs to justify decisions and clarify understanding. The survey will be developed in REDCap. Along with the survey, each round of the study will be accompanied by a cover letter to refamiliarise participants to the study, state the intentions of the round, and provide definitions for key concepts.

Development of the survey was informed by PM assessment tools (ie, AIRE[49] and QUALIFY[50] instruments), PM criteria used by organisations that develop and implement PMs (ie, the National Quality Forum[51] and NICE[25]), as well as syntheses of these sources.[21 23 30] As the PMs are the operationalisable form of the quality statements,[21 24 25 33–35] both will be judged by participants. Participants will be asked to judge each quality statement according to its importance (ie, relevant and of crucial value to the care of older adults in the ED) and actionability (ie, care can be done by providers in the ED setting with appropriate resources and tools). Then, they will be asked to judge each related PM according to its necessity (ie, the PM is necessary to evaluate the associated quality statement). The quality statements and PMs will be scored using these dimensions, on a Likert scale ranging from 1 to 9. Delphi participants will be advised to think of each 9-point scale being made up of three parts (ie, tertiles), where 1–3 could be used to record low ratings (ie, not at all important, actionable and/or necessary), 4–6 record average ratings (ie, somewhat important, actionable and/or necessary) and 7–9 record high ratings (ie, very important, actionable and/or necessary). For participants to provide an informed judgement, a rationale for the quality statement and PMs, including a brief summary of the evidence, will be provided.

Prior to implementation, the Delphi survey was piloted with clinical expert steering group members. The purpose of this process was to ensure that the PMs are clearly and precisely worded with unambiguous language, and that each set of PMs reflect the quality statement they are meant to measure. Those involved in the pilot were also asked questions such as, what did you like, what didn't you like, and how long did the survey take to complete? Adjustments were made to the survey based on feedback from all steering group members and agreed on by study coauthors. For example, as quality statements are defined as 'concise statements defining best practice in a specific context' it was agreed to call the quality statements 'best practices' in the Delphi survey to decrease the number of new concepts introduced to participants and to increase comprehensibility and readability.

Procedures

A minimum of three rounds will be undertaken to allow participants to receive feedback, revise previous responses, then stabilise responses.[41 52] A decision will be made after the third round if it is appropriate to stop according to a priori definitions of consensus and stability, due to the need to balance a high response rate with the need to achieve stability and consensus.

Defining consensus and stability

Establishing a priori criteria for defining consensus and stability reduces the subjectivity of stopping criteria.[43 45 52 53] There is no universal definition of consensus,[45] however, it is important to define the criteria for consensus that are suitable for the study objectives a priori.[43] In concordance with PM development frameworks,[30 54] we will use the RAND criteria for agreement to define consensus.[55] Consensus is defined as 80% of ratings within the 3-point tertile of the overall median. The lower tertile[1–3] represents scores that are 'not at all', the middle tertile[4–6] represents scores that are 'somewhat', and the upper

tertile[7–9] represents scores that are 'very' important, actionable and/or necessary.

Consensus has been stated to be meaningless if stability is not also achieved.[52 56] Therefore, a measurement of stability will be used as a stopping criterion for the Delphi process. In this study, stability will be defined as the consistency of responses between successive rounds (ie, no meaningful change).[41 52 56] Meaningful change is defined as a median change between tertiles *and* a greater than 15% change in the percentage of participants whose scores changed tertiles.[52 56]

Stopping and PM removal criteria

For the overall study, the criterion to stop the Delphi process is defined as no meaningful change in scores between the current and preceding round on at least 75% of quality statements and PMs assessed. Using a cut-off of stability in three-quarters of the quality statements and PMs acknowledges achieving 100% stability may not be feasible, and provides flexibility to balance participant workload.[41 52 56] Additionally, criteria for PM removal will be considered after the second and third rounds: To be removed from the process, a PM's scores must show no meaningful change from the previous round, and there must be consensus that the PM is not necessary (ie, overall panel median of 1 to 3, with 80% of ratings in the lower tertile). Regardless of scores, all quality statements will remain in all rounds.

Round 1

The objective of round 1 is to obtain participant demographic data to describe the sample and collect data on initial judgements on the quality statements and potential PMs. Participants will complete a Participant Details Form collecting information on age, sex, geographical location, professional background, highest level of education, work setting and years of experience. All PMs assessed in the first round will be retained for the next round.

Round 2

The objective of round 2 is to reach consensus on quality statements and PMs by enabling participants to reflect on their score considering the viewpoint of other experts in the group. All participants will be invited to round 2, including those who did not complete round 1, unless they have withdrawn from the study. This provides participants the opportunity to continue their involvement if they choose.[41] A personalised survey will be sent to each participant with: quantitative group results (ie, median, minimum and maximum ratings) presented numerically and graphically, qualitative feedback (ie, summary of participants' comments), and the participant's own response to illustrate their position in relation to the group. The PMs that meet stopping criteria will be removed from the process after round 2. All other PMs will remain in the third round to achieve or strengthen consensus on a final set of PMs.

Round 3

The objective of round 3 is to strengthen consensus on a final set of quality statements and PMs. All participants will be invited to round 3, unless they have withdrawn from the study. Similar to the second round, a new personalised survey will be sent to each participant again with the revised: quantitative group results presented numerically and graphically, qualitative feedback and the participant's own response to illustrate their position in relation to the group.

Potential subsequent round

If the stopping criteria are not met after three rounds, a fourth round will be considered. This will be done to gain stability in responses, as well as reach consensus or gain greater understanding about the lack of consensus. The Delphi will be stopped if responses are stable and consensus is still not achieved after the third round, signifying persistent disagreement, as these results may also provide informative insights to guide future research.[43]

Data analysis

Data will be analysed during the Delphi procedure to provide participants with feedback, to exclude PMs from subsequent rounds, and to inform the need for a fourth round. The final Delphi results will then be analysed to directly achieve the study purpose.

Data analyses conducted during the Delphi process

There are two types of data analysis that will be conducted during the Delphi process. First, data will be analysed after each round to generate feedback reports for participants as described above. Descriptive statistics will be used to summarise responses to be fed-back to participants. Content analysis will be used to summarise participants' comments into themes to be fed-back to the group.[41 57] Original wording from one expert that best represents the wording across participants with similar statements will be used where possible.[41]

Second, data will be analysed to assess stability as a criterion for stopping the Delphi process, as well as for potentially removing PMs after the second and third rounds. Data will be descriptively analysed with frequency distributions generated to determine if there is meaningful change in the central tendency (ie, median change between tertiles) or dispersion of scores (ie, ≥15% change in participants responses between tertiles) between the current and preceding round. Results will then be examined to determine if stopping or PM removal criteria are met.

Analysing final Delphi results

The primary analytical goal is to identify quality statements that are important and actionable for quality delirium care of older ED patients and generate a set of related PMs necessary to monitor and evaluate this care. This will be done using the concepts of importance, actionability and necessity to reach consensus. Descriptive statistics will be used to evaluate consensus. To be selected

for preliminary testing in future research, a quality statement and related PMs must reach consensus in the upper tertile (ie, overall panel median of 7–9, with 80% of ratings within the 3-point tertile of the overall median).

Once the Delphi is stopped and results are analysed, quality statements that are very important and actionable and related PMs that are very necessary according to group consensus will be selected for preliminary testing in future research. Those that achieve consensus just below the a priori thresholds for final selection but are believed to be relevant will be considered a posteriori as long as there is sufficient justification to do so.[45] This will be done by members of the study steering group by examining the analysed qualitative data, along with the quantitative results to help make these determinations.

Patient and public involvement

Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the 'study steering group' subsection for further details.

ETHICS AND DISSEMINATION

This study received approval from the University of Manitoba Health Research Ethics Board (ID HS25728 (H2022:340)). All data will be collected and stored online through the REDCap Survey server, which is a secure platform that abides by Canadian privacy laws and implements several security safeguards. Informed consent will be obtained from all participants before completing any surveys.

Engagement with diverse experts throughout this study will help to ensure that our results are translatable to everyday ED care. Knowledge translation and dissemination will be done through traditional (eg, conference presentations, peer-reviewed publications) and non-traditional (eg, ED Grand Rounds) strategies. This study will provide a set of quality statements and related PMs that are based on scientifically sound evidence, and are important, actionable and necessary to a diverse group of clinical experts. Results will be used in future research to test the feasibility of using the PMs to evaluate delirium care quality older ED patients receive and guide improvement efforts.

Author affiliations
[1]Community Health Sciences, University of Manitoba Faculty of Health Sciences, Winnipeg, Manitoba, Canada
[2]Clinical Epidemiology Program, Ottawa Health Research Institute, Ottawa, Ontario, Canada
[3]Department of Medicine, University of Ottawa, Ottawa, Ontario, Canada
[4]Emergency Medicine, University of Manitoba Max Rady College of Medicine, Winnipeg, Manitoba, Canada

Twitter Sarah Filiatreault @filiatrs

Acknowledgements The authors wish to thank and acknowledge Jessica Leigh RN, and Dr. Kerri Onotera (clinical experts), as well as Carli Rossall and one member who has chosen to remain anonymous (patient and family representatives) who agreed to be members of the study steering group.

Contributors SF conceived the research idea under the supervision of MD, SK, JG and AC. The manuscript was first drafted by SF and MD, additional content and draft reviews were provided by SK, JG and AC. All authors reviewed and approved the final version.

Funding SF: The Canadian Institutes of Health Research (CIHR) Frederick Banting and Charles Best Canada Graduate Scholarships-Doctoral Research Award (CGS-D) (Award #: FRN-175944), and the Canadian Nurses Foundation (CNF) Dr. Ann C. Beckingham award (Award #: N/A). JG: Canada Research Chair in Health Knowledge Transfer and Uptake (Grant #: 950-200986) and CIHR Foundation grant (Grant #: FDN-143269). SK: Manitoba Research Chair in Health System Innovation (Research Manitoba; Grant #: N/A).

Competing interests None declared.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iD
Sarah Filiatreault http://orcid.org/0000-0002-8442-1990

REFERENCES

1 Greenwald PW, Stern ME, Rosen T, *et al*. Trends in short-stay hospitalizations for older adults from 1990 to 2010: implications for geriatric emergency care. *Am J Emerg Med* 2014;32:311–4.
2 Greenwald PW, Estevez RM, Clark S, *et al*. The ED as the primary source of hospital admission for older (but not younger) adults. *Am J Emerg Med* 2016;34:943–7.
3 Samaras N, Chevalley T, Samaras D, *et al*. Older patients in the emergency Department: A review. *Ann Emerg Med* 2010;56:261–9. 10.1016/j.annemergmed.2010.04.015 Available: http://www.sciencedirect.com/science/article/pii/S0196064410003859
4 Canadian Institute for Health Information [CIHI]. *Health care in Canada, 2011: A focus on seniors and aging*. Ottawa, ON: CIHI, 2011. Available: https://secure.cihi.ca/free_products/HCIC_2011_seniors_report_en.pdf
5 Cole MG, Ciampi A, Belzile E, *et al*. Persistent delirium in older hospital patients: A systematic review of frequency and prognosis. *Age Ageing* 2009;38:19–26.
6 Hsieh SJ, Madahar P, Hope AA, *et al*. Clinical deterioration in older adults with delirium during early Hospitalisation: A prospective cohort study. *BMJ Open* 2015;5:e007496. 10.1136/bmjopen-2014-007496 Available: http://bmjopen.bmj.com/content/5/9/e007496.full.pdf+html
7 Griffith LE, Gruneir A, Fisher K, *et al*. Patterns of health service use in community living older adults with dementia and comorbid conditions: A population-based retrospective cohort study in Ontario, Canada. *BMC Geriatr* 2016;16:177. 10.1186/s12877-016-0351-x Available: https://doi.org/10.1186/s12877-016-0351-x
8 Kennedy M, Enander RA, Tadiri SP, *et al*. Delirium risk prediction, Healthcare use and mortality of elderly adults in the emergency Department. *J Am Geriatr Soc* 2014;62:462–9.
9 Barron EA, Holmes J. Delirium within the emergency care setting, occurrence and detection: a systematic review. *Emerg Med J* 2013;30:263–8. 10.1136/emermed-2011-200586 Available: http://emj.bmj.com/content/30/4/263

10 Boucher V, Lamontagne M-E, Nadeau A, *et al*. Unrecognized incident delirium in older emergency Department patients. *The Journal of Emergency Medicine* 2019;57:535–42. 10.1016/j.jemermed.2019.05.024 Available: https://www.sciencedirect.com/science/article/pii/S0736467919304196
11 Neufeld KJ, Thomas C. Delirium: definition, epidemiology, and diagnosis. *J Clin Neurophysiol* 2013;30:438–42. 10.1097/WNP.0b013e3182a73e31 Available: https://journals.lww.com/clinicalneurophys/Abstract/2013/10000/Delirium___Definition,_Epidemiology,_and_Diagnosis.3.aspx
12 Hirschman KB, Paik HH, Pines JM, *et al*. Cognitive impairment among older adults in the emergency Department. *West J Emerg Med* 2011;12:56–62.
13 Han JH, Shintani A, Eden S, *et al*. Delirium in the emergency Department: an independent Predictor of death within 6 months. *Ann Emerg Med* 2010;56:244–252. 10.1016/j.annemergmed.2010.03.003 Available: http://www.sciencedirect.com/science/article/pii/S0196064410002209
14 Cole MG, McCusker J, Bailey R, *et al*. Partial and no recovery from delirium after hospital discharge predict increased adverse events. *Age Ageing* January 8, 2017.
15 Bo M, Bonetto M, Bottignole G, *et al*. Postdischarge clinical outcomes in older medical patients with an emergency Department stay-associated delirium onset. *J Am Geriatr Soc* 2016;64:e18–9.
16 Émond M, Boucher V, Carmichael P-H, *et al*. Incidence of delirium in the Canadian emergency Department and its consequences on hospital length of stay: a prospective observational Multicentre cohort study. *BMJ Open* 2018;8:e018190.
17 Émond M, Grenier D, Morin J, *et al*. Emergency Department stay associated delirium in older patients. *Can Geriatr J* 2017;20:10–4.
18 Eagles D. Delirium in older emergency Department patients. *CJEM* 2018;20:811–2. 10.1017/cem.2018.468 Available: http://www.cambridge.org/core/journals/canadian-journal-of-emergency-medicine/article/delirium-in-older-emergency-department-patients/480D2CC4E413A8E20D097999BD441AF4
19 Donabedian A. *An introduction to quality assurance in health care*. New York, NY: Oxford University Press, 2003.
20 Institute of Medicine (US) Committee on Quality of Health Care in America [IOM]. *Crossing the quality chasm: A new health system for the 21st century*. Washington, DC: National Academies Press (US), 2001. Available: http://www.ncbi.nlm.nih.gov/books/NBK222274
21 Nothacker M, Stokes T, Shaw B, *et al*. Reporting standards for guideline-based performance measures. *Implement Sci* 2016;11:6. 10.1186/s13012-015-0369-z Available: http://implementationscience.biomedcentral.com/articles/10.1186/s13012-015-0369-z
22 Rubin HR, Pronovost P, Diette GB. The advantages and disadvantages of Process-Based measures of health care quality. *Int J Qual Health Care* 2001;13:469–74. 10.1093/intqhc/13.6.469 Available: https://doi.org/10.1093/intqhc/13.6.469
23 Busse R, Klazinga N, Panteli D, et al., eds. *Improving healthcare quality in Europe: Characteristics, effectiveness and implementation of different strategies*. United Kingdom: OECD, 2019.
24 Parmelli E, Langendam M, Piggott T, *et al*. Guideline-based quality assurance: a conceptual framework for the definition of key elements. *BMC Health Serv Res* 2021;21. 10.1186/s12913-021-06148-2 Available: https://doi.org/10.1186/s12913-021-06148-2
25 National Institute for Health and Care Excellence. *Health and social care directorate quality standards process guide*. Manchester: National Institute for Health and Care Excellence, 2020. Available: www.nice.org.uk
26 Burkett E, Martin-Khan MG, Gray LC. Quality indicators in the care of older persons in the emergency Department: A systematic review of the literature. *Australas J Ageing* 2017;36:286–98. 10.1111/ajag.12451 Available: http://onlinelibrary.wiley.com.proxy.hil.unb.ca/doi/10.1111/ajag.12451/abstract
27 Schnitker LM, Martin-Khan M, Burkett E, *et al*. Structural quality indicators to support quality of care for older people with cognitive impairment in emergency departments. *Acad Emerg Med* 2015;22:273–84. 10.1111/acem.12617 Available: http://doi.wiley.com/10.1111/acem.2015.22.issue-3
28 Schnitker LM, Martin-Khan M, Burkett E, *et al*. Process quality indicators targeting cognitive impairment to support quality of care for older people with cognitive impairment in emergency departments. *Acad Emerg Med* 2015;22:285–98. 10.1111/acem.12616 Available: http://doi.wiley.com/10.1111/acem.2015.22.issue-3
29 Schuster S, Singler K, Lim S, *et al*. Quality indicators for a geriatric emergency care (Geriq-ED) – an evidence-based Delphi consensus approach to improve the care of geriatric patients in the emergency Department. *Scand J Trauma Resusc Emerg Med* 2020;28:68. 10.1186/s13049-020-00756-3 Available: https://doi.org/10.1186/s13049-020-00756-3
30 Campbell SM, Kontopantelis E, Hannon K, *et al*. Framework and indicator testing protocol for developing and piloting quality indicators for the UK quality and outcomes framework. *BMC Fam Pract* 2011;12.
31 Hansen K, Boyle A, Holroyd B, *et al*. Updated framework on quality and safety in emergency medicine. *Emerg Med J* 2020;37:437–42. 10.1136/emermed-2019-209290 Available: http://emj.bmj.com/content/37/7/437
32 Schull MJ, Guttmann A, Leaver CA, *et al*. Prioritizing performance measurement for emergency Department care: consensus on evidence based quality of care indicators. *CJEM* 2011;13:300–9. 10.2310/8000.2011.110334 Available: https://www.cambridge.org/core/journals/canadian-journal-of-emergency-medicine/article/prioritizing-performance-measurement-for-emergency-department-care-consensus-on-evidencebased-quality-of-care-indicators/0C9DF3CCA4F8AAB29A64FC83F7ECF8D6
33 Institute of Medicine (US) Committee on Standards for Developing Trustworthy Clinical Practice Guidelines [IOM]. *Clinical Practice Guidelines We Can Trust*. Washington, DC: National Academies Press, 2011. Available: http://www.ncbi.nlm.nih.gov/books/NBK209539
34 Woolf SH, Grol R, Hutchinson A, *et al*. Potential benefits, limitations, and harms of clinical guidelines. *BMJ* 1999;318:7182:527–30.:.
35 Campbell SM, Braspenning J, Hutchinson A, *et al*. Research methods used in developing and applying quality indicators in primary care. *Qual Saf Health Care* 2002;11:358–64. 10.1136/qhc.11.4.358 Available: http://qualitysafetybeta.bmj.com/content/11/4/358
36 Langendam MW, Piggott T, Nothacker M, *et al*. Approaches of integrating the development of guidelines and quality indicators: a systematic review. *BMC Health Serv Res* 2020;20:875. 10.1186/s12913-020-05665-w Available: http://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-020-05665-w
37 Kötter T, Blozik E, Scherer M. Methods for the guideline-based development of quality indicators--a systematic review. *Implementation Sci* 2012;7. 10.1186/1748-5908-7-21 Available: http://implementationscience.biomedcentral.com/articles/10.1186/1748-5908-7-21
38 Filiatreault S, Grimshaw JM, Kreindler SA, *et al*. A critical appraisal and recommendation synthesis of delirium clinical practice guidelines relevant to the care of older adults in the emergency Department: an umbrella review. *J Eval Clin Pract* 14, 2023. 10.1111/jep.13883 Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/jep.13883
39 Boulkedid R, Abdoul H, Loustau M, *et al*. Using and reporting the Delphi method for selecting Healthcare quality indicators: A systematic review. *PLOS ONE* 2011;6:e20476. 10.1371/journal.pone.0020476 Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0020476
40 Brown BB. *Delphi Process: A Methodology Used for the Elicitation of Opinions of Experts*. Santa Monica, CA: RAND Corporation, 1968. Available: https://www.rand.org/pubs/papers/P3925.html
41 Keeney S, Hasson F, McKenna H. *The delphi technique in nursing and health research*. John Wiley & Sons, Ltd, 2011.
42 Murphy MK, Black NA, Lamping DL, *et al*. Consensus development methods, and their use in clinical guideline development. *Health Technol Assess* 1998;2:i–iv,
43 Jünger S, Payne SA, Brine J, *et al*. Guidance on conducting and reporting Delphi studies (CREDES) in palliative care: recommendations based on a methodological systematic review. *Palliat Med* 2017;31:684–706. 10.1177/0269216317690685 Available: https://doi.org/10.1177/0269216317690685
44 The EQUATOR network | enhancing the quality and transparency of health research. 2017. Available: http://www.equator-network.org/
45 Diamond IR, Grant RC, Feldman BM, *et al*. Defining consensus: a systematic review recommends Methodologic criteria for reporting of Delphi studies. *J Clin Epidemiol* 2014;67:401–9.
46 REDCap. 2021. Available: https://www.project-redcap.org/
47 University of Manitoba - information services and technology - enterprise applications - database services - Redcap consultation. 2021. Available: https://umanitoba.ca/computing/ist/service_catalogue/enterprise_app/database/3210.html
48 Rubin HR, Pronovost P, Diette GB. Methodology matters. from a process of care to a measure: the development and testing of a quality indicator. *Int J Qual Health Care* 2001;13:489–96.
49 de Koning JS, Kallewaard M, Klazinga NS. Prestatie-Indicatoren Langs de Meetlat – Het AIRE instrument. *TVGW* 2007;85:261–4. 10.1007/BF03078683 Available: https://doi.org/10.1007/BF03078683
50 Reiter A, Fischer B, Kötting J, *et al*. QUALIFY: ein instrument Zur Bewertung von Qualitätsindikatoren. Zeitschrift Für Ärztliche Fortbildung und Qualität Im Gesundheitswesen. *German J Qual Health Care* 2008;101:683–8. 10.1016/j.zgesun.2007.11.003

Available: https://www.sciencedirect.com/science/article/pii/S1431762107002862
51 National Quality Forum. NQF: measure evaluation criteria. 2021. Available: https://www.qualityforum.org/Measuring_Performance/Submitting_Standards/Measure_Evaluation_Criteria.aspx
52 von derHA. Consensus measurement in Delphi studies: review and implications for future quality assurance. *Technol Forecast Soc Change* 2012;79:1525–36. 10.1016/j.techfore.2012.04.013 Available: https://www.sciencedirect.com/science/article/pii/S0040162512001023
53 Holey EA, Feeley JL, Dixon J, *et al*. An exploration of the use of simple Statistics to measure consensus and stability in Delphi studies. *BMC Med Res Methodol* 2007;7:52. 10.1186/1471-2288-7-52 Available: https://doi.org/10.1186/1471-2288-7-52
54 Carinci F, Van Gool K, Mainz J, *et al*. Towards actionable International comparisons of health system performance: expert revision of the OECD framework and quality indicators. *Int J Qual Health Care* 2015;27:137–46.
55 Fitch K, Bernstein SJ, Aguilar MD, *et al*. *The RAND/UCLA appropriateness method user's manual*. Santa Monica, CA: RAND Corporation, 2001. Available: https://www.rand.org/pubs/monograph_reports/MR1269.html
56 Dajani JS, Sincoff MZ, Talley WK. Stability and agreement criteria for the termination of Delphi studies. *Technological Forecasting and Social Change* 1979;13:83–90. 10.1016/0040-1625(79)90007-6 Available: https://www.sciencedirect.com/science/article/pii/0040162579900076
57 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. 10.1111/nhs.12048 Available: http://onlinelibrary.wiley.com/doi/abs/10.1111/nhs.12048

Supplemental File 1. Summary of Nothacker et al. (2016) Criteria for Developing Guideline-Based PMs Aligned with Filiatreault et al. Research Phases

Criterion	Description	Key Attribute	Research Phase
Consider key attributes of scientifically sound, important/relevant, and feasible	Key attributes need to be taken into consideration during design and conduct of PM development and testing as identified in international approaches for PM appraisal	.	All Phases
Involve relevant stakeholders	Similar to principles of CPG development. Best method of involving patients is a subject for further research	Importance/ relevance	All Phases
Identify and critically appraise CPGs	Use high-quality CPGs as assessed with tool such as the AGREE-II, also ensure currency of included CPGs	Scientific soundness	Umbrella Review (completed)
Select/synthesize of CPG recommendations	State the strength of the evidence and/or grade of CPG recommendations	Scientific soundness	Umbrella Review (completed)
Process for PM selection from transformed recommendations	Use a formal consensus process with diverse panel of experts to reduce bias	Importance/ relevance & Scientific soundness	**Modified e-Delphi (current study)**
Identify intended uses	PMs to be prioritized for implementation will differ depending on the intended use (e.g., internal quality improvement versus public reporting)	Importance/ relevance	Future Research
Specify PMs	Unambiguously define numerator and denominator. State inclusion and exclusion criteria	Feasibility	Future Research
Conduct preliminary testing	No standard definition, but consider using an established process (e.g., QOF)	Scientific soundness & Feasibility	Future Research
Plan to review PMs	Ideally done simultaneously with CPG review and recommendation update	Scientific soundness	Future Research

Abbreviations: CPGs, clinical practice guidelines; AGREE-II, Appraisal of Guidelines, Research, and Evaluation-II instrument; PM, performance measure; QOF, Quality and Outcomes Framework. Reference: Nothacker M, et al. Reporting standards for guideline-based performance measures. Implementation Science. 2016;11(6). Available from: http://implementationscience.biomedcentral.com/articles/10.1186/s13012-015-0369-z

Supplementary File 2. CREDES Criteria for Delphi Studies

Item	Topic	Description	Line # Reported
Rationale for choice of Delphi method			
1	Justification	The choice of the Delphi technique as a method of systematically collating expert consultation and building consensus needs to be well justified. When selecting the method to answer a particular research question, it is important to keep in mind its constructivist nature	59-67
Planning & design			
2	Planning and process	The Delphi technique is a flexible method and can be adjusted to the respective research aims and purposes. Any modifications should be justified by a rationale and be applied systematically and rigorously	159-183
3	Definition of consensus	Unless not reasonable due to the explorative nature of the study, an a priori criterion for consensus should be defined. This includes a clear and transparent guide for action on (a) how to proceed with certain items or topics in the next survey round, (b) the required threshold to terminate the Delphi process and (c) procedures to be followed when consensus is (not) reached after one or more iterations	199-223
Study conduct			
4	Informational input	All material provided to the expert panel at the outset of the project and throughout the Delphi process should be carefully reviewed and piloted in advance in order to examine the effect on experts' judgements and to prevent bias	184-193
5	Prevention of bias	Researchers need to take measures to avoid directly or indirectly influencing the experts' judgements. If one or more members of the research team have a conflict of interest, entrusting an independent researcher with the main coordination of the Delphi study is advisable	317-318
6	Interpretation and processing of results	Consensus does not necessarily imply the 'correct' answer or judgement; (non)consensus and stable disagreement provide informative insights and highlight differences in perspectives concerning the topic in question	253-285
7	External validation	It is recommended to have the final draft of the resulting guidance on best practice be reviewed and approved by an external board or authority before publication and dissemination	NA
Reporting			
8	Purpose and rationale	The purpose of the study should be clearly defined and demonstrate the appropriateness of the use of the Delphi technique as a method to achieve the research aim. A rationale for the choice of the Delphi technique as the most suitable method needs to be provided	54-67
9	Expert panel	Criteria for the selection of experts and transparent information on recruitment of the expert panel, sociodemographic details including information on expertise regarding the topic in question, (non)response and response rates over the ongoing iterations should be reported	95-158

10	Description of methods	The methods employed need to be comprehensible; this includes information on preparatory steps (How was available evidence on the topic in question synthesised?), piloting of material and survey instruments, design of the survey instrument(s), the number and design of survey rounds, methods of data analysis, processing and synthesis of experts' responses to inform the subsequent survey round and methodological decisions taken by the research team throughout the process	46-53, 54-94, 159-198, 224-284
11	Procedure	Flow chart to illustrate the stages of the Delphi process, including a preparatory phase, the actual 'Delphi rounds', interim steps of data processing and analysis, and concluding steps	Figure 1
12	Definition/ attainment of consensus	It needs to be comprehensible to the reader how consensus was achieved throughout the process, including strategies to deal with non-consensus	NA for protocol
13	Results	Reporting of results for each round separately is highly advisable in order to make the evolving of consensus over the rounds transparent. This includes figures showing the average group response, changes between rounds, as well as any modifications of the survey instrument such as deletion, addition or modification of survey items based on previous rounds	NA for protocol
14	Discussion of limitations	Reporting should include a critical reflection of potential limitations and their impact of the resulting guidance	'Strengths and Limitations of this Study' bullet points
15	Adequacy of conclusions	The conclusions should adequately reflect the outcomes of the Delphi study with a view to the scope and applicability	NA for protocol
16	Publication and dissemination	The results should be clearly identifiable from the publication, including recommendations for transfer into practice and implementation. If the publication does not allow for a detailed presentation of either the output (e.g., CPG) or the methodological features of the applied Delphi technique, or both, reference to a more detailed presentation elsewhere should be made (e.g. availability of the full CPG from the authors or online; publication of a separate paper reporting on methodological details and particularities of the process (e.g. persistent disagreement and controversy on certain issues)). A dissemination plan should include endorsement of the guidance by professional associations and health care authorities to facilitate implementation	288-301

From: Jünger S, Payne SA, Brine J, Radbruch L, Brearley SG. Guidance on Conducting and REporting DElphi Studies (CREDES) in palliative care: Recommendations based on a methodological systematic review. Palliat Med [Internet]. 2017 Sep 1 [cited 2021 Aug 4];31(8):684–706. Available from: https://doi.org/10.1177/0269216317690685