BMJ Open

Existing guidance on reporting of consensus methodology: a systematic review to inform ACCORD guideline development

Esther J van Zuuren ,[1] Patricia Logullo,[2] Amy Price,[3,4] Zbys Fedorowicz,[5] Ellen L Hughes,[6] William T Gattrell[7]

To cite: van Zuuren EJ, Logullo P, Price A, *et al*. Existing guidance on reporting of consensus methodology: a systematic review to inform ACCORD guideline development. *BMJ Open* 2022;**12**:e065154. doi:10.1136/bmjopen-2022-065154

► 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-2022-065154).

Received 27 May 2022
Accepted 16 August 2022

For numbered affiliations see end of article.

Correspondence to
Dr Esther J van Zuuren; e.j.van_zuuren@lumc.nl

ABSTRACT

Objective To identify evidence on the reporting quality of consensus methodology and to select potential checklist items for the ACcurate COnsensus Reporting Document (ACCORD) project to develop a consensus reporting guideline.

Design Systematic review.

Data sources Embase, MEDLINE, Web of Science, PubMed, Cochrane Library, Emcare, Academic Search Premier and PsycINFO from inception until 7 January 2022.

Eligibility criteria Studies, reviews and published guidance addressing the reporting quality of consensus methodology for improvement of health outcomes in biomedicine or clinical practice. Reports of studies using or describing consensus methods but not commenting on their reporting quality were excluded. No language restrictions were applied.

Data extraction and synthesis Screening and data extraction of eligible studies were carried out independently by two authors. Reporting quality items addressed by the studies were synthesised narratively.

Results Eighteen studies were included: five systematic reviews, four narrative reviews, three research papers, three conference abstracts, two research guidance papers and one protocol. The majority of studies indicated that the quality of reporting of consensus methodology could be improved. Commonly addressed items were: consensus panel composition; definition of consensus and the threshold for achieving consensus. Items least addressed were: public patient involvement (PPI); the role of the steering committee, chair, cochair; conflict of interest of panellists and funding. Data extracted from included studies revealed additional items that were not captured in the data extraction form such as justification of deviation from the protocol or incentives to encourage panellist response.

Conclusion The results of this systematic review confirmed the need for a reporting checklist for consensus methodology and provided a range of potential checklist items to report. The next step in the ACCORD project builds on this systematic review and focuses on reaching consensus on these items to develop the reporting guideline.

Protocol registration https://osf.io/2rzm9.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ This systematic review used a comprehensive search of multiple databases without language restriction.
⇒ The included studies ranged from conference abstracts and protocols to guidelines and systematic reviews.
⇒ For full transparency and to promote discussion, all data retrieved are reported.
⇒ The data extraction form used may have missed a few potential reporting topics, but these will be recovered, in the following stages of the ACcurate COnsensus Reporting Document project, by additional reviews and the Delphi panel experience.
⇒ Conclusions are limited by the paucity of studies that provided substantial useful guidance.

INTRODUCTION

Healthcare providers face continuing challenges in making treatment decisions, particularly where available information on a clinical topic is limited, contradictory or non-existent. In such situations, alternative and complementary approaches underpinned by collective judgement and based on expert consensus may be used.[1–3]

A variety of approaches with differing methodological rigour can be used to achieve consensus-based decisions. These range from informal 'expert consensus meetings' to structured or systematic approaches such as the Delphi method and the Nominal Group Technique (NGT). These methods can be used for generating ideas or determining priorities and aim to achieve consensus through voting on a series of multiple-choice questions.[4–7] The voting process varies according to the method and may take place anonymously (as in Delphi) and/or face to face (in NGT and consensus conferences).[8–10] Key elements in the process include the use of valid and reliable methods to reach consensus and

subsequently their transparent reporting; however, these aspects are seldom clearly and explicitly reported.[3 11]

Reporting guidelines have been developed and are in use for the majority of study designs, for example, Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), Consolidated Standards of Reporting Trials (CONSORT) and Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) (for all existing reporting guidelines, see: https://www.equator-network.org/). However, no research reporting guideline exists for studies involving consensus methodology other than best practice guidance for Delphi studies in palliative care.[12] Guidelines should include 'a checklist, flow diagram or explicit text to guide authors in reporting a specific type of research, developed using explicit methodology'.[3]

Deficiencies in the reporting of consensus methods have been well documented in the literature and are referred to in the protocol for the ACcurate COnsensus Reporting Document (ACCORD) project, which aims to develop a reporting guideline for methods used to reach consensus.[13] In accordance with the EQUATOR Network guidance in the toolkit for the development of reporting guidelines, the next step for the ACCORD project was a review of the relevant literature, which would ultimately inform the voting process.[3]

Our objective was to undertake a thorough and comprehensive systematic review that seeks to identify evidence on the quality of reporting of consensus methodology, for subsequent development into a draft checklist of items for the ACCORD guideline. This ACCORD reporting guideline will assist the biomedical research and clinical practice community to describe the methods used to reach consensus in a complete, transparent and consistent manner.

METHODS

This manuscript conforms to the PRISMA statement[14] and follows a prespecified protocol.[13] The protocol was registered on 12 October 2021 at the Open Science Framework.[15]

Inclusion criteria

Eligible studies consisted of reviews and published guidance, which addressed the reporting quality of consensus methodology and aimed to improve health outcomes in biomedicine or clinical practice.

Exclusion criteria

Excluded were publications using consensus methods or describing consensus methods or discussing the advantages or disadvantages of frameworks, procedures or techniques to reach consensus, without specifically addressing reporting quality. Examples include guidelines developed through the use of consensus methodologies, such as reporting guidelines, clinical practice guidelines or core outcome set development studies. Editorials (usually brief opinion-based comments), letters about individual publications and commentaries on consensus methods outside the scope of biomedical research (eg, in the social sciences, economy, politics or marketing) were also excluded for this systematic review.

Literature search strategy and data sources

A systematic literature search was conducted on 7 January 2022 by a biomedical information specialist. The following bibliographical databases were searched: MEDLINE (OVID version), Embase (OVID version), PubMed, Web of Science, MEDLINE (Web of Science), Cochrane Library, Emcare (OVID version), PsycINFO (EbscoHOST version) and Academic Search Premier. The full search strategy is presented in online supplemental material 1.

We (EJvZ, ZF, PL and WTG) piloted four initial search strategies provided by the information specialist (JWS, see Acknowledgements section). The initial search strategy was sensitive and precise, producing the highest number of retrieved references (N=7951). After several rounds of checking through known relevant references and controlling for the effect of the performance of certain search terms, modifications were made, including the use of the most explicit terms in the most specific search fields. The performance of search terms was investigated from two vantage points: homonymy (same search term, but different meaning), and, particularly, loss-of-context (right meaning of the word, but not in the correct context). This extended search strategy not only provided extra 'signal' but also reduced the level of 'noise'. We chose to use specific rather than broad terms (eg, not using the singular terms 'delphi' and 'consensus' instead we included these words with relevant phrases or with other contextual words). In this way, the refined search strategy was better aligned with our inclusion criteria and the objectives of the systematic review.

Selection process

The final search results were uploaded to Rayyan (https://rayyan.ai) in the blind mode for independent screening by four review authors (EJvZ, ZF, PL and WTG) based on titles and abstracts. No language restrictions were applied. Records deemed eligible or without sufficient detail to make a clear judgement, we retrieved as full-text articles (EJvZ). The same four reviewers independently reassessed the eligibility of these full-text papers and any discrepancies were resolved through discussion. The references of the included studies were also checked for additional potentially eligible studies (EJvZ).

Data extraction, collection of items and synthesis

Study details and outcome data from the included studies were collected independently within Covidence (https://www.covidence.org/) by two authors using a piloted data extraction form (EJvZ and WTG). The data extraction form questions were compiled based on the review authors' own experiences with reporting quality evaluation and literature on consensus methodology.

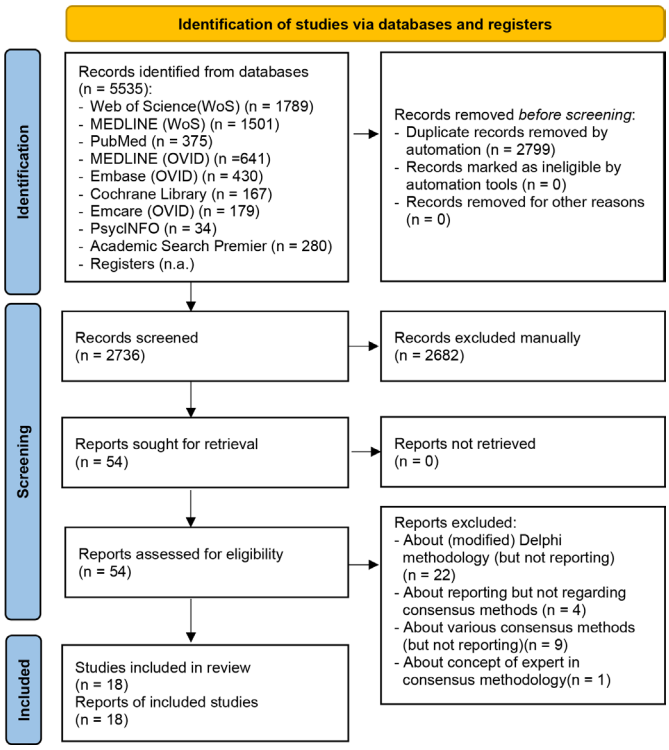

Figure 1 PRISMA 2020 flow diagram for new systematic reviews, including searches of databases, registers and other sources.[14]. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

Furthermore, two additional free text fields were created for extractors to present issues addressed by the included studies that were not captured by the other questions, and for others that the extractors felt were not directly addressed by the studies but were rather inferences about topics that could be potential issues in the reporting of consensus methods. Disagreements were discussed and reconciled by consultation with a third and fourth author (ZF and AP).

The following details were extracted: bibliographic details and reporting items including any suggestions and comments regarding reporting items. Reporting items were divided into the component parts of background, methods, results and discussion, each addressing key aspects of consensus methodology. We also included a section for additional items retrieved from the studies and not captured in the data extraction form. The complete data extraction form is found as online supplemental material 2.

The topics extracted and the methods used in the studies included are synthesised narratively, in text and tables and online supplemental material. No further analyses were carried out, but these will follow during the next stage of the ACCORD project as per protocol.[13]

Patient and public involvement

We involved patients, advocates and members of the lay public in the initial phases of this protocol,[13 15] as collaborators to develop this project and to coproduce the systematic review and coauthor the manuscript. They are collaborating with us by offering their experience with the use of consensus methods to develop guidelines and also systematic reviews. These contributors will work with us to disseminate the results.

RESULTS

Our searches across the databases identified 2599 articles and 137 further references to abstracts totalling 2736 references (after removal of duplicates) (see figure 1). A total of 2682 records were excluded after examination of titles and abstracts. Full-text copies of 54 studies were obtained for further assessment of eligibility, and finally, just 18 eligible studies were included. Checking of the references of these full-text publications did not yield any additional eligible articles.

Characteristics of included studies

Eighteen studies matched our prespecified eligibility criteria and were finally included in this review. These studies comprised five systematic reviews,[12 16–19] four reviews,[20–23] three research papers,[24–26] two research guidelines/guidance,[27 28] three conference abstracts[29–31] and one protocol.[32] Of the 18 included studies, 4 used Delphi plus other consensus methods[19 21 23 28] and the remaining 14 were primarily focused on only the Delphi method.[12 16–20 22 24–27 29 30]

Characteristics of excluded studies

A total of 36 studies were excluded.[7 8 33–66] The main reasons for their exclusion were: that they discussed (modified) Delphi methodology but did not include aspects of reporting[33–54]; that they covered reporting but not on consensus methodology[55–58]; that various other consensus methodologies were discussed but not their reporting[7 8 59–65] and that only the concept of experts in consensus methodology was discussed.[66]

Data extraction and narrative synthesis

The majority of studies indicated that reporting of consensus methods could be improved overall. The authors of these studies summarised some current limitations in reporting or proposed suggestions for improvement. Often there were common generic comments that noted reporting of consensus methodologies is inconsistent or lacks transparency. The studies provided few examples of areas that could be reported in more detail, such as: selection criteria for the participants and information about the participants; background information for panellists; definition of consensus; response rates after each round; description of level of anonymity or how anonymity was maintained and feedback between rounds (see table 1).

The studies we reviewed did not provide a systematic or standardised evaluation of the quality of reporting, but they did evaluate the literature critically and offered insights into the gaps of information about consensus. Fifteen papers made recommendations sometimes in the

Table 1 Data on reporting quality of consensus methodologies

Items that are not or not adequately reported in sufficient detail	
Selection criteria for participants/information about the participants[16 19 23 26 32]	Statement that anonymity was maintained or level of anonymity[20 21 25 28 29 32]
Literature review[20 21 31]	Type of consensus method used[29]
Background information for participants[20 21 25 28]	Threshold of consensus[29]
Recruitment strategies[19 22]	How questionnaire was developed[26]
Criteria for number of rounds[16 26]	Pretesting of instruments[19 32]
Stopping criteria[16 32]	Analysis procedure[24 32]
Feedback after rounds[17 20 21 23 25 26 28 31 32]	Changes to registered pre-analysis plan[24]
Rating scales used[31]	Reporting final number of list of items[32]
Criteria for dropping items[26]	Conflict of interest of panellists[29]
Response rates for each round[17 20 21 25 26 28 32]	Funding source[29]
Definition of consensus[17–19 21 23 25 26 28]	External support[29]
Level of consensus reached[19 31]	Generic comments that reporting needs improvement[12 17 26 30]

form of short lists—based solely on the authors' opinion, rather than using a systematic approach to reporting guidance development.[12 16–25 27 28 30 32] Detailed statements regarding quality of reporting are reproduced in online supplemental material 3 .

In table 2, we summarise the results of the data extraction, which correlates the corresponding aspects of consensus reporting ('items') to the studies that address them. The items in the table are presented in the format used in the data extraction form (see online supplemental material 2).

The most frequently addressed item in the included studies (16 times) was the composition of and the criteria for selecting the panellists, including their demographics; specifically, age, gender, specialty, years of experience and sociodemographic background. The aspects of clarity in, and the importance of, defining consensus and the corresponding thresholds to reach that consensus were addressed in 13 studies. The prespecified number of voting rounds and provision of feedback to the panellists at the end of each round were addressed in 10 and 11 of the studies, respectively.

None of the included studies reported or made reference to public patient involvement (PPI). The roles of the steering committee/chair/cochair were not defined in any of the included studies. Reporting of the time interval between voting rounds, panel members' conflicts of interest (COI) and funding sources, as well as the measures used to avoid the influence of COI on voting and decision-making, were minimally addressed.

Conversely, three studies addressed between 12 and 19 reporting items of the 30 items present in the data extraction form of this review,[12 19 28] whereas two studies covered only two or three items.[19 24] We identified a considerable number of other aspects of reporting that were proposed in the included studies, but which were not captured in our data extraction form. These included: 'justifications for deviating from the protocol', 'incentives for encouraging panellists to respond' and 'suggestions to add a flowchart of the consensus process'. All extracted data are found in online supplemental materials 4,5.

DISCUSSION

Although consensus methodology is widely used in healthcare and researchers do raise poor reporting as an issue, we were able to identify only 18 studies that commented on reporting quality and/or provided suggestions to improve the quality of reporting of consensus methodology. These included studies ranged from conference abstracts and protocols to guidelines and systematic reviews. Only four studies covered methods other than the Delphi method and, thus, providing very limited guidance on other consensus methodologies. We carried out a comprehensive search of the most commonly used databases for systematic reviews without language restriction. However, during peer review of the present manuscript, three studies were brought to our attention as potentially eligible for inclusion.[67–69] Two of the studies had been excluded at the screening stage.[67 68] After full-text evaluation, one of the articles did discuss reporting quality but failed to make that clear in the title or abstract[67]; however, the findings were consistent with our reported results. The second publication did not meet our eligibility criteria because it focused on studies of health economics rather than health outcomes.[68] Interestingly, the study identified similar gaps to the present study, but its scope is outside our protocol and research question. The third was not picked up during screening because the journal is not indexed in the nine predefined data sources for the searches.[69]

The data extraction form may have missed a few potential reporting topics—which will be recovered, in the next stages of the ACCORD project, by additional reviews and the Delphi panel experience. Furthermore, one study was published after our search date, showing that

Table 2 Studies providing guidance for reporting items in the extraction form of this systematic review

Reporting items	Studies that provide guidance	
Background	Number	References
1.1 Rationale for choosing a consensus method over other methods	4	12 25 27 28
1.2 Clearly defined objective	6	12 17 18 20 27 28
Methods		
2.1 Review of existing evidence informing consensus study	5	20 21 27 28 31
2.2 Inclusion and exclusion criteria of the literature search	3	17 20 22
2.3 Composition of the panel	16	12 16–23 25–30 32
2.4 Public patient involvement (PPI)	0	
2.5 Panel recruitment	4	12 17 22 23
2.6 Defining consensus and the threshold for achieving consensus	13	12 17–21 23–29
2.7 Decision of item approval	3	12 17 27
2.8 Number of voting rounds	10	12 16 18 20 21 23 26–28 32
2.9 Rationale for number of voting rounds	8	16 20–23 25 26 28
2.10 Time between voting rounds	1	17
2.11 Additional methods used alongside consensus	2	17 23
2.12 Software or tools used for voting	1	25
2.13 Anonymity of panellists and how this was maintained	7	16 20–22 25 28 29
2.14 Feedback to panellists at the end of each round	11	17 19–22 25–29 31
2.15 Synthesis/analysis of responses after voting rounds	5	12 22–24 30
2.16 Pilot testing of study material/instruments	3	12 22 28
2.17 Role of the steering committee/chair/co-chair/facilitator	0	
2.18 Conflict of interest or funding received	4	12 29 30 32
2.19 Measures to avoid influence by conflict of interest	1	12
Results		
3.1 Results of the literature search	1	12
3.2 Number of studies found as supporting evidence	0	
3.3 Response rates per voting round	5	12 21 22 25 30
3.4 Results shared with respondents	9	12 17 20 25–28 30 31
3.5 Dropped items	5	12 16 18 26 32
3.6 Collection, synthesis and comments from panellists	5	12 17 22 28 31
3.7 Final list of items (eg, for guideline or reporting guideline)	4	12 22 30 31
Discussion		
4.1 Limitations and strengths of the study	5	12 20 25 27 28
4.2 Applicability, generalisability, reproducibility	3	12 17 26

the development of reporting guidelines for consensus methodologies is an active area, with more studies being published on the topic continuously, which could inform future stages or updates of ACCORD.[70] Comments regarding deficient reporting from the included studies varied from generic statements such as 'reporting could be improved' to rather specific comments of which aspects of consensus methods were inadequately or not reported. Far more detailed data were provided regarding guidance to improve reporting quality or suggestions for items that require reporting. Both composition and characteristics of the panel, and defining consensus and threshold for achieving assessment received, were consistently addressed and appeared to be critical items that should be reported in sufficient detail. Feedback to the panel might be considered an important aspect of ensuring ongoing engagement with the panellists, transparency and replicability of methods; thus, it was somewhat surprising to see just 11 of the 18 studies consider this an element of consensus methodology worth reporting.

Some items were not addressed in any of the studies, specifically PPI, which is currently considered a key element in the shared decision-making process and is a component of guideline development.[71] Just four

studies made reference to the COI of panel members and project funding. COI of panellists, as well as of chair, cochair and steering committee, can directly or indirectly impact and influence decision-making during the various steps of consensus methodology. As such, COI remains under-reported and is often inconsistently described.[72] This also raises concerns about the measures that can be taken to mitigate the potential influence of COI and to ensure that those panellists who do have relevant interests are, for example, not able to vote on pertinent items. For full transparency and to promote discussion, all data retrieved are reported as supplementary material (online supplemental materials 3–5).

Although conclusions are limited by the paucity of studies, a few were particularly informative. The first was a systematic review on the use and reporting of the Delphi method for selecting healthcare indicators.[17] Specifically, this review not only provided guidance for planning and using the Delphi procedure but also additionally formulated general recommendations for reporting. The second study was a guidance report on consensus methods such as Delphi and NGT, which were used in medical education research.[28] The authors reported that there is a lack of 'standardisation in definitions, methodology and reporting' and proposed items for researchers to consider when using consensus methods to improve methodological rigour as well as the reporting quality. However, it is worth noting that none of these studies followed the EQUATOR Network guidance for the development of a reporting guideline.[3]

The third study we would like to highlight is the Guidance on Conducting and REporting DElphi Studies (CREDES) in palliative care, which was based on a methodological systematic review.[12] This study focused on the development of guidance in palliative care, although it may not be suitable for extrapolation to other biomedical areas. Furthermore, this study only considered the Delphi methodology, whereas we included studies covering consensus processes involving non-Delphi-based methods or 'modified Delphi' in our review (and in the ACCORD project overall). However, many of the suggestions made regarding the design and conduct of Delphi studies in addition to recommendations for reporting are equally applicable to our ACCORD project. These items will be used and integrated into the next step of the project, which is the development of a reporting checklist on consensus methods.

Two additional studies proved to be of particular value.[21 25] One provided a preliminary Delphi checklist to be used for Outcome Measures in Rheumatology.[25] The other concluded, in a scoping review that consensus methods are 'poorly standardised and inconsistently used' and exposed reporting flaws in consensus reports.[21]

CONCLUSION

The principal objectives of this systematic review were to conduct a comprehensive search and to identify the existing evidence on the quality of reporting of consensus methodology. As such, we have been able to gather together all relevant studies, summarise the existing research and highlight key gaps in the current evidence base on consensus methods. This systematic review will ultimately inform the generation of a draft checklist of items for the development steps of the ACCORD reporting guideline.

Author affiliations
[1]Department of Dermatology, Leiden University Medical Center, Leiden, Zuid-Holland, Netherlands
[2]Nuffield Department of Orthopaedics, Rheumatology and Muskuloskeletal Sciences, Centre for Statistics in Medicine, University of Oxford and EQUATOR Network UK Centre, Oxford, UK
[3]Stanford Anesthesia, Informatics and Media Lab, Stanford University School of Medicine, Stanford, California, USA
[4]The BMJ, London, UK
[5]Veritas Health Sciences Consultancy, Huntingdon, UK
[6]Sciwright Limited, Somerset, UK
[7]Global Medical Affairs, Ipsen, Abingdon, Oxfordshire, UK

Twitter Esther J van Zuuren @ezzoef and Patricia Logullo @patlogullo

Acknowledgements We would like to thank Jan W. Schoones of the Directorate of Research Policy (formerly: Walaeus Library), Leiden University Medical Centre, Leiden, The Netherlands, for his assistance in the development of the search strategy. Furthermore, we would like to thank the other members of the ACCORD project Steering Committee for their contributions to the protocol: Amrit Pali Hungin, Christopher C Winchester, David Tovey, Keith Goldman, Rob Matheis and Niall Harrison. We are grateful for the editorial guidance and reviewer insights from the BMJ Open team. We agree their observations improved the manuscript.

Contributors EJvZ, PL, ZF and WTG contributed to the screening and agreed on the inclusion of studies. EJvZ and WTG extracted data from the included studies. AP, ZF and ELH contributed to the discussion of extracted data and interpretation. EJvZ was the major contributor in the review of studies, data extraction, interpretation of findings as well as writing of the manuscript. All authors read the final manuscript, provided feedback and approved the final manuscript. EJvZ is the guarantor.

Funding PL contributed to this study using her time funded by CRUK. WTG contributed his time with the agreement of his employers. Apart from these authors, this study was conducted without external funding for the development of the study design; for the collection, analysis and interpretation of the data; or for the writing of the report. As ZF (Veritas Health Sciences Consultancy) acted as an independent consultant for this work, his support for data extraction and tabulation was funded by Oxford PharmaGenesis. Support for Open Access was also provided by Oxford PharmaGenesis. The decision to submit the paper for publication was made solely by the authors.

Competing interests PL is a member of the UK EQUATOR Centre, an organisation that promotes the use of reporting guidelines, many of which are developed using consensus methods, and she is personally involved in the development of other reporting guidelines. ELH has worked with Ogilvy Health UK on consensus projects. WTG is a former employee of Ipsen and is now employed by Bristol Myers Squib. AP is an editor at the BMJ and is a senior research scientist at Stanford University, where she is responsible for advising and seeking funding on Delphi and other research studies. EJvZ and ZF have no conflict of interest.

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.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information. All key data for this study are included in this article or uploaded as online supplementary information. The ACCORD protocol has been listed on the EQUATOR website (Reporting guidelines under development for other study designs | The EQUATOR Network (equator-

network.org)) and registered with the Open Science Framework (https://osf.io/2rzm9).

ORCID iD

Esther J van Zuuren http://orcid.org/0000-0002-4780-0182

REFERENCES

1 Raine R, Sanderson C, Black N. Developing clinical guidelines: a challenge to current methods. *BMJ* 2005;331:631–3.
2 Djulbegovic B, Guyatt G. Evidence vs consensus in clinical practice guidelines. *JAMA* 2019;322:725–6.
3 Moher D, Schulz KF, Simera I, *et al*. Guidance for developers of health research reporting guidelines. *PLoS Med* 2010;7:e1000217.
4 Dalkey N, Helmer O. An experimental application of the Delphi method to the use of experts. *Manage Sci* 1963;9:458–67.
5 Murphy MK, Black NA, Lamping DL, *et al*. Consensus development methods, and their use in clinical Guideline development. *Health Technol Assess* 1998;2:1–88.
6 Foth T, Efstathiou N, Vanderspank-Wright B, *et al*. The use of Delphi and nominal group technique in nursing education: a review. *Int J Nurs Stud* 2016;60:112–20.
7 Nair R, Aggarwal R, Khanna D. Methods of formal consensus in classification/diagnostic criteria and Guideline development. *Semin Arthritis Rheum* 2011;41:95–105.
8 McMillan SS, King M, Tully MP. How to use the nominal group and Delphi techniques. *Int J Clin Pharm* 2016;38:655–62.
9 Grant S, Armstrong C, Khodyakov D. Online Modified-Delphi: a potential method for continuous patient engagement across stages of clinical practice Guideline development. *J Gen Intern Med* 2021;36:1746–50.
10 Helmer-Hirschberg O. *Analysis of the future: the Delphi method*. Santa Monica, CA: RAND Corporation, 1967. https://www.rand.org/pubs/papers/P3558.htm
11 Arakawa N, Bader LR. Consensus development methods: considerations for national and global frameworks and policy development. *Res Social Adm Pharm* 2022;18:2222–9.
12 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.
13 Gattrell WT, Hungin AP, Price A, *et al*. ACCORD guideline for reporting consensus-based methods in biomedical research and clinical practice: a study protocol. *Res Integr Peer Rev* 2022;7:3.
14 Page MJ, McKenzie JE, Bossuyt PM, *et al*. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71.
15 Gattrell W, Hungin AP, BWMA A. Consensus-Based methods in biomedical research and clinical practice: the ACCORD study protocol for establishing a reporting guideline. Available: https://osf.io/e5w9n/ [Accessed 19 April 2022].
16 Banno M, Tsujimoto Y, Kataoka Y. The majority of reporting guidelines are not developed with the Delphi method: a systematic review of reporting guidelines. *J Clin Epidemiol* 2020;124:50–7.
17 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.
18 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.
19 Wang X, Chen Y, Yang N, *et al*. Methodology and reporting quality of reporting guidelines: systematic review. *BMC Med Res Methodol* 2015;15:74.
20 Chan TM, Yarris LM, Humphrey-Murto S. Delving into delphis. *CJEM* 2019;21:167–9.
21 Humphrey-Murto S, Varpio L, Wood TJ, *et al*. The use of the Delphi and other consensus group methods in medical education research: a review. *Acad Med* 2017;92:1491–8.
22 Paré G, Cameron A-F, Poba-Nzaou P, *et al*. A systematic assessment of rigor in information systems ranking-type Delphi studies. *Inf Manage* 2013;50:207–17.
23 Waggoner J, Carline JD, Durning SJ. Is there a consensus on consensus methodology? descriptions and recommendations for future consensus research. *Acad Med* 2016;91:663–8.
24 Grant S, Booth M, Khodyakov D. Lack of preregistered analysis plans allows unacceptable data mining for and selective reporting of consensus in Delphi studies. *J Clin Epidemiol* 2018;99:96–105.
25 Humphrey-Murto S, Crew R, Shea B, *et al*. Consensus building in OMERACT: recommendations for use of the Delphi for core outcome set development. *J Rheumatol* 2019;46:1041–6.
26 Niederberger M, Spranger J. Delphi technique in health sciences: a MAP. *Front Public Health* 2020;8:457.
27 Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. *J Adv Nurs* 2000;32:1008–15.
28 Humphrey-Murto S, Varpio L, Gonsalves C, *et al*. [Using]. *Med Teach* 2017;39:14–19.
29 Gattrell WT, Clements SJ, Sheard D. Quality assessment of guidelines/recommendations developed using Delphi methodology. *Curr Med Res Opin* 2019;35:40.
30 Ng J. Delphi method: a qualitative approach for quantitative results. *Value Health* 2018;21:S54 https://www.valueinhealthjournal.com/article/S1098-3015(18)30747-2/pdf.
31 Resemann HK, Clements S, Griffiths A. Reporting of Delphi methods to achieve consensus on guidelines in rare diseases. *Curr Med Res Opin* 2018;34:37. doi:1080/03007995.2018.1440990
32 Banno M, Tsujimoto Y, Kataoka Y. Reporting quality of the Delphi technique in reporting guidelines: a protocol for a systematic analysis of the EQUATOR network library. *BMJ Open* 2019;9:e024942.
33 Bellali T, Karamitri I. The Delphi research methodology and its applications in the healthcare sciences. *Arch Hellen Med* 2011;28:48 http://www.mednet.gr/archives/2011-6/pdf/839.pdf
34 Bijl R. Delphi in a future scenario study on mental health and mental health care. *Futures* 1992;24:232–50.
35 Boel A, Navarro-Compán V, Landewé R, *et al*. Two different invitation approaches for consecutive rounds of a Delphi survey led to comparable final outcome. *J Clin Epidemiol* 2021;129:31–9.
36 Bowles N. The Delphi technique. *Nurs Stand* 1999;13:32–6.
37 Cramer CK, Klasser GD, Epstein JB, *et al*. The Delphi process in dental research. *J Evid Based Dent Pract* 2008;8:211–20.
38 de Loë RC, Melnychuk N, Murray D, *et al*. Advancing the state of policy Delphi practice: a systematic review evaluating methodological evolution, innovation, and opportunities. *Technol Forecast Soc Change* 2016;104:78–88.
39 Drumm S, Bradley C, Moriarty F. 'More of an art than a science'? the development, design and mechanics of the Delphi technique. *Res Social Adm Pharm* 2022;18:2230–6.
40 Fink-Hafner D, Dagan T, Doušak M. Delphi method: strengths and weaknesses. *Metodoloski Zv* 2019;16:1–19.
41 Franklin KK, Hart JK, Generation I. Idea generation and exploration: benefits and limitations of the policy Delphi research method. *Innov High Educ* 2006;31:237–46.
42 Guzys D, Dickson-Swift V, Kenny A, *et al*. Gadamerian philosophical hermeneutics as a useful methodological framework for the Delphi technique. *Int J Qual Stud Health Well-being* 2015;10:26291.
43 Hallowell MR, Gambatese JA. Qualitative research: application of the Delphi method to CEM research. *J Constr Eng Manag* 2010;136:99–107.
44 Holloway K. Doing the E-Delphi: using online survey tools. *Comput Inform Nurs* 2012;30:347–50.
45 Hung H-L, Altschuld JW, Lee Y-F. Methodological and conceptual issues confronting a cross-country Delphi study of educational program evaluation. *Eval Program Plann* 2008;31:191–8.
46 Ibiyemi AO, Adnan YM, Daud MN. The validity of the classical Delphi applications for assessing the industrial sustainability-correction factor: an example study. *Foresight* 2016;18:603–24.
47 Keeney S, Hasson F, McKenna H. Consulting the oracle: ten lessons from using the Delphi technique in nursing research. *J Adv Nurs* 2006;53:205–12.
48 Keeney S, Hasson F, McKenna HP. A critical review of the Delphi technique as a research methodology for nursing. *Int J Nurs Stud* 2001;38:195–200.

49 Kennedy HP. Enhancing Delphi research: methods and results. *J Adv Nurs* 2004;45:504–11.
50 Khodyakov D, Grant S, Denger B, *et al*. Practical considerations in using online Modified-Delphi approaches to engage patients and other stakeholders in clinical practice Guideline development. *Patient* 2020;13:11–21.
51 Murry JW, Hammons JO. Delphi: a versatile methodology for conducting qualitative research. *Rev High Ed* 1995;18:423–36.
52 Olsen AA, Wolcott MD, Haines ST, *et al*. How to use the Delphi method to aid in decision making and build consensus in pharmacy education. *Curr Pharm Teach Learn* 2021;13:1376–85.
53 Vernon W. The Delphi technique: a review. *Int J Ther Rehabil* 2009;16:69–76.
54 Von der Gracht HA. Consensus measurement in Delphi studies review and implications for future quality assurance. *Technol Forecast Soc Change* 2012;79:1525–36.
55 Bennett C, Khangura S, Brehaut JC, *et al*. Reporting guidelines for survey research: an analysis of published guidance and reporting practices. *PLoS Med* 2010;8:e1001069.
56 Han SJ, Oh MC, Sughrue ME, *et al*. Reporting standard compliance in publications of vestibular schwannoma patients treated with microsurgery. *Otol Neurotol* 2012;33:648–50.
57 Moher D, Weeks L, Ocampo M, *et al*. Describing reporting guidelines for health research: a systematic review. *J Clin Epidemiol* 2011;64:718–42.
58 Murray A, Junge A, Robinson PG, *et al*. International consensus statement: methods for recording and reporting of epidemiological data on injuries and illnesses in golf. *Br J Sports Med* 2020;54:1136–41.
59 Campbell SM, Cantrill JA. Consensus methods in prescribing research. *J Clin Pharm Ther* 2001;26:5–14.
60 Durkin J, Usher K, Jackson D. Using consensus from experts to inform a shared understanding of subjective terms. *Nurse Res* 2019;27:46–9.
61 Fink A, Kosecoff J, Chassin M, *et al*. Consensus methods: characteristics and guidelines for use. *Am J Public Health* 1984;74:979–83.
62 Halcomb E, Davidson P, Hardaker L. Using the consensus development conference method in healthcare research. *Nurse Res* 2008;16:56–71.
63 James D, Warren-Forward H. Research methods for formal consensus development. *Nurse Res* 2015;22:35–40.
64 Rotondi AJ, Kvetan V, Carlet J, *et al*. Consensus conferences in critical care medicine. methodologies and impact. *Crit Care Clin* 1997;13:417–39.
65 van Teijlingen E, Pitchforth E, Bishop C, *et al*. Delphi method and nominal group technique in family planning and reproductive health research. *J Fam Plann Reprod Health Care* 2006;32:249–52.
66 Baker J, Lovell K, Harris N. How expert are the experts? An exploration of the concept of 'expert' within Delphi panel techniques. *Nurse Res* 2006;14:59–70.
67 Sinha IP, Smyth RL, Williamson PR. Using the Delphi technique to determine which outcomes to measure in clinical trials: recommendations for the future based on a systematic review of existing studies. *PLoS Med* 2011;8:e1000393.
68 Iglesias CP, Thompson A, Rogowski WH, *et al*. Reporting guidelines for the use of expert judgement in model-based economic evaluations. *Pharmacoeconomics* 2016;34:1161–72.
69 Toma C, Picioreanu I. The Delphi technique: methodological considerations and the need for reporting guidelines in medical journals. *Int J Public Health Res* 2016;4 http://www.openscienceonline.com/journal/archive2?journalId=718&paperId=3586
70 Spranger J, Homberg A, Sonnberger M. Reporting guidelines for Delphi techniques in health sciences: a methodological review. Berichterstattungsleitlinien für Delphi-Verfahren in den Gesundheitswissenschaften: ein methodologisches review. *Z Evid Fortbild Qual Gesundh wesen (ZEFQ) Available online since* June 16th 2022.
71 National Institute for Health and Care Excellence. NICE Guideline [NG197]: Shared decision making. Available: https://www.nice.org.uk/guidance/ng197
72 Dunn AG, Coiera E, Mandl KD, *et al*. Conflict of interest disclosure in biomedical research: a review of current practices, biases, and the role of public registries in improving transparency. *Res Integr Peer Rev* 2016;1:1:1.

ACCORD - January 7th, 2022

Regular references:

Total: 2599 references, sourced from:

- Web of Science - core collection: 1775
- MEDLINE (Web of Science): 1501 - 202 unique
- PubMed: 375 - 219 unique
- MEDLINE (OVID): 641 - 174 unique
- Embase (OVID): 331 - 66 unique
- Cochrane Library: 131 - 77 unique
- Emcare (OVID): 179 - 29 unique
- Academic Search Premier: 280 - 23 unique
- PsycINFO: 173 - 34 unique

Meeting abstract references:

Total: 137 references, sourced from:

- Web of Science: 14
- Embase (OVID): 99 - 90 unique
- Cochrane Library: 36 - 33 unique

Known references:

- PubMed: 27841062 26796090 25587865 26395179 24581294
- MEDLINE (Web of Science): PMID=(27841062 OR 26796090 OR 25587865 OR 26395179 OR 24581294)
- Web of Science Core Collection: UT=(000393885800003 OR 000375153500022 OR 000376181900007 OR 000361506400001 OR 000334256400007 OR 000309802600012 OR 000321232400002 OR 000309802600012 OR 000465105500070)

Databases:

Web of Science Core Collection and MEDLINE (Web of Science)
http://isiknowledge.com/wos

((TI=("Delphi Technique" OR "Delphi Technique" OR "Delphi techniques" OR "Delphi method" OR "Delphi methods" OR "Delphi study" OR "Delphi studies" OR "Delphi survey" OR "Delphi surveys" OR "Delphi consensus" OR "Delphi based consensus" OR "Delphi questionnaire" OR "Delphi questionnaires" OR "Delphi research" OR "Delphi review" OR "Delphi reviews" OR "Delphi process" OR "Delphi processes" OR "Delphi based" OR "Delphi procedure" OR "Delphi procedures" OR "Delphi assessment" OR "Delphi assessments" OR "Delphi approach" OR "Delphi approaches" OR "Delphi panel" OR "Delphi panels" OR "Delphi round" OR "Delphi rounds" OR "Delphi analysis" OR

"Delphi expert" OR "Delphi experts" OR "Delphi consultation" OR "Delphi methodology" OR "nominal group technique" OR "nominal group techniques" OR "nominal group" OR "nominal groups" OR "nominal grouping" OR "consensus recommendation" OR "consensus recommendations" OR "consensus development" OR "consensus activity" OR "consensus activities" OR "Consensus Development Conference" OR "Consensus Development" OR "Consensus methodology" OR "consensus method*" OR "RAND" OR (("Guidelines" OR "guideline") NEAR/2 ("consensus" OR "delphi"))) OR AB=("Delphi Technique" OR "Delphi Technique" OR "Delphi techniques" OR "Delphi method" OR "Delphi methods" OR "Delphi study" OR "Delphi studies" OR "Delphi survey" OR "Delphi surveys" OR "Delphi consensus" OR "Delphi based consensus" OR "Delphi questionnaire" OR "Delphi questionnaires" OR "Delphi research" OR "Delphi review" OR "Delphi reviews" OR "Delphi process" OR "Delphi processes" OR "Delphi based" OR "Delphi procedure" OR "Delphi procedures" OR "Delphi assessment" OR "Delphi assessments" OR "Delphi approach" OR "Delphi approaches" OR "Delphi panel" OR "Delphi panels" OR "Delphi round" OR "Delphi rounds" OR "Delphi analysis" OR "Delphi expert" OR "Delphi experts" OR "Delphi consultation" OR "Delphi methodology" OR "nominal group technique" OR "nominal group techniques" OR "nominal group" OR "nominal groups" OR "nominal grouping" OR "consensus recommendation" OR "consensus recommendations" OR "consensus development" OR "consensus activity" OR "consensus activities" OR "Consensus Development Conference" OR "Consensus Development" OR "Consensus methodology" OR "consensus method*" OR "RAND") OR AK=("Delphi Technique" OR "Delphi Technique" OR "Delphi techniques" OR "Delphi method" OR "Delphi methods" OR "Delphi study" OR "Delphi studies" OR "Delphi survey" OR "Delphi surveys" OR "Delphi consensus" OR "Delphi based consensus" OR "Delphi questionnaire" OR "Delphi questionnaires" OR "Delphi research" OR "Delphi review" OR "Delphi reviews" OR "Delphi process" OR "Delphi processes" OR "Delphi based" OR "Delphi procedure" OR "Delphi procedures" OR "Delphi assessment" OR "Delphi assessments" OR "Delphi approach" OR "Delphi approaches" OR "Delphi panel" OR "Delphi panels" OR "Delphi round" OR "Delphi rounds" OR "Delphi analysis" OR "Delphi expert" OR "Delphi experts" OR "Delphi consultation" OR "Delphi methodology" OR "nominal group technique" OR "nominal group techniques" OR "nominal group" OR "nominal groups" OR "nominal grouping" OR "consensus recommendation" OR "consensus recommendations" OR "consensus development" OR "consensus activity" OR "consensus activities" OR "Consensus Development Conference" OR "Consensus Development" OR "Consensus methodology" OR "consensus method*" OR "RAND" OR (("Guidelines" OR "guideline") NEAR/2 ("consensus" OR "delphi")))) AND (TI=("quality of reporting" OR "reporting quality" OR "reporting qualities" OR "selective reporting" OR "poor reporting" OR "poor reported" OR "reporting guideline" OR "reporting" OR ("reporting" AND ("quality" OR "selective" OR "poor" OR "weak" OR "manner" OR "rigor" OR "improv*")) OR "Data Accuracy" OR "quality assessment" OR "strengths" OR "strength" OR "weaknesses" OR "weakness" OR "research method" OR "research methods" OR "research method*") OR AK=("quality of reporting" OR "reporting quality" OR "reporting qualities" OR "selective reporting" OR "poor reporting" OR "poor reported" OR ("reporting" NEAR/5 ("quality" OR "selective" OR "poor" OR "weak" OR "manner" OR "rigor" OR "improv*")) OR "Data Accuracy" OR

"Research Report standards" OR "quality assessment" OR "strengths" OR "strength" OR "weaknesses" OR "weakness" OR "research method" OR "research methods" OR "research method*") OR AB=("quality of reporting" OR "reporting quality" OR "reporting qualities" OR "selective reporting" OR "poor reporting" OR "poor reported" OR ("reporting" NEAR/5 ("quality" OR "selective" OR "poor" OR "weak" OR "manner" OR "rigor" OR "improv*")) OR "Data Accuracy" OR "Research Report standards" OR "quality assessment" OR "strengths" OR "strength" OR "weaknesses" OR "weakness")))

PubMed
http://www.ncbi.nlm.nih.gov/pubmed?otool=leiden

(("Delphi Technique"[majr] OR "Delphi Technique"[ti] OR "Delphi techniques"[ti] OR "Delphi method"[ti] OR "Delphi methods"[ti] OR "Delphi study"[ti] OR "Delphi studies"[ti] OR "Delphi survey"[ti] OR "Delphi surveys"[ti] OR "Delphi consensus"[ti] OR "Delphi based consensus"[ti] OR "Delphi questionnaire"[ti] OR "Delphi questionnaires"[ti] OR "Delphi research"[ti] OR "Delphi review"[ti] OR "Delphi reviews"[ti] OR "Delphi process"[ti] OR "Delphi processes"[ti] OR "Delphi based"[ti] OR "Delphi procedure"[ti] OR "Delphi procedures"[ti] OR "Delphi assessment"[ti] OR "Delphi assessments"[ti] OR "Delphi approach"[ti] OR "Delphi approaches"[ti] OR "Delphi panel"[ti] OR "Delphi panels"[ti] OR "Delphi round"[ti] OR "Delphi rounds"[ti] OR "Delphi analysis"[ti] OR "Delphi expert"[ti] OR "Delphi experts"[ti] OR "Delphi consultation"[ti] OR "Delphi methodology"[ti] OR "nominal group technique"[ti] OR "nominal group techniques"[ti] OR "nominal group"[ti] OR "nominal groups"[ti] OR "nominal grouping"[ti] OR "consensus recommendation"[ti] OR "consensus recommendations"[ti] OR "consensus development"[ti] OR "consensus activity"[ti] OR "consensus activities"[ti] OR "consensus methodology"[ti] OR "consensus method*"[ti] OR "Consensus Development Conferences as Topic"[majr] OR "RAND"[ti] OR ("Guidelines as Topic"[majr:noexp] AND ("consensus"[tw] OR "delphi"[tw]))) AND ("reporting"[ti] OR "quality of reporting"[tw] OR "reporting quality"[tw] OR "reporting qualities"[tw] OR "selective reporting"[tw] OR "poor reporting"[tw] OR "poor reported"[tw] OR "poorly reported"[tw] OR "Research Report/standards"[majr] OR "Research Design/standards"[mesh] OR "Research Design"[majr:noexp] OR "Writing/standards"[mesh] OR "Writing"[majr] OR "research method"[ti] OR "research methods"[ti] OR "research method*"[ti]))

MEDLINE via OVID
http://gateway.ovid.com/ovidweb.cgi?T=JS&MODE=ovid&NEWS=n&PAGE=main&D=medall

((exp *"Delphi Technique"/ OR "Delphi Technique".ti OR "Delphi techniques".ti OR "Delphi method".ti OR "Delphi methods".ti OR "Delphi study".ti OR "Delphi studies".ti OR "Delphi survey".ti OR "Delphi surveys".ti OR "Delphi consensus".ti OR "Delphi based consensus".ti OR "Delphi questionnaire".ti OR "Delphi questionnaires".ti OR "Delphi research".ti OR "Delphi review".ti OR "Delphi reviews".ti OR "Delphi

process".ti OR "Delphi processes".ti OR "Delphi based".ti OR "Delphi procedure".ti OR "Delphi procedures".ti OR "Delphi assessment".ti OR "Delphi assessments".ti OR "Delphi approach".ti OR "Delphi approaches".ti OR "Delphi panel".ti OR "Delphi panels".ti OR "Delphi round".ti OR "Delphi rounds".ti OR "Delphi analysis".ti OR "Delphi expert".ti OR "Delphi experts".ti OR "Delphi consultation".ti OR "Delphi methodology".ti OR "nominal group technique".ti OR "nominal group techniques".ti OR "nominal group".ti OR "nominal groups".ti OR "nominal grouping".ti OR "consensus recommendation".ti OR "consensus recommendations".ti OR "consensus development".ti OR "consensus activity".ti OR "consensus activities".ti OR "consensus methodology".ti OR "consensus method*".ti OR exp *"Consensus Development Conferences as Topic"/ OR "RAND".ti OR ("Guidelines as Topic"/ AND ("consensus".mp OR "delphi".mp)) OR (("Guidelines".mp OR "guideline".mp) ADJ2 ("consensus".mp OR "delphi".mp))) AND ("reporting".ti OR "quality of reporting".mp OR "reporting quality".mp OR "reporting qualities".mp OR "selective reporting".mp OR "poor reporting".mp OR "poor reported".mp OR "poorly reported".mp OR "Research Report/standards"/ OR exp "Research Design"/st OR *"Research Design"/ OR exp "Writing"/st OR exp *"Writing"/ OR "research method".ti OR "research methods".ti OR "research method*".ti OR ("reporting" ADJ8 ("quality" OR "selective" OR "poor" OR "weak" OR "manner" OR "rigor" OR "improv*")).mp))

Embase
http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=main&MODE=ovid&D=oemezd

((exp *"Delphi Study"/ OR "Delphi Technique".ti OR "Delphi techniques".ti OR "Delphi method".ti OR "Delphi methods".ti OR "Delphi study".ti OR "Delphi studies".ti OR "Delphi survey".ti OR "Delphi surveys".ti OR "Delphi consensus".ti OR "Delphi based consensus".ti OR "Delphi questionnaire".ti OR "Delphi questionnaires".ti OR "Delphi research".ti OR "Delphi review".ti OR "Delphi reviews".ti OR "Delphi process".ti OR "Delphi processes".ti OR "Delphi based".ti OR "Delphi procedure".ti OR "Delphi procedures".ti OR "Delphi assessment".ti OR "Delphi assessments".ti OR "Delphi approach".ti OR "Delphi approaches".ti OR "Delphi panel".ti OR "Delphi panels".ti OR "Delphi round".ti OR "Delphi rounds".ti OR "Delphi analysis".ti OR "Delphi expert".ti OR "Delphi experts".ti OR "Delphi consultation".ti OR "Delphi methodology".ti OR "nominal group technique".ti OR "nominal group techniques".ti OR "nominal group".ti OR "nominal groups".ti OR "nominal grouping".ti OR "consensus recommendation".ti OR "consensus recommendations".ti OR "consensus development".ti OR "consensus activity".ti OR "consensus activities".ti OR "consensus methodology".ti OR "consensus method*".ti OR exp *"Consensus Development"/ OR "RAND".ti OR (("Guidelines".ti,ab OR "guideline".ti,ab) ADJ2 ("consensus".ti,ab OR "delphi".ti,ab))) AND ("reporting".ti OR "quality of reporting".ti,ab OR "reporting quality".ti,ab OR "reporting qualities".ti,ab OR "selective reporting".ti,ab OR "poor reporting".ti,ab OR "poor reported".ti,ab OR "poorly reported".ti,ab OR *"Methodology"/ OR *"Writing"/ OR "research method".ti OR "research methods".ti OR "research method*".ti OR ("reporting" ADJ8 ("quality" OR "selective" OR "poor" OR "weak" OR "manner" OR "rigor" OR "improv*")).ti,ab))

Cochrane
https://www.cochranelibrary.com/advanced-search/search-manager

(("Delphi Technique" OR "Delphi Technique" OR "Delphi techniques" OR "Delphi method" OR "Delphi methods" OR "Delphi study" OR "Delphi studies" OR "Delphi survey" OR "Delphi surveys" OR "Delphi consensus" OR "Delphi based consensus" OR "Delphi questionnaire" OR "Delphi questionnaires" OR "Delphi research" OR "Delphi review" OR "Delphi reviews" OR "Delphi process" OR "Delphi processes" OR "Delphi based" OR "Delphi procedure" OR "Delphi procedures" OR "Delphi assessment" OR "Delphi assessments" OR "Delphi approach" OR "Delphi approaches" OR "Delphi panel" OR "Delphi panels" OR "Delphi round" OR "Delphi rounds" OR "Delphi analysis" OR "Delphi expert" OR "Delphi experts" OR "Delphi consultation" OR "Delphi methodology" OR "nominal group technique" OR "nominal group techniques" OR "nominal group" OR "nominal groups" OR "nominal grouping" OR "consensus recommendation" OR "consensus recommendations" OR "consensus development" OR "consensus activity" OR "consensus activities" OR "Consensus Development Conference" OR "Consensus Development" OR "Consensus methodology" OR "consensus method*" OR "RAND" OR (("Guidelines" OR "guideline") NEAR/2 ("consensus" OR "delphi"))) AND ("quality of reporting" OR "reporting quality" OR "reporting qualities" OR "selective reporting" OR "poor reporting" OR "poor reported" OR ("reporting" NEAR/5 ("quality" OR "selective" OR "poor" OR "weak" OR "manner" OR "rigor" OR "improv*")) OR "Data Accuracy" OR "Research Report standards" OR "quality assessment" OR "strengths" OR "strength" OR "weaknesses" OR "weakness" OR "research method" OR "research methods" OR "research method*")):ti,ab,kw

Emcare
http://ovidsp.ovid.com/ovidweb.cgi?T=JS&NEWS=n&CSC=Y&PAGE=main&D=emcr

((exp *"Delphi Study"/ OR "Delphi Technique".ti OR "Delphi techniques".ti OR "Delphi method".ti OR "Delphi methods".ti OR "Delphi study".ti OR "Delphi studies".ti OR "Delphi survey".ti OR "Delphi surveys".ti OR "Delphi consensus".ti OR "Delphi based consensus".ti OR "Delphi questionnaire".ti OR "Delphi questionnaires".ti OR "Delphi research".ti OR "Delphi review".ti OR "Delphi reviews".ti OR "Delphi process".ti OR "Delphi processes".ti OR "Delphi based".ti OR "Delphi procedure".ti OR "Delphi procedures".ti OR "Delphi assessment".ti OR "Delphi assessments".ti OR "Delphi approach".ti OR "Delphi approaches".ti OR "Delphi panel".ti OR "Delphi panels".ti OR "Delphi round".ti OR "Delphi rounds".ti OR "Delphi analysis".ti OR "Delphi expert".ti OR "Delphi experts".ti OR "Delphi consultation".ti OR "Delphi methodology".ti OR "nominal group technique".ti OR "nominal group techniques".ti OR "nominal group".ti OR "nominal groups".ti OR "nominal grouping".ti OR "consensus recommendation".ti OR "consensus recommendations".ti OR "consensus development".ti OR "consensus activity".ti OR "consensus activities".ti OR "consensus methodology".ti OR "consensus method*".ti OR exp *"Consensus Development"/ OR "RAND".ti OR (("Guidelines".ti,ab OR "guideline".ti,ab) ADJ2 ("consensus".ti,ab OR "delphi".ti,ab))) AND ("reporting".ti

OR "quality of reporting".ti,ab OR "reporting quality".ti,ab OR "reporting qualities".ti,ab OR "selective reporting".ti,ab OR "poor reporting".ti,ab OR "poor reported".ti,ab OR "poorly reported".ti,ab OR *"Methodology"/ OR *"Writing"/ OR "research method".ti OR "research methods".ti OR "research method*".ti OR ("reporting" ADJ8 ("quality" OR "selective" OR "poor" OR "weak" OR "manner" OR "rigor" OR "improv*")).ti,ab))

Academic Search Premier
http://search.ebscohost.com/login.aspx?authtype=ip,uid&profile=lumc&defaultdb=aph

((TI("Delphi Technique" OR "Delphi Technique" OR "Delphi techniques" OR "Delphi method" OR "Delphi methods" OR "Delphi study" OR "Delphi studies" OR "Delphi survey" OR "Delphi surveys" OR "Delphi consensus" OR "Delphi based consensus" OR "Delphi questionnaire" OR "Delphi questionnaires" OR "Delphi research" OR "Delphi review" OR "Delphi reviews" OR "Delphi process" OR "Delphi processes" OR "Delphi based" OR "Delphi procedure" OR "Delphi procedures" OR "Delphi assessment" OR "Delphi assessments" OR "Delphi approach" OR "Delphi approaches" OR "Delphi panel" OR "Delphi panels" OR "Delphi round" OR "Delphi rounds" OR "Delphi analysis" OR "Delphi expert" OR "Delphi experts" OR "Delphi consultation" OR "Delphi methodology" OR "nominal group technique" OR "nominal group techniques" OR "nominal group" OR "nominal groups" OR "nominal grouping" OR "consensus recommendation" OR "consensus recommendations" OR "consensus development" OR "consensus activity" OR "consensus activities" OR "Consensus Development Conference" OR "Consensus Development" OR "Consensus methodology" OR "consensus method*" OR "RAND" OR (("Guidelines" OR "guideline") N2 ("consensus" OR "delphi"))) OR KW("Delphi Technique" OR "Delphi Technique" OR "Delphi techniques" OR "Delphi method" OR "Delphi methods" OR "Delphi study" OR "Delphi studies" OR "Delphi survey" OR "Delphi surveys" OR "Delphi consensus" OR "Delphi based consensus" OR "Delphi questionnaire" OR "Delphi questionnaires" OR "Delphi research" OR "Delphi review" OR "Delphi reviews" OR "Delphi process" OR "Delphi processes" OR "Delphi based" OR "Delphi procedure" OR "Delphi procedures" OR "Delphi assessment" OR "Delphi assessments" OR "Delphi approach" OR "Delphi approaches" OR "Delphi panel" OR "Delphi panels" OR "Delphi round" OR "Delphi rounds" OR "Delphi analysis" OR "Delphi expert" OR "Delphi experts" OR "Delphi consultation" OR "Delphi methodology" OR "nominal group technique" OR "nominal group techniques" OR "nominal group" OR "nominal groups" OR "nominal grouping" OR "consensus recommendation" OR "consensus recommendations" OR "consensus development" OR "consensus activity" OR "consensus activities" OR "Consensus Development Conference" OR "Consensus Development" OR "Consensus methodology" OR "consensus method*" OR "RAND" OR (("Guidelines" OR "guideline") N2 ("consensus" OR "delphi")))) AND (TI("quality of reporting" OR "reporting quality" OR "reporting qualities" OR "selective reporting" OR "poor reporting" OR "poor reported" OR "reporting guideline" OR "reporting" OR ("reporting" AND ("quality" OR "selective" OR "poor" OR "weak" OR "manner" OR "rigor" OR "improv*")) OR "Data Accuracy" OR "quality assessment" OR "strengths" OR "strength" OR "weaknesses" OR "weakness" OR "research method" OR "research methods" OR "research method*") OR

KW("quality of reporting" OR "reporting quality" OR "reporting qualities" OR "selective reporting" OR "poor reporting" OR "poor reported" OR ("reporting" N5 ("quality" OR "selective" OR "poor" OR "weak" OR "manner" OR "rigor" OR "improv*")) OR "Data Accuracy" OR "Research Report standards" OR "quality assessment" OR "strengths" OR "strength" OR "weaknesses" OR "weakness" OR "research method" OR "research methods" OR "research method*") OR AB("quality of reporting" OR "reporting quality" OR "reporting qualities" OR "selective reporting" OR "poor reporting" OR "poor reported" OR ("reporting" N5 ("quality" OR "selective" OR "poor" OR "weak" OR "manner" OR "rigor" OR "improv*")) OR "Data Accuracy" OR "Research Report standards" OR "quality assessment" OR "strengths" OR "strength" OR "weaknesses" OR "weakness")))

PsycINFO
http://search.ebscohost.com/login.aspx?authtype=ip,uid&profile=lumc&defaultdb=psyh

((TI("Delphi Technique" OR "Delphi Technique" OR "Delphi techniques" OR "Delphi method" OR "Delphi methods" OR "Delphi study" OR "Delphi studies" OR "Delphi survey" OR "Delphi surveys" OR "Delphi consensus" OR "Delphi based consensus" OR "Delphi questionnaire" OR "Delphi questionnaires" OR "Delphi research" OR "Delphi review" OR "Delphi reviews" OR "Delphi process" OR "Delphi processes" OR "Delphi based" OR "Delphi procedure" OR "Delphi procedures" OR "Delphi assessment" OR "Delphi assessments" OR "Delphi approach" OR "Delphi approaches" OR "Delphi panel" OR "Delphi panels" OR "Delphi round" OR "Delphi rounds" OR "Delphi analysis" OR "Delphi expert" OR "Delphi experts" OR "Delphi consultation" OR "Delphi methodology" OR "nominal group technique" OR "nominal group techniques" OR "nominal group" OR "nominal groups" OR "nominal grouping" OR "consensus recommendation" OR "consensus recommendations" OR "consensus development" OR "consensus activity" OR "consensus activities" OR "Consensus Development Conference" OR "Consensus Development" OR "Consensus methodology" OR "consensus method*" OR "RAND" OR (("Guidelines" OR "guideline") N2 ("consensus" OR "delphi"))) OR AB("Delphi Technique" OR "Delphi Technique" OR "Delphi techniques" OR "Delphi method" OR "Delphi methods" OR "Delphi study" OR "Delphi studies" OR "Delphi survey" OR "Delphi surveys" OR "Delphi consensus" OR "Delphi based consensus" OR "Delphi questionnaire" OR "Delphi questionnaires" OR "Delphi research" OR "Delphi review" OR "Delphi reviews" OR "Delphi process" OR "Delphi processes" OR "Delphi based" OR "Delphi procedure" OR "Delphi procedures" OR "Delphi assessment" OR "Delphi assessments" OR "Delphi approach" OR "Delphi approaches" OR "Delphi panel" OR "Delphi panels" OR "Delphi round" OR "Delphi rounds" OR "Delphi analysis" OR "Delphi expert" OR "Delphi experts" OR "Delphi consultation" OR "Delphi methodology" OR "nominal group technique" OR "nominal group techniques" OR "nominal group" OR "nominal groups" OR "nominal grouping" OR "consensus recommendation" OR "consensus recommendations" OR "consensus development" OR "consensus activity" OR "consensus activities" OR "Consensus Development Conference" OR "Consensus Development" OR "Consensus methodology" OR "consensus method*" OR "RAND") OR KW("Delphi Technique" OR "Delphi

Technique" OR "Delphi techniques" OR "Delphi method" OR "Delphi methods" OR "Delphi study" OR "Delphi studies" OR "Delphi survey" OR "Delphi surveys" OR "Delphi consensus" OR "Delphi based consensus" OR "Delphi questionnaire" OR "Delphi questionnaires" OR "Delphi research" OR "Delphi review" OR "Delphi reviews" OR "Delphi process" OR "Delphi processes" OR "Delphi based" OR "Delphi procedure" OR "Delphi procedures" OR "Delphi assessment" OR "Delphi assessments" OR "Delphi approach" OR "Delphi approaches" OR "Delphi panel" OR "Delphi panels" OR "Delphi round" OR "Delphi rounds" OR "Delphi analysis" OR "Delphi expert" OR "Delphi experts" OR "Delphi consultation" OR "Delphi methodology" OR "nominal group technique" OR "nominal group techniques" OR "nominal group" OR "nominal groups" OR "nominal grouping" OR "consensus recommendation" OR "consensus recommendations" OR "consensus development" OR "consensus activity" OR "consensus activities" OR "Consensus Development Conference" OR "Consensus Development" OR "Consensus methodology" OR "consensus method*" OR "RAND" OR (("Guidelines" OR "guideline") N2 ("consensus" OR "delphi")))) AND (TI("quality of reporting" OR "reporting quality" OR "reporting qualities" OR "selective reporting" OR "poor reporting" OR "poor reported" OR "reporting guideline" OR "reporting" OR ("reporting" AND ("quality" OR "selective" OR "poor" OR "weak" OR "manner" OR "rigor" OR "improv*")) OR "Data Accuracy" OR "quality assessment" OR "strengths" OR "strength" OR "weaknesses" OR "weakness" OR "research method" OR "research methods" OR "research method*") OR KW("quality of reporting" OR "reporting quality" OR "reporting qualities" OR "selective reporting" OR "poor reporting" OR "poor reported" OR ("reporting" N5 ("quality" OR "selective" OR "poor" OR "weak" OR "manner" OR "rigor" OR "improv*")) OR "Data Accuracy" OR "Research Report standards" OR "quality assessment" OR "strengths" OR "strength" OR "weaknesses" OR "weakness" OR "research method" OR "research methods" OR "research method*") OR AB("quality of reporting" OR "reporting quality" OR "reporting qualities" OR "selective reporting" OR "poor reporting" OR "poor reported" OR ("reporting" N5 ("quality" OR "selective" OR "poor" OR "weak" OR "manner" OR "rigor" OR "improv*")) OR "Data Accuracy" OR "Research Report standards" OR "quality assessment" OR "strengths" OR "strength" OR "weaknesses" OR "weakness")))

Author, year	
Assessor	

Background 1.1 Does the study suggest anything about how or if consensus papers should report the context or rationale for choosing a consensus method over other methods?	
Background 1.2 Does the study suggest anything about how/what or if consensus papers should report the objectives of the consensus exercise?	

Methods 2.1 Does the study suggest anything about how/what or if consensus papers should report regarding: A literature search/strategy?	
Methods 2.2 Does the study the suggest anything about how/what or if consensus papers should report regarding: Inclusion and exclusion criteria for the literature search?	
Methods 2.3 Does the study suggest anything of what or if consensus report should report on panel composition, n of participants, expertise, origin? Prespecified?	
Methods 2.4 Does the study suggest anything of how or if PPI (public patient involvement) activity should be reported?	
Methods 2.5 Does the study suggest anything about what or if consensus papers should report regarding panel recruitment strategies, invitations? Any level of detail specified?	
Methods 2.6 Does the study suggest how or if consensus papers should report the consensus criteria/threshold (or the level of agreement considered to reach consensus)?	
Methods 2.7 Does the study suggest how or if consensus papers should report how decision of approval of an item will be made?	
Methods	

2.8 Does the study suggest anything about what level of detail should be reported regarding the number of Delphi rounds or if this should be reported?	
Methods 2.9 Does the study suggest anything about what level of detail should be reported regarding the criteria used for defining the number of rounds? (why 2-3 or more e.g.) or if this should be reported?	
Methods 2.10 Does the study suggest anything about the details that should be reported regarding the time between rounds, if it should be prespecified or if this should be reported?	
Methods 2.11 Does the study suggest anything about details that should be reported of the names of the techniques of non-Delphi methods used to gather participants' inputs and reach consensus?	
Methods 2.12 Does the study suggest anything of what or in which detail should be reported regarding tool or electronic system used for Delphi? (If Delphi was used)? Or if this should be reported?	
Methods 2.13 Does the study suggest anything about how or in what level of detail the anonymity of participants (in Delphi or other methods) has to be reported? Or if this should be reported?	
Methods 2.14 Does the study suggest anything about how to report, and in what level of detail, the feedback for panellists (in Delphi rounds or other methods) process? Or if this should be reported?	
Methods 2.15 Does the study suggest anything about how or if data synthesis/analysis should be reported (from any consensus method used and how this was calculated statistically) and in what level of detail?	
Methods 2.16 Does the study suggest anything about how or if piloting should be reported and in what level of detail (e.g. understanding of consensus items, platforms used, tools used)?	
Methods	

2.17 Does the study suggest anything about how or if the role of Steering Committee members should be reported?	
Methods 2.18 Does the study suggest anything on what or if should be described regarding COI or funding?	
Methods 2.19 Does the study suggest anything on what should be described of how is dealt with COI of panellist (not allowed to vote when there is COI)? Or if this should be described	

Results 3.1 Does the study suggest anything on how to report the initial evidence search (presentation of results of the literature review)?	
Results 3.2 Does the study suggest anything on how to report n of studies found?	
Results 3.3 Does the study recommend which detail should be used when reporting panellists drop-outs (numbers and reasons)? Or if this should be reported?	
Results 3.4 Does the study suggest how or if approval rates per item shared with respondents for each round should be reported in the Results section?	
Results 3.5 Does the study suggest anything about in which detail the items that have been dropped should be reported? (reasons e.g.) Or if this should be reported?	
Results 3.6 Does the study make any recommendation on how to report the collection, synthesis and use of comments from panellists? Or if this should be reported?	
Results 3.7 Does the study suggest regarding how the final list of items (for clinical guideline or reporting guideline) should be reported? Or if this should be reported?	

Discussion 4.1 Does the paper suggest anything about reporting the limitations and strengths of the study and how? Or if this should be reported?	
Discussion	

4.2 Does the paper suggest anything about what or in which detail the applicability generalisability, and reproducibility of the study should be reported? Or if this should be reported?	

5.1 Any other item proposed by the paper that is not captured in other columns?	
5.2 Any other item not proposed by the paper, but you think that could be added (not fitting the categories above)?	

Examples of text with well reported methods/results (for E&E document) - write NA if none was cited or found by you	
Additional comments from assessor	

Data on reporting quality (*recommendations in italics)*

Study	What is stated regarding reporting quality?
Banno 2019[32]	• "The reporting quality of the Delphi technique in reporting guidelines is unknown even though the use of the Delphi technique was recommended in the guidance for reporting guidelines." (Note: This is a protocol for the systematic review of 2020.) *4 quality score items are summarised of Delphi methods used in reporting guidelines*.
Banno 2020[16]	• "Reproducible criteria of participants, number of rounds, criteria for dropping items, and stopping criteria other than rounds were found for 87%, 97%, 69%, and 13%, respectively of reporting guidelines developed with the Delphi method. The total score of reporting quality was 2 or more in 94% of reporting guidelines using the Delphi method." *4 quality score items are summarised of Delphi methods used in reporting guidelines*.
Boulkedid 2011[17]	• "Study reports did not consistently provide details that are important for interpreting the results. For example, only 39% of studies reported that individual feedback was given between rounds and the method used to define a consensus was specified in only 77% studies. Moreover, response rates for all rounds were reported in only 31% of studies. Information on both points is needed to evaluate the validity and credibility of the results. If the Delphi method is incompletely described this may affect the overall quality of the final consensus and the selected indicators are unlikely to gain the level of credibility needed for adoption I clinical practice." • "The Delphi procedure is valuable for achieving a consensus about issues where none existed previously. However, our findings indicate a need for improving the use and reporting of this technique." *Table 5 provides recommendations for reporting the Delphi procedure.*
Chan 2019[20]	• "This lack of clear definition has led to considerable confusion and substantial variation in the quality of reporting of Delphi studies" • "One-third of medical education Delphi studies failed to report that a literature review on the topic of interest had been conducted , and over half failed to report key aspects such as what background information was provided to participants; the response rate for each round; what formal feedback of group rating was shared between rounds; a statement that anonymity was maintained; and a clear definition of consensus." • "Lack of clarity in the report in the reporting of procedures and methodological choices associated with the modified Delphi studies can prevent readers from effectively appraising and interpreting findings." • "Methodological rigor and transparent reporting are essential to assure readers that the consensus results are applicable to their environment, and to translate expert opinion into practice." *Box 1 provides recommendations to improve reporting*.
Diamond 2014[18]	• "Definitions of consensus vary widely and are poorly reported. Improved criteria for reporting of methods of Delphi studies are required." • "Methodologic criteria are proposed for the reporting of Delphi studies." • "Despite the fact that the most Delphi studies in our cohort had consensus as their aim, in only a minority of the Delphi studies reviewed was consensus defined with a specific criterion. Furthermore, this criterion was the reason for termination of the Delphi process, usually on the basis of an *a priori* definition." • "We believe that there is a need to improve the reporting of Delphi studies, along the lines of a CONSORT-like guideline, as is used for randomized controlled trials." *Methodologic criteria are proposed for the reporting of Delphi studies.*

Gattrell 2019[29]	"At present there are a lack of standard, validated reporting guidelines for publications reporting the results of Delphi panel studies." Quality assessment: Methodological quality • The type of Delphi technique used, or the modifications to the method, was not outlined in all publications (included in 62/90 publications; 68.9%). • Just over half of all publications stated that there was some diversity amongst participants and clearly outlined the methods for the selection of panellists. • Agreement and consensus thresholds should be defined prior to study commencement, but in 40% of publications it was unclear, or not stated whether these thresholds were predefined. • Anonymised responses are typically conveyed back to the group after each round, but this was clearly reported in less than half (38.9%) of publications. Quality assessment: Reporting quality and transparency (Figure 3b). • The funding source was not clearly disclosed in over a third of publications, and almost twice as many publications did not clearly disclose the funder's role. • Conflicts of interest were clearly described in most publications (included in 79/90 publications; 87.8%). • Clear disclosure of external support was not evident in the majority of the publications.
Grant 2018[24]	• "Specifying the analysis procedure for consensus is therefore a critical consideration when designing consensus-oriented Delphi processes in health research." • "Without prespecifying their analysis procedures in a study registry, health researchers conducting consensus-oriented Delphi processes can mine for and selectively report the most desirable set of items reaching consensus and even present the reported analysis as the only one conducted. Undisclosed flexibility in data collection, analysis, and reporting is a growing concern in empirical research." • "Without preregistering and reporting all of the attempted analysis procedures and when they were attempted, the extent and impact of researchers trying different analysis procedures is nearly impossible for peer reviewers, editors, and consumers of Delphi research to assess." • "To be completely registered, the preanalysis plan should precisely describe the essential elements of the analysis procedure for determining consensus (see Box 2)." • "Researchers should use existing guidance on reporting completed Delphi processes to provide sufficient information for comparing the final article to the registered preanalysis plan [1,12,42], with particular attention in the final article to any changes from the preanalysis plan in the items, rating criteria, analytic procedure (measure and threshold), and data and participants included in the analysis." *Box 2 provides a minimum set of items to include in prospectively registered preanalysis plans for consensus-oriented Delphi processes.*
Hasson 2017[27]	• "Figure 1 Areas for reporting on the Delphi survey technique." • "In Delphi surveys there exists no consistent method for reporting findings (Schmidt 1997) and a review of the literature showed that a number of approaches have been used." • "The following diagram attempts to outline those sections that researchers should report upon when using the Delphi. This will help readers to judge the reliability of the method and the results obtained." *Followed by a checklist of issues, which could be used by researchers.*

Humphrey-Murto 2017[21]	• "The authors set out to describe the use of consensus methods in medical education research and to assess the reporting quality of these methods and results." • "Improved criteria for reporting are needed." • "Our findings suggest that the reporting quality and standardization of consensus methods in medical education research varies greatly. The following areas appeared particularly problematic and were often left out or poorly described in the articles we reviewed: conducting a literature review to inform the consensus method; providing background information to participants; reporting the number of participants after each round; describing the level of anonymity used in the study; providing participants with feedback of group ratings; and articulating the definition of consensus used in the study." *Recommendations for improvements in these areas are provided in Discussion.*
Humphrey-Murto 2017[28]	• "Consensus group methods are widely used in research to identify and measure areas where incomplete evidence exists for decision-making. Despite their widespread use, these methods are often inconsistently used and reported." • "This paper and associated Guide aim to describe these methods and to highlight common weaknesses in methodology and reporting." • "The AMEE Guide describes these methods to provide a "how to" approach, highlight common weaknesses in methodology and reporting, and outline recommendations for reporting future consensus based studies." • "Four recent reviews using the Delphi in health care and policy-related research have systematically explored deficiencies in the use and reporting of consensus group methods. Collectively, these studies have noted deficiencies regarding: information provided to the participants at the start of Delphi, reporting response rates, feedback to participants, level of anonymity, outcomes after each round and the definition of consensus." *This guide provides recommendations for improvement of reporting.*
Humphrey-Murto 2019[25]	• "Studies using the Delphi for selecting performance indicators for healthcare, for medical and nursing education, or for determining outcomes to measure in clinical trials, often fail to adequately report sufficient methodological detail. Examples include poor reporting of background information provided to participants, response rates for all rounds, level of anonymity, formal feedback between rounds, and the definition of consensus." *OMERACT Delphi consensus checklist is provided in Figure 1.*
Jünger 2017[12]	• "Substantial variation was found concerning the quality of the study conduct and the transparency of reporting of Delphi studies used for the development of best practice guidance in palliative care. Since credibility of the resulting recommendations depends on the rigorous use of the Delphi technique, there is a need for consistency and quality both in the conduct and reporting of studies. To allow a critical appraisal of the methodology and the resulting guidance, a reporting standard for Conducting and Reporting of DElphi Studies (CREDES) is proposed." Study adds in Box 3"*Recommendations for the Conducting and REporting of DElphi Studies (CREDES)."*
Ng 2018[30]	• "Given the variance in the use of Delphi method, reporting guidelines could help improve reporting of this research, and thereby allow readers to be aware of the accuracy of data and conclusions." • "We anticipate the implementation of this will promote transparent and accurate reporting of research using Delphi method for obtaining quantitative data." *A set of reporting guidelines is proposed.*
Niederberger 2020[26]	• "Significant weaknesses exist in the quality of the reporting."

	• "Criteria for evaluating the quality of their execution and reporting also appear to be necessary." • "A specific definition of the underlying Delphi technique was found in 61% (ID11) and 88.2% (ID4) of the Delphi articles investigated." • "Most of the Delphi studies analyzed in the reviews reported on the number of participating experts. The rates for the initial round were between 84% (ID6) and 100% (ID12). Four of the reviews investigated whether the number of experts was stated for each round (ID4, ID7, ID11, ID12). In one review based on 10 Delphi studies from health sciences (ID7), the authors discovered that the number of experts per round was stated in all articles. A review of 48 studies in a medical context indicated that the number of invited experts was stated less frequently with each round (ID6). Seven of the 12 reviews investigated whether the backgrounds of the experts had been reported, what kind of expertise they possessed, and the criteria according to which they were selected (ID1, ID3, ID4, ID6, ID9, ID11, ID12). One review of Delphi techniques in a health context determined that the criteria for selecting the experts was reproduced in 65 of 100 articles (65%) (ID3) included in that particular review. In other reviews with a more specific focus, such as on health care, palliative medicine, or health promotion, the rates were higher at 69% (ID11), 70% (ID9) and 79% (ID1), respectively. Based on the results of the reviews, the criteria by which the experts were selected and approached was not always clear. In one review of 100 studies from the care sector, the proportion of articles with unclear selection criteria was 11.2% (ID4), while the proportion was 93.3% in a review of 15 studies from the clinical sector (ID12)." • "Seven of the 12 reviews determined whether and when consensus was defined in the Delphi studies (ID1, ID3, ID4, ID6, ID9, ID11, ID12). The number of studies in which consensus was defined in the article was between 73.5% (ID3) and 83.3% (ID9) in the reviews." • "The authors of seven reviews investigated whether the number of Delphi rounds was published (ID1, ID3, ID4, ID6, ID9, ID11, ID12). The number of Delphi rounds was stated in most of the Delphi studies (e.g., ID1 82.5%, ID4 91%, ID6 100%, ID9 49.3%, ID12 93.3%). Six of the reviews included a report of the generation of the questionnaire (ID1, ID4, ID6, ID9, ID11, ID12). They demonstrated that up to 96.3% of the investigated articles reported on how the items for the questionnaire were developed (ID1). In contrast, this rate stood at 33.3% in the review of palliative care articles (ID9). The authors of two reviews investigated the question of how the items were changed during the Delphi process based on the judgments submitted by the experts (ID3, ID12). In one of the reviews, the authors indicated that 59% of the analyzed articles had defined criteria for dropping items (ID3). In another review, the authors stated that all of the investigated Delphi studies included a report of "what was asked in each round" (ID12, p. 2). The authors of the reviews reported about the feedback in most of the Delphi studies (ID11 67.9%, ID12 93.3%). The information provided about the response rate per Delphi round was less (ID1 and ID4 39%). According to the results of the reviews, around half of the studies did not provide information about the feedback design between the Delphi rounds (ID1 40%, ID4 55.1%, ID6 37.7% ID12 40%). According to the authors of the review on health promotion, the process—from formulating the issue being investigated through to the development of the questionnaire—was in general similar to a "black box," and the methodological quality of the survey instrument was almost impossible to evaluate using the published information (ID11, p. 318)." • "Our results also indicate deficits both in carrying out and also reporting Delphi techniques."

	• "The findings in the reviews we analyzed indicated that there is no uniform process for carrying out and reporting Delphi techniques."
Paré 2013[22]	• "Thirty-one percent of the articles in our sample provided a detailed description of the expert recruitment and selection process, 43% provided only limited details, and 26% did not provide any details." • "All of the articles in our database (n = 42) specified the criteria that were used to select the panel of experts. Position is by far the most used criteria (71%), followed by relevant professional experience (57%), geographic location (7%), and education level (5%)." • "38% of the studies provided detailed information about the participating experts [e.g., 44], 40% provided minimal information [e.g., 2], and 22% did not provide any description". • "The anonymity of the experts was reported in virtually all of the studies (95%) in our sample." • "Only 29% of all of the studies reported the response rate to the initial request for participation." • "35 studies (83%)reported the size of the panels. The majority of the studies (n = 21) reported a panel size between 7 and 30, only one study reported a size of 6 or less, and 13 studies reported panel sizes above 30. Nine studies (19%) examined multiple panels of experts." • "Only 17% of these Delphi studies reported that a pretest of the instruments had been conducted." • "24 studies out of 27 (89%) reported the brainstorming instructions that were sent to the experts." • "Only 8 studies (30%) reported the use of this recommendation. (i.e. Have the experts comment and validate the consolidated list)." • "The vast majority of the studies (85%) reported the final number of items at the end of phase 1." • "Among the 25 studies that did not include this phase (i.e. narrowing down phase), 68% explicitly justified this choice (e.g., the number of items at the end of phase 1 are equal or less than 20 as suggested by Schmidt." • "All 17 studies clearly described the narrowing down instructions that were given to the experts." • "65% of the studies clearly specified their item selection rule." • "Most of the studies (82%) reported the final number of items at the end of the second phase." • "All 42 articles described clearly the ranking instructions that were provided to the experts." • "Almost all of the studies (95%) in our sample reported the statistics that were used for data analysis." • "31% of the studies in our database specified a clear stopping rule." • "Only 15 studies (36%) reported the final consensus rate." • "29 of the 42 studies had multiple rounds of ranking. Of these, the feedback that was provided to the experts in between the rounds included the mean ranks of items (69% of studies), an interpretation of the Kendall's W coefficient (3%), the expert's prior responses (59%), and the comments made by the other experts (38%)." *Recommendations regarding what to report are provided throughout the Results section as well as in the Discussion.*
Resemann 2018[31]	• "Reporting of the Delphi method was critiqued against the AGREE Reporting Checklist." • "All studies reported consensus results. The majority (8/11 [73%]) used a two-stage modified Delphi method, while the remainder used a classic three-stage process. Literature searches guided the development of statements for Delphi panel review in the majority of studies, but only 2/11 (18%) conducted

	systematic literature reviews and merely 6/11 (55%) of studies reported the number of statements assessed. Furthermore, 7/11 (64%) did not report collecting panellist feedback to inform subsequent Delphi stages, 5/11 (45%) of studies did not describe the rating scales used, and 2/11 (18%) omitted reporting the level of consensus reached" • "There is a need for improved reporting of Delphi methods".
Waggoner 2016[23]	• "Despite the widespread utility of consensus methods and the variety of approaches available, there is a lack of guidelines for conducting such studies. This lack of stringency in guidelines for conducting consensus studies has led to variability not only in reporting results but in conducting the studies themselves." • "Many studies describe their methods for collecting data and that they did have a benchmark that would point to a consensus, but a lack of a description of the analytical techniques is apparent in many studies." • "In addition to the lack of descriptive techniques in these articles, there is a wide range of criteria that points to consensus. How these particular benchmarks are determined is also not a topic in many of the studies. Given the lack of current research, we believe that the methodology used I subsequent studies should be described more thoroughly in the manuscript." • "We set out to determine best practices for conducting such research as well as reporting on results in the hopes that future studies are more reliable and valid." *This article provides guidance for reporting of various consensus methods.*
Wang 2015[19]	• "Adoption of reporting guidelines is associated with improved reporting quality of research." • "For example, 28 % of the included guidelines reported no information about consensus, and 57 % were silent about how the feedback after consensus was dealt with." • "In addition to the methodology, only 31 % reported formal consensus method." • "Among guidelines developed through consensus, 30 (50 %) reported group member identification and 31 (52 %) reported member recruitment. Of those who identified members, 27 (45 %) reported specialties of experts, 20 (32 %) described information of members, such as names and institutions, and four (7 %) gave the selection criteria. For those who recruited members, even (12 %) described the recruit methods, for instance, through e-mail, study co-chairs, or group decision. In guidelines developed by a working group, 22 (37 %) reported the number of experts participating in guideline development (median 32, range 3–115). Eleven (18 %) guidelines reported the endpoint of consensus process, which were all terminated after a fixed number of rounds (Table 2). In addition, the inclusion criteria of items were given in eight (13 %) guidelines. For example, items meeting the median score of eight or higher in the final round were included." • "11 (18 %) described the pilot methods, seven (12 %) described the feedback information requirement and five (8 %) gave the methods for feedback collection." • "More than 30 % of the reporting guidelines did not report consensus. For those who did, details of consensus methods were poorly reported." • "Consensus methods should be supported by developers, and the reporting of the methods should be improved." *Recommendations for Consensus methods are provided, but more about improvement of applying and reporting using all other reporting guidelines, but some items are applicable for consensus methodology as well (e.g. reporting COI and funding.*

Background 1.1 Does the study suggest anything about how or if consensus papers should report the context or rationale for choosing a consensus method over other methods?	1) Research problem clearly defined and topic and method justification should be reported [Hasson 2000, Figure 1 and page 1013] 2) Selection of one consensus method over another should be evident if the purpose is clearly stated. [Humphrey-Murto 2017 Med Teach page 16] 3) What is the rationale for selecting the Delphi procedure? [Humphrey-Murto 2019, Figure 1] 4) The choice of the Delphi technique as a method of systematically collating expert consultation and building consensus needs to be well justified. A rationale for the choice of the Delphi technique as the most suitable method needs to be provided [Jünger 2017, Box 3, items 1 and 8]
Background 1.2 Does the study suggest anything about how/what or if consensus papers should report the objectives of the consensus exercise?	1) Define the study objective [Boulkedid 2011, Table 5 page 7] 2) Define the purpose of the study [Chan 2019, Box 1] 3) Is the objective of the Delphi study to present results (eg, a list or statement) reflecting the consensus of the group, or does the study aim to merely quantify the level of agreement? [Diamond 2014, Table 6 and page 403] If the aim of the Delphi study is to elicit consensus, then a clear definition for what constitutes consensus should be provided a priori together with threshold values that specify when consensus is reached. If the investigators plan to only quantify the degree of consensus, but not have consensus as a criterion to stop the Delphi study, this should also be explicitly stated [Diamond 2014, page 406] 4) Research problem clearly defined and topic and method justification should be reported [Hasson 2020, Figure 1 and page 1013] 5) Authors must provide a clear purpose for their study or line of inquiry [Humphrey-Murto 2017 Med Teach, page 16] 6) 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 [Jünger 2017, item 8]

	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" [Jünger 2017, item 2]

Methods 2.1 Does the study the suggest anything about how/what or if consensus papers should report regarding: A literature search/strategy?	1) Describe the selection and preparation of the scientific evidence for the participants [Chan 2019, Box 1] 2) A literature review should be reported [Hasson 2000, Figure 1] 3) "We suggest that this important step must be described", but they don't say how. [Humphrey-Murto 2017 AMA, page 1493 and 1496 Partially] 4) Describe the selection and preparation of the scientific evidence for the participants [Humphrey-Murto 2017 Med Teach, page 16] 5) Only implying it should happen and be reported [Resemann 2018]
Methods 2.2 Does study the suggest anything about how/what or if consensus papers should report regarding: Inclusion and exclusion criteria for the literature search?	1) Clear definition of the selection criteria and/or the definition used in the Delphi questionnaire; criteria for selection should be reported [Boulkedid 2011, Table 5, Appendix S1 item 2] 2) Describe how items were selected for inclusion in questionnaire, in sufficient detail [Chan 2019, Box 1] 3) Clear selection criteria should be prespecified [Paré 2013 page 210]
Methods 2.3 Does the study suggest anything of what or if consensus report should report on panel composition, n of participants, expertise, origin? Prespecified?	1) The method used to select participants is stated. Number and type of participant subgroups (eg, patients, generalists and experts) are needed [Banno 2019, page 2 item 1] 2) The method to include and exclude participants was described. The number and type of participant subgroups (e.g., patients, generalists, and experts) were essential to record [Banno 2020, page 52 item 1] 3) How the experts were chosen (e.g., willingness to participate, expertise, or membership in an organization); Composition and characteristics of the panel, number of participants (diagram of participant flow); number invited, how they were chosen, whether they were described (age, sex, specialty), years of experience, single or from multiple

	specialties, inclusion of multiple stakeholders, types of stakeholders [Boulkedid 2011, page 2, Table 5, Appendix S1 item 9-15] 4) Describe how participants were selected and their qualifications. Include description of facilitator credentials [Chan 2019, Box 1] 5) Were criteria for participants reproducible? How will participants be selected or excluded? [Diamond 2014, Table 5 and 6] 6) Was there heterogeneity in panel membership and is the method for selection of experts clearly defined [Gattrell 2019, Table 1] 7) Expert selection process and characteristics should be reported in detail [Hasson 2000, page 1009, 1013] 8) How many participants were involved? We noted that the type of expertise required of participants was usually not clearly described [Humphrey-Murto 2017 AMA, page 1493 and 1494] 9) Describe how the participants were selected and their qualifications: if the NGT or RAND/UCLA is used, describe facilitator's credentials. Whatever the makeup of the expert panel, the authors must provide a rationale and justify their choices [Humphrey-Murto 2017 Med Teach] 10) How many stakeholder/participant groups will be involved in each step? Provide a rationale for inclusion or exclusion and define the stakeholder groups [Humphrey-Murto 2019, Fig 4] 11) 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 [Jünger 2017, Box 3 9] 12) Describing expert panel selection with eligibility criteria and including conflicts of interest [Ng 2018] 13) The number of experts in each round should be stated. The backgrounds of the experts should be reported, what kind of expertise they possessed, and the criteria according to which they were selected [Niederberger 2020, page 4]

	14) Explicit procedures for expert selection; Clear selection criteria; Clear selection criteria should be prespecified and may include the candidates' years of related experience, or tenure in a position that is relevant to the subject under study Report the response rate to the initial call for participation; provide detailed information about the participating experts (profile) to better allow judgments about their credibility [Paré 2013, page 210, Table 3] 15) Explain how groups were chosen. Consensus Development Panels: Panel composition: the panel should be made up of experts in the field; the publication should report on how they were chosen and why; [Waggoner 2016, page 665, 667] 16) Implied by mentioning that detailed information on participants was lacking in some reporting guidelines. Page 5 Report specialties of experts, names and institutions, the selection criteria [Wang 2015]
Methods 2.4 Does the study suggest anything of how or if PPI (public patient involvement) activity should be reported	No data
Methods 2.5 Does the study suggest anything about what or if consensus papers should report regarding panel recruitment strategies, invitations? Any level of detail specified?	1) The use of specific methods to encourage the experts to respond (e.g., stamped addressed envelope for returning the questionnaire and financial compensation) [page 2] and recommendation to report whether special techniques were used to invite participants [Boulkedid 2011, Appendix S1 item 21] 2) Criteria for the selection of experts and transparent information on recruitment of the expert panel, socio- demographic details including information on expertise regarding the topic in question, (non)response and response rates over the ongoing iterations should be reported" [Jünger 2017, Box 3, 9] 3) provide a detailed description of the expert recruitment and selection process [Paré 2013, page 215 first bullet on the right] 4) method of obtaining participants should be described [Waggoner 2016, page 667]
Methods	1) The method used to define a consensus among panel members; , whether the percentage of agreement was determined; Whether a cut-off (e.g., median value) was used to select indicators [page 2] Consensus definition at each

2.6 Does the study suggest how or if consensus papers should report the consensus criteria/threshold (or the level of agreement considered to reach consensus)?	round [page 7, Appendix item 28] how was consensus obtained [page 7, Appendix item 28] definition of consensus should be reported [Boulkedid 2011, table 5] 2) Clearly describe how consensus was defined [Chan 2019, Box 1] 3) Need to define criteria for consensus and to document the degree of agreement together with the results of the Delphi process. Should be defined a priori. [Diamond 2014, page 404 and table 6] 4) Was the agreement/consensus threshold predefined? [Gattrell 2019, table 1] 5) Box 2 Specific threshold for the chosen measure (e.g., median of at least 7 on a nine-point scale and an interquartile range of less than 2) [Grant 2018, p 97] 6) Determine the criteria and the meaning of `consensus' in relation to the studies [Hasson 2020, page 1013] 7) No. They do state that "articulating the definition of consensus used" was identified as "particularly problematic and were often left out or poorly described", and that "the most concerning issue we identified was that consensus was often not defined a priori. Only 43.2% of the articles we reviewed reported their definition of consensus at the start of the study." But they do not suggest how to report. [Humphrey-Murto 2017 AMA] 8) Clearly describe how consensus was defined [Humphrey-Murto 2017 Med Teach, page 18] 9) suggests definition of consensus should be reported [Humphrey-Murto 2019, table 1, also fig 1 and page 1044] 10) 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". Definition and 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". "If an a priori definition of consensus is not realistic due to the explorative nature of the study, it should be identified and established by the research team in the course of the process." [Jünger 2017, item 12]

	11) How was consensus defined and measured? What role did the stability of the answers play? [Niederberger 2020, Table 2] Whether and when consensus was defined in the Delphi studies. Was consensus defined a priori in advance of development of the questionnaire. [Niederberger 2020, Table 5] How was consensus measured, e.g. percentage agreement, units of central tendency (especially median) or a combination of percent agreement within a certain range and for a certain threshold. [Niederberger 2020, page 6] 12) NGT explain criteria used to determine how and when a consensus was met Consensus Development Panels: Explain what constituted consensus and how this was assessed. [Waggoner 2016, page 665] Delphi Explain what constituted consensus and how this was assessed. [Waggoner 2016, page 667] 13) The endpoint of consensus [Wang 2015, page 5]
Methods 2.7 Does the study suggest how or if consensus papers should report how decision of approval of an item will be made?	1) Whether the percentage of agreement was determined [page 2] We recorded the method used to define a consensus among panel members, whether the percentage of agreement was determined, and whether a cut-off (e.g., median value) was used to select [Boulkedid 2011, Appendix S1 item 16 (technique method)] 2) Reporting on each round separately illustrates clearly the array of themes generated in round one and gives an indication of the strength of support for each round. The presentations of findings are important and findings from subsequent rounds should be reported in a summarized format to indicate the relative standing of each of the opinions. [Hasson 2020, page 1013] 3) (Non)response and response rates over the ongoing iterations should be reported [Jünger 2017, item 9]
Methods 2.8 Does the study suggest anything about what level of detail should be reported regarding the number of Delphi rounds or if this should be reported?	1) Was the number of rounds to be performed stated (not how it should be reported, but implies it should be) [Banno 2019, page 2 under item 2] 2) Was the number of rounds to be performed stated? [Banno 2020, 3.4, table 3] 3) Describe the number of rounds planned [Chan 2019, Box 1]

	4) Specify a maximum number of rounds [page 404] what was the reason to stop the delphi [Diamond 2014, table 3] What criteria will be used to determine to stop the Delphi process or will the Delphi be run for a specific number of rounds only [Diamond 2014, table 6, table 1 item 2] 5) number and outline per round should be reported also page 1013 [Hasson 2020, fig 1] 6) Describe the number of rounds planned and/or criteria for terminating the process [Humphrey-Murto 2017 Med Teach, page 17] 7) Only implying that x number of rounds are necessary [Humphrey-Murto 2017 AMA] 8) The methods employed need to be comprehensible; information about the number and design of survey rounds, [Jünger 2017, Box 3 item 10] 9) Not specifically under item 4 in table 2 report of the specific process used? How many rounds were used in the Delphi technique [Niederberger 2020] 10) If a study goes beyond the agreed number of rounds (review suggests 2 rounds are required), this should be explained [Waggoner 2016, page 667]
Methods 2.9 Does the study suggest anything about what level of detail should be reported regarding the criteria used for defining the number of rounds? (why 2-3 or more e.g.) or if this should be reported?	1) Implied in Banno 2020 The prespecified criteria for stopping the Delphi process, other than a statement of the number of rounds, were clarified [Banno 2020] 2) Describe the number of rounds planned and criteria for terminating the process [Chan 2019, Box 1] 3) Describe the number of rounds planned and/or criteria for terminating the process [Humphrey-Murto 2017 Med Teach, page 17] 4) They, imply that the number of rounds is an important thing to report -- but they do not state this as a suggestion.[Humphrey-Murto 2017 AMA] 5) Will the number of rounds be decided a priori? If not determined a priori, what are the criteria for terminating the process? [Humphrey-Murto 2019, Fig 1]

	6) What was the rationale for the number of rounds; when was the number of rounds defined [Niederberger 2020, page 6] 7) Table 3 Report the stopping [Paré 2013] 8) For delphi: if a study goes beyond two rounds, explain reason for doing so; [Waggoner 2016, page 667]
Methods 2.10 Does the study suggest anything about the details that should be reported regarding the time between rounds, if this should be prespecified in advance, or if this should be reported?	1) The time taken to complete the Delphi procedure was recorded [Boulkedid 2011, page 2]
Methods 2.11 Does the study suggest anything about details that should be reported of the names of the techniques of non-Delphi methods used to gather participants' inputs and reach consensus ?	1) Whether the meeting was held before, after, or between Delphi rounds and what the participants did during the meeting [Boulkedid 2011, page 2]
Methods 2.12 Does the study suggest anything of what or in which detail	1) What software will be used to administer the Delphi? [Humphrey-Murto 2019, fig 1]

should be reported regarding tool or electronic system used for Delphi? (If Delphi was used)? Or if this should be reported?	
Methods 2.13 Does the study suggest anything about how or in what level of detail the anonymity of participants (in Delphi or other methods) has to be reported? Or if this should be reported?	1) No, only that it is a limitation of this study that the quality score did not include that. So actually they feel it should be reported how anonymity was maintained [Banno 2020] 2) Describe how anonymity was defined [Chan 2019, Box 1] 3) Were responses anonymized [Gattrell 2019, table 1] 4) It suggests that conducting anonymous iterative mail or e-mail questionnaire rounds is one of the steps [p 1491]. While the authors may have assumed that readers would understand that anonymity was part of their study design, we suggest that they state this, given the variability in approaches that have been labelled as modified consensus methods. [Humphrey-Murto 2017 AMA, page 1497] 5) Describe how anonymity was maintained. Authors must clearly state how this was accomplished. It is achieved through the use of mail outs in Delphi and RAND/UCLA and private ranking in NGT. [Humphrey-Murto 2017 Med Teach, page 18] 6) How will anonymity be maintained? [Humphrey-Murto 2019, fig 1] 7) Ensure the anonymity of the participants. The anonymity of the experts was reported in virtually all of the studies [Paré 2013]
Methods 2.14 Does the study suggest anything about how to report, and in what level of detail, the feedback for panellists (in	1) Whether the experts were informed of both the response of the group and their own individual response (individual feedback) to each item. The type of feedback, which was defined as qualitative when a summary of the panel's comments was sent to each participant and quantitative when simple statistical summaries illustrating the collective opinion (e.g., central tendency and variance) were sent to each participant [page 2] After each round, each participant should be given the panel results (median, lowest, and highest ratings), the participant's response, and a summary of all comments received. These data inform each participant of his or her position relative to the rest of the group, thus assisting in decisions about replies during future Delphi rounds. [Boulkedid 2011, page 8] It has been recommended that

Delphi rounds or other methods) process? Or if this should be reported?	feedback should include qualitative comments and statistical measures [citation 51, Murphy 1998]. More specifically, we determined whether the experts were informed of both the response of the group and their own individual response (individual feedback) to each item [Boulkedid 2011] 2) Describe the type of feedback provided after each round [Chan 2019, Box 1] 3) Were participants' responses in each round reported back to the group, and were responses anonymized? [Gattrell 2019, Table 1] 4) Give attention to issues which guide data collection: the discovery of opinions, the process of determining the most important issues referring to the design of the initial round, and the management of opinions [Hasson 2020, page 1013] 5) Was formal feedback provided? If so, was the feedback described? [page 1493], areas that need to be improved with reporting providing participants with feedback of group ratings [Humphrey-Murto 2017 AMA, page 1494] 6) Describe the type of feedback provided after each round [page 18]. Feedback to participants can include quantitative and/or qualitative data. It also involves two types of agreement: the extent to which individual participants agree with an issue, and the extent to which participants agree with one another. Quantitative feedback may include summary statistics such as the participants' score, participants' medians, range of scores and the proportion of participants selecting each point on a scale. Participants are provided an opportunity to change their ranking, but it should be made clear that they do not need to conform. Researchers may ask the participants who are outliers to provide written justification for their choices (qualitative data) [Humphrey-Murto 2017 Med Teach] 7) What type of feedback will participants received after each round? [2019] indicates feedback between rounds should include individuals' scores for each item and the distribution of votes by participant group. Some, however, preferred to view aggregated feedback as well as feedback to individual participants [Humphrey-Murto 2019 Yes page 1042, table 1] 8) How was the feedback designed? [Niederberger 2020, table 2] 9) Citation [Schmidt, 54] recommends three relevant pieces of feedback that can be provided to experts in phase 3 in addition to mean ranks, namely, the interpretation of Kendall's W from the previous round, the percentage of experts

	placing each item in the top half of their list and the relevant comments that were made by the other panellists [Paré 2013, page 213] 10) They imply that it should be reported that panellist feedback was collected to inform subsequent Delphi rounds [Resemann 2018] 11) not about reporting but they state "57 % were silent about how the feedback after consensus was dealt with." suggesting that they felt it needs to be reported. [page 2] only that some reporting guidelines described the feedback information requirement, or gave the methods for feedback collection [Wang 2015, page 6]
Methods 2.15 Does the study suggest anything about how or if data synthesis/analysis should be reported (from any consensus method used and how this was calculated statistically) and in what level of detail?	1) It is important that standards and norms for prospectively defining analysis plans are needed to improve the credibility of Delphi processes for informing health research, practice, and policy [Grant 2018, page 97] 2) The methods employed need to be comprehensible; information about methods of data analysis, processing and synthesis of experts' responses to inform the subsequent survey round [Box 3] {Jünger 2017] 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." [Jünger 2017, item 13] 3) Detailing statistical analyses and interpretation in arriving at final agreed values [Ng 2018, item 7] 4) The statistical analyses should be reported [Paré 2013, page 211] 5) Consensus Development Panels: Statistical analysis: must be reasonable for the research question, and should be as rigorous as possible [Waggoner 2016, page 665]
Methods 2.16 Does the study suggest anything about how or if piloting should be reported and in what	1) Pilot testing with a small group of individuals is suggested before implementation [Humphrey-Murto 2017 Med Teach, page 16] 2) 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. [Box 3] The methods employed need to be comprehensible; this includes information on preparatory steps (How was

level of detail (e.g. understanding of consensus items, platforms used, tools used)?	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 [Jünger 2017] 3) Pre-test task instructions and questionnaire instruments [Paré 2013]
Methods 2.17 Does the study suggest anything about how or if the role of Steering Committee members should be reported?	No data
Methods 2.18 Does the study suggest anything on what or if should be described regarding COI or funding?	1) 'Sources of funding (industry, non-industry)'as items associated with reporting quality [Banno 2019, page 2] 2) Is the funding source clearly disclosed? [table 1] Is the role of the funder clearly disclosed? [table 1] Is the funding of any external support (e.g. with the Delphi panel meeting/questionnaires, or medical writing support for the final manuscript) clearly disclosed? [Gattrell 2019] 3) "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" [Jünger 2017] 4) Describing expert panel selection with eligibility criteria and including conflicts of interest [Ng 2018]
Methods 2.19 Does the study suggest anything on what should be described of how is dealt with COI of panellist (not allowed	1) No. It only deals with COI as a planning/methodological procedure, not reporting. "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"[Jünger 2017]

to vote when there is COI)? Or if this should be described	

Results 3.1 Does the study suggest anything on how to report the initial evidence search (presentation of results of the literature review)?	1) No, but they suggest it should be reported [Jünger 2017]
Results 3.2 Does the study suggest anything on how to report n of studies found?	No data
Results 3.3 Does the study recommend which detail should be used when reporting panellists drop-outs (numbers and reasons)? Or if this should be reported?	1) No but it states that number the response rate for the first round dropped to 170 (66.1%). [page 1494]; areas that need improvement in reporting the number of participants after each round [page 1496] Other analyses of consensus methods research found similar poor reporting of this feature, with 7% to 39% of studies reporting response rates for all rounds of data collection [Humphrey-Murto 2017 AMA] 2) Fig 1 step 7 How will non-responders be managed, i.e. will they be excluded in subsequent rounds What response rate will be acceptable for each stakeholder group in each round? [Humphrey-Murto 2019] 3) 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 [Jünger 2017, Box 3] 4) Outlining participation and attrition rates for each round [Ng 2018]

	5) report the response rate to the initial request for participation, the size of the panel and the retention rate; [Paré 2013, page 215 3rd bullet]
Results 3.4 Does the study suggest how or if approval rates per item shared with respondents for each round should be reported in the Results section?	1) Response rate for each round [Boulkedid 2011, Table 5 on page 7] 2) Yes Box 1 report response rates and results after each round [Chan 2019] 3) Response rates for each round should be reported, presentation of total of issues generated in round 1, and presentation of results in round 2 indicating strength of support [Hasson 2000, figure 1 and page 1013] 4) Report response rates and results after each round [Humphrey-Murto 2017 Med Teach, page 18] 5) it should report response rates for all rounds [Humphrey-Murto 2019, page 1042] 6) 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." [Jünger 2017, item 13]Criteria for the selection of experts and transparent information on recruitment of the expert panel, socio- demographic details including information on expertise regarding the topic in question, (non)response and response rates over the ongoing iterations should be reported". [Jünger 2017] 7) Reporting both quantitative results and textual comments for each round of analysis [Ng 2018] 8) How high was the response rate from the experts both when initially approached and also for the individual rounds [Niederberger 2020, Table 2] 9) Level of consensus should be reported [Resemann 2018]
Results 3.5 Does the study suggest anything about in which detail the items that have been dropped should	1) Were the criteria for dropping clear; are stopping criteria, other than rounds, reported [Banno 2019, item 3 and 4] 2) Were the criteria for dropping items clear? (yes, no, or not applicable) [Banno 2020, 2.6 item 3] 3) Clear criteria for dropping or combining items should also be specified based on the level of agreement or disagreement with individual items. One of the limitations of a priori specification is that certain items may fall just below the

be reported? (reasons e.g.) Or if this should be reported?	threshold for what is fundamentally an arbitrary cut off. In the event that items, believed to be important fell just below the threshold for inclusion in the study, the authors could consider including these items as posteriori considerations provided that sufficient justification was provided. [page 405] Suggested quality criteria: Were criteria for dropping items clear; Stopping criteria other than rounds specified? [Table 5] Were items dropped? What criteria will be used to determine which items to drop? [Diamond 2014, Table 6] 4) No, but they state 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 and Definition and 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 [Jünger 2017 in Box 3] 5) Were criteria defined for dropping items [Niederberger 2020, page 6]
Results 3.6 Does the study make any recommendation on how to report the collection, synthesis and use of comments from panellists? Or if this should be reported?	1) It has been recommended that feedback should include qualitative comments and statistical measures [Murphy 1998, 51]. After each round, each participant should be given the panel results (median, lowest, and highest ratings), the participant's response, and a summary of all comments received [Boulkedid 2011] 2) Describe the type of feedback provided after each round. Quantitative feedback may include summary statistics such as the participants' score, participants' medians, range of scores and the proportion of participants selecting each point on a scale. Participants are provided an opportunity to change their ranking, but it should be made clear that they do not need to conform [Humphrey-Murto 2017 Med Teach] 3) 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 [Jünger 2017, item 13] 4) Ask experts to justify their rankings. Have experts comment and validate consolidated list [page 210 Table 3]. Did experts consolidate the list of items; Did experts comment on and validate the list of items; Was the final number of items reported. Report whether panel members had the opportunity to justify or clarify their own reasoning and to comment on the responses of the other experts as well as on the progress of the panel as a whole. [Paré 2013, page 213].

	Were panellists able to revise previous statements [Paré 2013] 5) No, but implied that it should be: did not report collecting panellist feedback to inform subsequent Delphi stages [Resemann 2018]
Results 3.7 Does the study suggest regarding how the final list of items (for clinical guideline or reporting guideline) should be reported? Or if this should be reported?	1) Partially. It says it should be detailed and disseminated, but it does not suggest how (in what format) it should be reported [Jünger 2017] 2) Suggests "detailing statistical analyses and interpretation in arriving at final agreed values" [Ng 2018] 3) Report final number of items [Paré 2013, page 210 Table 3] 4) No but again imply "reported the number of statements assessed." [Resemann 2018]

Discussion 4.1 Does the paper suggest anything about reporting the limitations and strengths of the study and how? Or if this should be reported?	1) Address potential methodological issues (e.g lack of consensus) or limitations in the discussion (e.g. low response rate) [Chan 2019, Box 1] 2) Interpretation of consensus gained/not gained [Hasson 2020, page 1009] 3) In the discussion the authors should address issues that may have impacted the results such as poor response rates between rounds, lack of participation from a select group or geographic region, or lack of consensus. [Humphrey-Murto 2017 Med Teach, page 18] 4) Methodological issues should be reported [Humphrey-Murto 2019, figure 1] 5) Reporting should include a critical reflection of potential limitations and their impact of the resulting guidance". [Jünger 2017]
Discussion 4.2 Does the paper suggest anything about what or in	1) Page 5: is considered a good measure if it meets criteria including reliability, sensitivity, specificity, and feasibility (or applicability) [20,31]. The common use of these characteristics can facilitate acceptance and implementation of indicators developed [Boulkedid 2011]

which detail the applicability generalisability, and reproducibility of the study should be reported? Or if this should be reported?	2) The conclusions should adequately reflect the outcomes of the Delphi study with a view to the scope and applicability of the resulting practice guidance. [Jünger 2017, item 15] 3) It is also necessary to discuss the critical and rationalistic criteria for the validity and reliability of the studies and the more constructivist characteristics of credibility, transparency, and transferability. [Niederberger 2020, page 8]

5.1 Any other item proposed by the paper that is not captured in other columns?	1) Were criteria for dropping items clear? Are stopping criteria, other than rounds, specified [Banno 2019] 2) Differences between the protocol and the article [Banno 2020, 2.9] 3) Geographic scope of the survey [page 2]. Main methods used to send the questionnaires (e.g., mail, E-mail, or fax). [Boulkedid 2011, page 7] The formulation of the questionnaire items (e.g., open questions, rating of quality indicators, or both). [Boulkedid 2011] Whether the quality indicators were rated (in which case, we recorded the minimum and maximum values on the rating scale). [Boulkedid 2011] A flow chart of quality indicators (figure showing the output and input indicators at each round) and/or for a written description of indicator flow. [Boulkedid 2011, page 3] Quality indicators used in the first round versus the end of the last round. [Boulkedid 2011, page 3] Availability of the questionnaires in the article itself or in an appendix [Boulkedid 2011, page 3] Whether selection criteria changed between rounds [Boulkedid 2011, page 5] Whether panelists were able to make comments. [Boulkedid 2011, page 6] Whether there was a meeting; at what stage it took place and how people participated [Boulkedid 2011] Response rate for each round [Boulkedid 2011, page 7] preparation in advance of starting Delphi (outcome indicators, structure indicators, process indicators) [Boulkedid 2011, In appendix S1, item 1] METHODS We evaluated the relationship between the response rate and the use of specific methods to encourage the experts to respond (e.g., stamped addressed envelope for returning the questionnaire and financial compensation). Also on maybe we should add item regarding encouragement of participants [Boulkedid 2011, page 2, page 5 right column] Geographic scope of Delphi consensus procedure [Boulkedid 2011,item 20 of appendix and table 5] Question format (open questions, rating scale?) Also in table 5 how were questions formulated? [Boulkedid 2011, item 24

	appendix] Rating scale [Boulkedid 2011, item 25] Methods used to send questionnaire (email fax, mail) [Boulkedid 2011, table 5] Time to complete questionnaire reporting of differences in response rate in rounds [Boulkedid 2011] Number of rounds necessary to reach consensus [Boulkedid 2011] Duration of the procedure [Boulkedid 2011] Is questionnaire added as appendix? [Boulkedid 2011] For Discussion: Validity [Boulkedid 2011] 4) Outline each step of the process. If modifications were made, provide a rationale for your choices. [Chan 2019] Describe the selection and preparation of the scientific evidence for the participants. [Chan 2019] Include a description of the facilitator's credentials. [Chan 2019] What background material was provided to participants. [Chan 2019] What formal feedback of group rating was shared between rounds [Chan 2019] 5) Specify stopping criteria in the absence of consensus [Diamond 2014] 6) Were the questions formulated or validated by an expert panellist [Gattrell 2019] 7) Researchers conducting consensus-oriented Delphi processes should prospectively and completely register the intended procedure for identifying which items reach consensus. [Grant 2018] The analysis procedure for determining consensus for Delphi processes should be chosen a priori ideally before starting the first round but at the very latest before completing data collection to improve the validity of findings. [Grant 2018] Health researchers conducting consensus-oriented Delphi processes should commit themselves in advance to an analytic procedure for determining which items reach consensus before they see the actual data (or, ideally, before they even collect the data). [Grant 2018] Registrations should be in a publicly available and independently controlled platform that time-stamps entries [Grant 2018] 8) "Copy of each round questionnaire illustrated" [Hasson 2020] statistical interpretation for the reader [Hasson 2020] appendices to include the questionnaires [Hasson 2020] For Discussion interpretations of consensus gained/not gained reliability and validity [Hasson 2020]

9) *Page 1493(2) Was background information provided to the participants? pg 1496 areas appeared particularly problematic and were often left out or poorly described: providing background information to participants
AND so a clear description of what information was provided and in what format is important
* (3) Was the consensus method used for item generation, ranking, or both?
* (11) Was consensus forced?
Was mail/e-mail polling or face-to-face questioning used? [Humphrey-Murto 2017 AMA]

10) Outline each step of the process: if modifications were made, provide a rationale for the choices made. Providing justification for the choices made will also add credibility. [Humphrey-Murto 2017 Med Teach]

11) Background provided to participants, what is level of detail provided [Humphrey-Murto 2019]
Figure 1 clear outline of the overall process involved and where Delphi fits [Humphrey-Murto 2019, figure 1]
How sample size is determined of participants [Humphrey-Murto 2019, figure 1]

12) Any modifications should be justified by a rationale and be applied systematically and rigorously [Jünger 2017, Box 3]
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 [Jünger 2017]
It is recommended to have the final draft of the resulting guidance on best practice in palliative care reviewed and approved by an external board or authority before publication and dissemination [Jünger 2017, Box 3]
information about methodological decisions taken by the research team throughout the process Jünger 2017, Box 3]
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 [Jünger 2017, Box 3]
Publication and dissemination [Jünger 2017, Box 3]

13) Item 2-4 and 9 appending revised questionnaires [Ng 2018]

14) Specific definition of underlying Delphi technique (or as I thought it is important to define exactly what method is used, especially if a modified method is used this needs to be very clear [Niederberger 2020]
What role did the stability of the answers play? [Niederberger 2020, table 2]
Questionnaire and scale development How were the questionnaires and the specific items for a Delphi technique

	developed? [Niederberger 2020] Nevertheless, it is important to precisely describe, justify, and methodologically reflect on any modifications [Niederberger 2020] How were the questionnaires and the specific items for a Delphi technique developed? [Niederberger 2020, Table 2] Were items identified from empirical analyses such as qualitative interviews or focus groups that were completed in advance or were taken from existing guidelines. [Niederberger 2020, Complementary AND page 6 Was the first (qualitative) round of questions in the Delphi process used to generate the items for a standardized questionnaire. [Niederberger 2020, Complementary AND page 6] 15) Was the final number of items reported [Paré 2013, Table 3] Were items randomly ordered [Paré 2013, Table 3] 16) Describe the rating scales used [Resemann 2018] the number of statements assessed should be reported [Resemann 2018] 17) For nominal group process, the research question used to prompt the panel must be clear and concise to obtain valid suggestions from panel members. [Waggoner 2016, page 665] The heterogeneity should be reported [Waggoner 2016, page 665] Evaluation of reliability [Waggoner 2016, page 665] 18) Meeting attendance; format (e.g. face-to-face); agenda preparation; materials sent to participants prior to meeting; duration of meeting [Wang 2015, page 5] Flow diagram [Wang 2015, page 3] Should we add something regarding other consensus methods including an item regarding face to face meetings? [Wang 2015, page 5]
5.2 Any other item not proposed by the paper, but you think that could be added (not fitting the categories above)?	1) Are stopping criteria, other than rounds, specified? [Banno 2019, page 2] 2) Information letter explaining the method and the reasons their participation to the whole process would be necessary, as well as a form for collecting their consent to complete the entire Delphi process. [Boulkedid 2011] 3) "Round 1: presentation of total number of issues generated" [Hasson 2020] 4) This paper was "pointing fingers", showing what was wrong, without suggesting solutions. However, we can be inspired by the critics to build the following list of items: 1) Purpose of the consensus study Whether a literature review was done to support the selection of items [Humphrey-Murto 2017 AMA] 5) Length of the background provided [Humphrey-Murto 2019]

	Purpose of study: outcome/diagnosis/intervention? [Humphrey-Murto 2019]

Examples of text with well reported methods/results (for E&E document) - write NA if none was cited or found by you	1) Page 7 Table 5 [Boulkedid 2011] 2) Box 1 [Chan 2019] 3) Might have a look at table 6 [Diamond 2014] 4) Table 1 [Gattrell 2019] 5) Parts of Fig 1 and checklist page 1013 [Hasson 2020] 6) Table 1 lists "exemplary publications" for nominal group process, consensus development panel and Delphi technique Page 667 references studies that were "Very descriptive" of the statistical techniques used. [Waggoner 2016]
Additional comments from assessor	1) Limited value; protocol for Banno 2020 [Banno 2019] 2) Of limited use. The authors developed a 4-point quality score that they applied to Delphi publications [Banno 2020] 3) Excellent resource [Boulkedid 2011] 4) Focusses on defining consensus [Diamond 2014] 5) Congress poster only [Gattrell 2019] 6) Study used RAND's ExpertLens as the Delphi platform [Grant 2018] 7) 1497: The lack of consensus on consensus methods makes it imperative that researchers provide clear and detailed reporting of the methods they used and that they justify these choices. [Humphrey-Murto 2017]

	8) Page 1044 A suggestion to improv uniformity is to use a software program that provides structure and help with reporting all relevant outcomes (e.g. DelphiManager, http://comet-initiative.org/delphimanager/) [Humphrey-Murto 2019] 9) Very informative [Jünger 2017] 10) The study focusses on information systems. Arguably, this is not within the inclusion criteria for the search [Paré 2013] 11) Review covers nominal group process, consensus development panel and Delphi technique [Waggoner 2016] 12) Study looked at the reporting quality of reporting guidelines [Wang 2015]

ACcurate COnsensus Reporting Document (ACCORD): Summary of extracted data from literature search

Table of Contents

van Zuuren EJ, *et al. BMJ Open* 2022; 12:e065154. doi: 10.1136/bmjopen-2022-065154

Section: Background

1. Background

Data extraction question	Articles	Checklist item(s) with brief explanation
1.1. Does the study suggest anything about how or if consensus papers should report the context or rationale for choosing a consensus method over other methods?	Hasson F, *et al. J Adv Nurs* 2000[1] Humphrey-Murto S, *et al*. *Med Teach* 2017[2] Humphrey-Murto S, *et al. J Rheumatol* 2019[3] Jünger S, *et al*. *Palliat Med* 2017[4]	State the rationale for use of consensus method over other options. *Should consider other consensus methods as well as other methodology types.*
1.2. Does the study suggest anything about how/what or if consensus papers should report the objectives of the consensus exercise?	Hasson F, *et al. J Adv Nurs* 2000[1] Humphrey-Murto S, *et al*. *Med Teach* 2017[2] Jünger S, *et al*. *Palliat Med* 2017[4] Boulkedid R, *et al. PLoS One* 2011[5] Chan TM, *et al*. *CJEM* 2019[6] Diamond IR, *et al*. *J Clin Epidemiol* 2014[7]	Clearly define study objectives. *Could include presentation of group consensus, or just to quantify the level of agreement.*

Section: Methods

2. Methods

Data extraction question	Articles	Checklist item(s) with brief explanation
2.1. Does the study suggest anything about how/what or if consensus papers should report regarding: A literature search/strategy?	Hasson F, *et al. J Adv Nurs* 2000[1] Humphrey-Murto S, *et al*. *Med Teach* 2017[2] Chan TM, *et al*. *CJEM* 2019[6] Humphrey-Murto S, *et al*. *Acad Med* 2017[8] Resemann HK, *et al*. *Curr Med Res Opin* 2018[9]	A) Describe the strategy for reviewing the existing scientific evidence that informed the study. *If no existing literature is available, the extent of the search should be described.* B) Describe how existing scientific evidence will be provided to the participants. *If different participant groups are involved, it should be stated which information will be provided to which group.*
2.2. Does the study suggest anything about how/what or if consensus papers should report regarding: Inclusion and exclusion criteria for the literature search?	Boulkedid R, *et al. PLoS One* 2011[5] Chan TM, *et al*. *CJEM* 2019[6] Paré G, *et al. Inf Manag* 2013[10]	Describe the process of the literature search. *Should include inclusion and exclusion criteria, and state whether these were prespecified.*
2.3. Does the study suggest anything of what or if consensus report should report on panel composition, n of participants, expertise, origin? Prespecified?	Hasson F, *et al. J Adv Nurs* 2000[1] Humphrey-Murto S, *et al*. *Med Teach* 2017[2] Humphrey-Murto S, *et al. J Rheumatol* 2019[3] Jünger S, *et al*. *Palliat Med* 2017[4] Boulkedid R, *et al. PLoS One* 2011[5] Chan TM, *et al*. *CJEM* 2019[6] Diamond IR, *et al*. *J Clin Epidemiol* 2014[7] Humphrey-Murto S, *et al*. *Acad Med* 2017[8] Paré G, *et al. Inf Manag* 2013[10] Banno M, *et al*. *J Clin Epidemiol* 2019[11] Banno M, *et al*. *J Clin Epidemiol* 2020[12] Gattrell WT, *et al*. *Curr Med Res Opin* 2019[13]	A) Describe the structure of the study's participants. *Should describe inclusion of a Chair/Co-chairs, steering committee, and subgroups, if applicable.* B) Explain how panel participants were selected. *Should state who was responsible for panellist selection, the selection criteria applied, the justification for choosing panellist numbers and selection criteria, and whether criteria were prespecified.*

Section: Methods

Data extraction question	Articles	Checklist item(s) with brief explanation
	Ng J. *Value Health* 2018[14] Niederberger M, *et al. Front Public Health* 2020[15] Waggoner J, *et al*. *Acad Med* 2016[16] Wang X, *et al. BMC Med Res Methodol* 2015[17]	*C)* Describe the composition of the panel. *Should include number of participants at all stages of the process, sociodemographics (e.g. age, sex, specialty, type and duration of relevant experience). Should also describe panel subgroups, if relevant.* D) Describe the expertise of the panel. *Should include the definition of "expert" and description of any public or patients involved.* E) Describe the facilitator(s), if used. *Should include type and duration of relevant experience, and the role played in the process.*
2.4. Does the study suggest anything of how or if PPI (public patient involvement) activity should be reported	No data	Describe the role and involvement of any public or patients. *Should detail the stage(s) at which they were involved, and their roles and contributions.*
2.5. Does the study suggest anything about what or if consensus papers should report regarding panel recruitment strategies, invitations? Any level of detail specified?	Jünger S, *et al*. *Palliat Med* 2017[4] Boulkedid R, *et al. PLoS One* 2011[5] Paré G, *et al. Inf Manag* 2013[10] Waggoner J, *et al*. *Acad Med* 2016[16]	Describe how the panel members were recruited. *Could include communication/advertisement method(s) and locations.*
2.6. Does the study suggest how or if consensus papers should report the consensus criteria/threshold (or the level of agreement considered to reach consensus)?	Hasson F, *et al. J Adv Nurs* 2000[1] Humphrey-Murto S, *et al*. *Med Teach* 2017[2] Humphrey-Murto S, *et al. J Rheumatol* 2019[3] Jünger S, *et al*. *Palliat Med* 2017[4] Boulkedid R, *et al. PLoS One* 2011[5] Chan TM, *et al*. *CJEM* 2019[6] Diamond IR, *et al*. *J Clin Epidemiol* 2014[7] Humphrey-Murto S, *et al*. *Acad Med* 2017[8] Gattrell WT, *et al*. *Curr Med Res Opin* 2019[13]	A) Define the consensus measure to be used. *Could include percentage agreement, units of central tendency (e.g. median), a categorical rating (e.g. Agree/Strongly agree) or a combination of percent agreement within a certain range.* B) State the threshold for the group achieving consensus. *Should include whether the threshold was*

Section: Methods

Data extraction question	Articles	Checklist item(s) with brief explanation
	Niederberger M, *et al. Front Public Health* 2020[15] Waggoner J, *et al*. *Acad Med* 2016[16] Wang X, *et al. BMC Med Res Methodol* 2015[17] Grant S, *et al*. *J Clin Epidemiol* 2018[18]	*pre-defined and highlight any threshold variation between rounds, with explanation for the change. If the intention is to quantify the degree of consensus but not to use consensus as a stop criterion for the study, this should be stated.*
2.7. Does the study suggest how or if consensus papers should report how decision of approval of an item will be made?	Hasson F, *et al. J Adv Nurs* 2000[1] Jünger S, *et al*. *Palliat Med* 2017[4] Boulkedid R, *et al. PLoS One* 2011[5]	Explain how final consensus was reached. *Should describe the evolution of themes between voting rounds, if applicable.*
2.8. Does the study suggest anything about what level of detail should be reported regarding the number of Delphi rounds or if this should be reported?	Hasson F, *et al. J Adv Nurs* 2000[1] Humphrey-Murto S, *et al*. *Med Teach* 2017[2] Jünger S, *et al*. *Palliat Med* 2017[4] Chan TM, *et al*. *CJEM* 2019[6] Diamond IR, *et al*. *J Clin Epidemiol* 2014[7] Humphrey-Murto S, *et al*. *Acad Med* 2017[8] Banno M, *et al*. *J Clin Epidemiol* 2019[11] Banno M, *et al*. *J Clin Epidemiol* 2020[12] Niederberger M, *et al. Front Public Health* 2020[15] Waggoner J, *et al*. *Acad Med* 2016[16]	State how many voting rounds were conducted. *Should include whether the number of rounds was prespecified, and whether this was an absolute or a maximum. If the maximum was exceeded, should explain the reasoning for doing so.*
2.9. Does the study suggest anything about what level of detail should be reported regarding the criteria used for defining the number of rounds? (why 2-3 or more e.g.) or if this should be reported?	Humphrey-Murto S, *et al*. *Med Teach* 2017[2] Humphrey-Murto S, *et al. J Rheumatol* 2019[3] Chan TM, *et al*. *CJEM* 2019[6] Humphrey-Murto S, *et al*. *Acad Med* 2017[8] Paré G, *et al. Inf Manag* 2013[10] Banno M, *et al*. *J Clin Epidemiol* 2020[12] Niederberger M, *et al. Front Public Health* 2020[15] Waggoner J, *et al*. *Acad Med* 2016[16]	Explain the rationale for choosing the number of voting rounds. *Should also describe the stop criteria, if used, and whether these were prespecified.*
2.10. Does the study suggest anything about the details that should be reported regarding the time between rounds, if this should be	Boulkedid R, *et al. PLoS One* 2011[5]	Describe the time period between voting rounds. *Should include whether the period was*

Section: Methods

Data extraction question	Articles	Checklist item(s) with brief explanation
prespecified in advance, or if this should be reported?		*prespecified and highlight differences between inter-round periods, if applicable.*
2.11. Does the study suggest anything about details that should be reported of the names of the techniques of non-Delphi methods used to gather participants' inputs and reach consensus?	Boulkedid R, *et al. PLoS One* 2011[5] Waggoner J, *et al*. *Acad Med* 2016[16]	Describe any additional methods used alongside the consensus process. *Should include all that were used, e.g. a self-administered questionnaire combined with a group meeting. Should also explain how the consensus process fitted into the overall study methodology.*
2.12. Does the study suggest anything of what or in which detail should be reported regarding tool or electronic system used for Delphi? (If Delphi was used)? Or if this should be reported?	Humphrey-Murto S, *et al. J Rheumatol* 2019[3]	Describe any tools used to administer the voting. *Could detail electronic platforms, if used.*
2.13. Does the study suggest anything about how or in what level of detail the anonymity of participants (in Delphi or other methods) has to be reported? Or if this should be reported?	Humphrey-Murto S, *et al*. *Med Teach* 2017[2] Humphrey-Murto S, *et al. J Rheumatol* 2019[3] Chan TM, *et al*. *CJEM* 2019[6] Humphrey-Murto S, *et al*. *Acad Med* 2017[8] Paré G, *et al. Inf Manag* 2013[10] Banno M, *et al*. *J Clin Epidemiol* 2020[12] Gattrell WT, *et al*. *Curr Med Res Opin* 2019[13]	Detail how anonymity of voters was maintained. *Could involve use of mail-outs in a standard Delphi procedure, blinding on an electronic platform, or private ranking in the NGT.*
2.14. Does the study suggest anything about how to report, and in what level of detail, the feedback for panellists (in Delphi rounds or other methods) process? Or if this should be reported?	Hasson F, *et al. J Adv Nurs* 2000[1] Humphrey-Murto S, *et al*. *Med Teach* 2017[2] Humphrey-Murto S, *et al. J Rheumatol* 2019[3] Boulkedid R, *et al. PLoS One* 2011[5] Chan TM, *et al*. *CJEM* 2019[6] Humphrey-Murto S, *et al*. *Acad Med* 2017[8] Resemann HK, *et al*. *Curr Med Res Opin* 2018[9] Paré G, *et al. Inf Manag* 2013[10] Gattrell WT, *et al*. *Curr Med Res Opin* 2019[13] Niederberger M, *et al. Front Public Health* 2020[15]	Explain how voting feedback was provided to panellists at the end of each round. *Could include summaries of group voting and/or their own individual responses. Should state whether feedback will be quantitative and/or qualitative, and whether it will be anonymised. If no feedback was provided, this should be stated.*

Section: Methods

Data extraction question	Articles	Checklist item(s) with brief explanation
	Wang X, *et al. BMC Med Res Methodol* 2015[17]	
2.15. Does the study suggest anything about how or if data synthesis/analysis should be reported (from any consensus method used and how this was calculated statistically) and in what level of detail?	Jünger S, *et al*. *Palliat Med* 2017[4] Paré G, *et al. Inf Manag* 2013[10] Ng J. *Value Health* 2018[14] Waggoner J, *et al*. *Acad Med* 2016[16] Grant S, *et al*. *J Clin Epidemiol* 2018[18]	Detail methods used to process responses after each voting round. *Could include statistical analysis methods, if used.*
2.16. Does the study suggest anything about how or if piloting should be reported and in what level of detail (e.g. understanding of consensus items, platforms used, tools used)?	Humphrey-Murto S, *et al*. *Med Teach* 2017[2] Jünger S, *et al*. *Palliat Med* 2017[4] Paré G, *et al. Inf Manag* 2013[10]	Describe any piloting of the study materials and/or survey instruments. *Should include the number of individuals in the pilot group and the rationale for their selection. Should also explain any changes made as a result of the pilot. If no pilot was conducted, this should be stated.*
2.17. Does the study suggest anything about how or if the role of Steering Committee members should be reported?	No data	Describe the role(s) of the Steering Committee in the process. *Should also detail the involvement of the Chair/Co-chairs, subgroups, or individual members at relevant stages of the process, if different from the group as a whole.*
2.18. Does the study suggest anything on what or if should be described regarding COI or funding?	Jünger S, *et al*. *Palliat Med* 2017[4] Banno M, *et al*. *J Clin Epidemiol* 2019[11] Gattrell WT, *et al*. *Curr Med Res Opin* 2019[13] Ng J. *Value Health* 2018[14]	A) Disclose any COI of the panellists *Should specify COI of each participant in the panel.* B) Disclose any funding received and the role of the funder. *Should specify the role of the funding source(s), e.g. involvement in the study concept/design, participation of the Steering Committee, for conducting the consensus process, medical writing support for its reporting.*

Section: Methods

Data extraction question	Articles	Checklist item(s) with brief explanation
2.19. Does the study suggest anything on what should be described of how is dealt with COI of panellist (not allowed to vote when there is COI)? Or if this should be described	Jünger S, *et al*. *Palliat Med* 2017[4]	Describe measures taken to avoid influence by any conflicts of interest (COI). *Should include disclosure of COI and how this was accounted for in the methodology, e.g. by limiting voting in case of a specific COI, adjudication by an independent researcher.*

Section: Results

3. Results

Data extraction question	Articles	Checklist item(s) with brief explanation
3.1. Does the study suggest anything on how to report the initial evidence search (presentation of results of the literature review)?	Jünger S, *et al*. *Palliat Med* 2017[4]	Describe how existing scientific evidence was provided to the participants. *Should include relevant specifics of the literature search, e.g. n of studies reported, to provide relevant context for the results. If different participant groups were involved, it should be stated which information was provided to which group.*
3.2. Does the study suggest anything on how to report n of studies found?	No data	Describe the results of the search and number of included studies.
3.3. Does the study recommend which detail should be used when reporting panellists drop-outs (numbers and reasons)? Or if this should be reported?	Humphrey-Murto S, *et al. J Rheumatol* 2019[3] Jünger S, *et al*. *Palliat Med* 2017[4] Humphrey-Murto S, *et al*. *Acad Med* 2017[8] Paré G, *et al. Inf Manag* 2013[10] Ng J. *Value Health* 2018[14]	A) State the response rates for each voting round. *Should specify n as well as percent, or otherwise indicate attrition/retention rates.* B) State the reasons cited for voter drop-outs at each stage of the process. *Could be provided as an aggregated summary or as individual responses. If this information was not collected, this should be stated.* C) Describe measures undertaken to maintain acceptable response rates. *If threshold rates differ between stakeholder groups, these should be described with explanation.*

Section: Results

Data extraction question	Articles	Checklist item(s) with brief explanation
3.4. Does the study suggest how or if approval rates per item shared with respondents for each round should be reported in the Results section?	Hasson F, *et al. J Adv Nurs* 2000[1] Humphrey-Murto S, *et al*. *Med Teach* 2017[2] Humphrey-Murto S, *et al. J Rheumatol* 2019[3] Jünger S, *et al*. *Palliat Med* 2017[4] Boulkedid R, *et al. PLoS One* 2011[5] Chan TM, *et al*. *CJEM* 2019[6] Resemann HK, *et al*. *Curr Med Res Opin* 2018[9] Ng J. *Value Health* 2018[14] Niederberger M, *et al. Front Public Health* 2020[15]	Describe which results that were shared with respondents after each voting round were reported in the final manuscript. *Could include response rates, the type of information presented, summaries of group voting and/or individual responses. If this information is not provided, this should be stated together with the rationale.*
3.5. Does the study suggest anything about in which detail the items that have been dropped should be reported? (reasons e.g.) Or if this should be reported?	Jünger S, *et al*. *Palliat Med* 2017[4] Diamond IR, *et al*. *J Clin Epidemiol* 2014[7] Banno M, *et al*. *J Clin Epidemiol* 2019[11] Banno M, *et al*. *J Clin Epidemiol* 2020[12] Niederberger M, *et al. Front Public Health* 2020[15]	A) List any voting items that were dropped. B) Explain the rationale for dropping any voting items. *Should state whether the criteria for dropping any items were prespecified.*
3.6. Does the study make any recommendation on how to report the collection, synthesis and use of comments from panellists? Or if this should be reported?	Humphrey-Murto S, *et al*. *Med Teach* 2017[2] Jünger S, *et al*. *Palliat Med* 2017[4] Boulkedid R, *et al. PLoS One* 2011[5] Resemann HK, *et al*. *Curr Med Res Opin* 2018[9] Paré G, *et al. Inf Manag* 2013[10]	Describe how responses were processed prior to reporting. *Should describe methods by which responses were analysed, aggregated or summarised, include whether any statements were revised between voting rounds, and state by whom the information was processed.*
3.7. Does the study suggest regarding how the final list of items (for clinical guideline or reporting guideline) should be reported? Or if this should be reported?	Jünger S, *et al*. *Palliat Med* 2017[4] Resemann HK, *et al*. *Curr Med Res Opin* 2018[9] Paré G, *et al. Inf Manag* 2013[10] Ng J. *Value Health* 2018[14]	Report the final outcomes. *Could be quantitative (e.g. summary statistics, score means, medians and/or ranges) and/or qualitative (e.g. aggregated themes from comments). Should be clear, accurately represent the consensus methodology used, and relevant to the field.*

Section: Discussion

4. Discussion

Data extraction question	Articles	Checklist item(s) with brief explanation
4.1. Does the paper suggest anything about reporting the limitations and strengths of the study and how? Or if this should be reported?	Hasson F, *et al. J Adv Nurs* 2000[1] Humphrey-Murto S, *et al*. *Med Teach* 2017[2] Humphrey-Murto S, *et al. J Rheumatol* 2019[3] Jünger S, *et al*. *Palliat Med* 2017[4] Chan TM, *et al*. *CJEM* 2019[6]	Discuss the study's methodological strengths and limitations. *Should address issues that may impact results, e.g. response rates or representation.*
4.2. Does the paper suggest anything about what or in which detail the applicability generalisability, and reproducibility of the study should be reported? Or if this should be reported?	Jünger S, *et al*. *Palliat Med* 2017[4] Boulkedid R, *et al. PLoS One* 2011[5] Niederberger M, *et al. Front Public Health* 2020[15]	A) Discuss the reliability of the study. B) Discuss the sensitivity of the study. C) Discuss the specificity of the study. D) Discuss the applicability of the study. E) Discuss the validity of the study.

Section: Additional topics

5. Additional topics

Data extraction question: Any other item proposed by the paper that is not captured in previous sections?

Articles	Checklist item(s) with brief explanation
Humphrey-Murto S, *et al*. *Med Teach* 2017[3] Jünger S, *et al*. *Palliat Med* 2017[4] Boulkedid R, *et al. PLoS One* 2011[5] Paré G, *et al. Inf Manag* 2013[10] Banno M, *et al*. *J Clin Epidemiol* 2020[12]	Explain any deviations from the planned protocol. *Should include any affected stages, including but not limited to change in panel number or composition, number of voting rounds, stopping criteria, statistical plan, reporting of outcomes.*
Boulkedid R, *et al. PLoS One* 2011[5] Resemann HK, *et al*. *Curr Med Res Opin* 2018[9]	Describe the formulation of questions. *Should include the type of questions, e.g. open questions, numerical rating, level of agreement rating. If rating questions were used, the scale range should be stated, and whether respondents were able to leave additional comments after rating items.*
Boulkedid R, *et al. PLoS One* 2011[5] Wang X, *et al. BMC Med Res Methodol* 2015[17]	Describe any group meetings that were held. *Should state at what stage the meeting took place, objectives/purpose, format (e.g. face-to-face or virtual), pre-read materials shared, attendance, location, duration, and how individuals participated.*
Hasson F, *et al. J Adv Nurs* 2000[1] Boulkedid R, *et al. PLoS One* 2011[5] Ng J. *Value Health* 2018[14]	List any items included in the appendix accompanying the main report. *Could include e.g. full voting questions from each round with response rates, or information provided to the panel as pre-reads or to summarise voting rounds.*
Boulkedid R, *et al. PLoS One* 2011[5]	State how the survey was presented to participants. *For example, as hard copy or via digital platform; could include description of email or mailing process. Should describe any randomisation procedures for questions, if used. If questions were not randomised, this should be stated.*
Boulkedid R, *et al. PLoS One* 2011[5]	Describe incentives for encouraging responses. *Should list any specific methods, e.g. paid return postage for the questionnaire or financial compensation.*
Boulkedid R, *et al. PLoS One* 2011[5]	State the period in which the process was conducted.
Grant S, *et al*. *J Clin Epidemiol* 2018[18]	Describe any prospective registrations for the consensus process.

Section: Additional topics

Articles	Checklist item(s) with brief explanation
	Should include the platform on which it was registered and a link, if applicable. If the process was not registered, this should be stated.
Jünger S, *et al*. *Palliat Med* 2017[4]	Describe any external peer review prior to publication. *Should name the authority, state the rationale for their review, and describe any modifications made as a result of their review.*
Humphrey-Murto S, *et al*. *Med Teach* 2017[3] Jünger S, *et al*. *Palliat Med* 2017[4]	Describe the overall process using a flow chart or diagram.
Paré G, *et al. Inf Manag* 2013[10] Niederberger M, *et al. Front Public Health* 2020[15]	Explain how the initial voting items in the consensus were developed. *Could describe e.g. development from empirical analyses, qualitative interviews, advance focus groups, brainstorming, or existing guidelines. Should state who consolidated the information and developed the voting items.*
Boulkedid R, *et al. PLoS One* 2011[5]	Describe the procedure for collecting participants' consent to complete the full consensus process. *Could briefly describe any forms used and how the data were collected and stored.*

Section: References

References

1. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. *J Adv Nurs* 2000;32:1008-15. doi: 10.1046/j.1365-2648.2000.t01-1-01567.x.
2. Humphrey-Murto S, Varpio L, Gonsalves C, et al. Using consensus group methods such as Delphi and Nominal Group in medical education research. *Med Teach* 2017;39:14-9. doi: 10.1080/0142159X.2017.1245856.
3. Humphrey-Murto S, Crew R, Shea B, et al. Consensus Building in OMERACT: Recommendations for Use of the Delphi for Core Outcome Set Development. *J Rheumatol* 2019;46:1041-6. doi: 10.3899/jrheum.181094.
4. 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(8):684-706. doi: 10.1177/0269216317690685.
5. 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. doi: 10.1371/journal.pone.0020476.
6. Chan TM, Yarris LM, Humphrey-Murto S. Delving into Delphis. *CJEM* 2019;21:167-9. doi: 10.1017/cem.2019.3.
7. 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. doi: 10.1016/j.jclinepi.2013.12.002.
8. Humphrey-Murto S, Varpio L, Wood TJ, et al. The Use of the Delphi and Other Consensus Group Methods in Medical Education Research: A Review. *Acad Med* 2017;92:1491-8. doi: 10.1097/ACM.0000000000001812.
9. Resemann HK, Clements S, Griffiths A, et al. Reporting of Delphi methods to achieve consensus on guidelines in rare diseases. *Curr Med Res Opin* 2018;34 (Suppl 1):37.
10. Paré G, Cameron AF, Poba-Nzaou P, et al. A systematic assessment of rigor in information systems ranking-type Delphi studies. *Inf Manag* 2013;50:207-17 doi: 10.1016/j.im.2013.03.003.
11. Banno M, Tsujimoto Y, Kataoka Y. Reporting quality of the Delphi technique in reporting guidelines: a protocol for a systematic analysis of the EQUATOR Network Library. *J Clin Epidemiol* 2019;9:e024942.
12. Banno M, Tsujimoto Y, Kataoka Y. The majority of reporting guidelines are not developed with the Delphi method: a systematic review of reporting guidelines. *J Clin Epidemiol* 2020;124;50-7. doi: 10.1016/j.jclinepi.2020.04.010.
13. Gattrell WT, Clements SJ, Sheard D. Quality assessment of guidelines/recommendations developed using Delphi methodology. *Curr Med Res Opin* 2019;35 (Suppl 2):40. doi: 10.1080/03007995.2019.1583496.
14. Ng J. Delphi method: A qualitative approach for quantitative results. *Value Health* 2018;21:S54. Available at https://www.valueinhealthjournal.com/article/S1098-3015(18)30747-2/pdf.
15. Niederberger M, Spranger J. Delphi Technique in Health Sciences: A Map. *Front Public Health* 2020;8:457. doi: 10.3389/fpubh.2020.00457.

Section: References

16. Waggoner J, Carline JD, Durning SJ. Is There a Consensus on Consensus Methodology? Descriptions and Recommendations for Future Consensus Research. *Acad Med* 2016;91:663-8. doi: 10.1097/ACM.0000000000001092.
17. Wang X, Chen Y, Yang N, et al. Methodology and reporting quality of reporting guidelines: systematic review. *BMC Med Res Methodol* 2015;15:74. doi: 10.1186/s12874-015-0069-z.
18. Grant S, Booth M, Khodyakov D. Lack of preregistered analysis plans allows unacceptable data mining for and selective reporting of consensus in Delphi studies. *J Clin Epidemiol* 2018;99:96-105. doi: 10.1016/j.jclinepi.2018.03.007.