Journal of Multidisciplinary Healthcare Dovepress

Open Access Full Text Article REVIEW

A Systematic Scoping Review of Essential Methodological Elements for Developing a Tool to Improve the Reporting of Consensus Studies in Classification, Diagnostic Criteria, and Guidelines Development

Yimy F Medina[1,2], Cindy V Mendieta[1,3], Natalia Prieto[2], María Laura Acosta Felquer[4], Enrique R Soriano[4]

[1]Department of Clinical Epidemiology and Biostatistics, PhD Program in Clinical Epidemiology, Pontificia Universidad Javeriana, Bogotá, Colombia; [2]Department of Internal Medicine, Rheumatology Unit, Faculty of Medicine, Universidad Nacional de Colombia, Bogotá, Colombia; [3]Nutrition and Biochemistry Department, Faculty of Sciences, Pontificia Universidad Javeriana, Bogotá, Colombia; [4]Internal Medicine Department, Rheumatology Section, Internal Medicine Service, Instituto Universitario Hospital Italiano, Buenos Aires, Argentina

Correspondence: Yimy F Medina, Email yimfmedinav@unal.edu.co

Introduction: A consensus is a general agreement among group members that is pivotal in gathering expert input for classification, diagnostic criteria, and guideline development. However, the absence of established methodological standards presents challenges in conducting and analyzing these studies.

Objective: This scoping review explored the evidence on essential elements in consensus studies to create a list of candidate items for a standardized reporting tool. This tool is intended to improve the critical appraisal and methodological rigor of consensus studies.

Methods: A systematic scoping review was conducted using predetermined criteria for study selecting studies and extracting data. A comprehensive literature search was performed without imposing date restrictions, covering multiple databases, including Medline, Embase, LILACS, SciELO, and up to March 2022. We included only English-language publications and excluded incomplete articles and conference reports. The risk of bias was assessed using the CASP checklist, and the study selection and data extraction were performed independently by two researchers in duplicate.

Results: We identified 8360 references; 20 publications were included for data extraction. The majority (70%) used the Delphi method, and the remainder (30%) employed the modified Delphi method. Inconsistencies in reporting conflicts of interest and consensus timing were observed. Other methodologies, such as RAND/UCLA and Nominal Group Technique were excluded due to methodological limitations. Most studies exhibited a low risk of bias.

Discussion: Our findings underscored the need for more standardization in definitions, methodology, and reporting within consensus studies. To address these gaps, we developed a checklist of key reporting items aimed at improving the planning, execution, and reporting consensus studies. Although the developed checklist requires validation, it offers a practical framework to enhance methodological transparency and reliability.

Conclusion: Deficiencies and variability in consensus methodologies reporting underscore the need for a standardized approach. We propose the adoption of a checklist to strengthen the robustness of consensus studies, supporting advances in classification, diagnostic criteria, and guideline development.

Keywords: consensus study, reporting guidelines, scoping review, methodology, completeness of reporting, excellent reporting

Received: 28 June 2024
Accepted: 15 November 2024
Published: 8 December 2024

Introduction

A consensus is defined as a general agreement, implicit or explicitly expressed, among the members of a group.[1] A Consensus study is a method to obtain input from a group of experts and seeks to elicit consensus on the topic under investigation.[2] The Delphi Technique utilizes structured interactions between group members, referred to as the panel, through the use of questionnaires rather than direct face-to-face communication. This method allows for the preservation of participant anonymity, when necessary, making it particularly useful in situations where unbiased input or sensitive feedback is desired.[3] The Nominal Group Technique is a structured method of face-to-face group interaction that empowers participants by providing a platform for them to express their perspectives and ensuring their opinions are recognized by the group. The Nominal Group Technique promotes equitable participation, valuing all contributions, which enhances the overall quality of decision-making. By fostering an inclusive environment, it integrates diverse viewpoints into the collective understanding, ultimately facilitating the resolution of complex issues through a systematic and collaborative process.[4]

The Delphi and Nominal Group Techniques are highly versatile methods that can be employed in conjunction with other approaches to enhance the depth and scope of inquiry. Variations of these techniques may be tailored based on the availability of research, participant time constraints, or the degree of clarification, consensus, or generalizability needed for the specific topic under investigation.[5] The RAND/UCLA (a collaboration between the RAND Corporation and the University of California, Los Angeles) method combines elements of the Delphi and Nominal Group Techniques to identify consensus for future recommendations and highlight areas of disagreement for further research. Experts play a key role in literature analysis and recommendation drafting, with an iterative rating process enhancing the relevance and consistency of rankings. This approach allows for minority opinions, promotes group interaction, and prevents dominance by any single expert.[1]

Clinical practice guidelines are crucial in assisting healthcare professionals, in decision-making and improving patient health outcomes.[3] In addition, diagnostic and classification criteria should be based on evidence from rigorously conducted controlled studies. Formal group consensus methods have been developed to organize subjective judgments and synthesize them with the available evidence.[6] Incorporating a checklist enhances the understandability, reproducibility, and generalizability of studies, while also improving the reporting, execution, and dissemination of research employing diverse methodologies. In general terms, these methods have their corresponding checklist defined according to the kind of research. This is the case for example with PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)[5,7] and AMSTAR 2 (A Measurement Tool to Assess Systematic Reviews) for systematic reviews,[8] CONSORT for clinical trials,[9] and STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) for observational studies.[10] Checklist Increase consistency in designing and publishing studies by providing a framework. They help the user to be systematic by ensuring that all important factors or considerations are taken into account.[11,12]

The development of consensus methodology guidelines is essential to enhance the quality, transparency, and reproducibility of research, particularly in healthcare and clinical decision-making. Poor reporting of consensus methods can impair the interpretation and application of findings, lead to miscommunication, and reduce confidence in the results.[13] Robust guidelines address critical challenges such as unclear definitions of consensus thresholds, inadequate feedback processes, and non-transparent recruitment of experts, which can otherwise hinder the reliability of consensus studies.[14] In areas where evidence is limited or conflicting, consensus methods enable experts to integrate diverse knowledge and provide guidance in uncertain scenarios.[15] By promoting standardized methodologies and clear reporting practices, these guidelines improve the consistency and impact of consensus-based research, fostering better decision-making and ensuring that healthcare practices are supported by reliable evidence.[14]

Objectives

This study aimed to identify the key elements commonly used in consensus studies, focusing on their design, execution, and reporting phases. Our primary objective was to explore these elements across diverse consensus studies, while the secondary objective was to compile candidate items for developing a structured checklist. By identifying and organizing these elements, we aim to establish a foundation for a comprehensive reporting tool that enhances methodological rigor and transparency in consensus research, thereby improving its reliability and impact.

We addressed the following research questions:

What are the key methodological elements involved in the design, execution, and reporting of consensus studies used for classification, diagnostic criteria, and guideline development?

Which candidate items, identified through consensus methodologies in the literature, can serve as the basis for developing a structured reporting tool or framework to improve the transparency and rigor of consensus studies?

Methods

We conducted a systematic scoping review following the methodological framework outlined in the Handbook of Scoping Methods Reviews[16] and its update.[17] We followed our registered protocol, which is available on the Open Science Framework (https://osf.io/v8mqw). To guide and report this review, we adhered to the PRISMA Extension for Scoping Reviews (PRISMA-ScR) guidelines.[18]

Eligibility Criteria

Studies Meeting the Following Criteria Were Included in This Review (Box 1)

Studies that utilized consensus methodologies, including the Delphi method, Modified Delphi, Nominal Group Technique, or RAND/UCLA approaches.

Original research studies that adhered to criteria outlined within consensus methodology.

Studies were deemed relevant if they incorporated essential elements described in the literature on consensus methodology.

Studies involving human participants, with no restrictions on publication year or country of origin.

Studies Exclusion Criteria:

The following types of studies were excluded:

Studies employing methods other than the Delphi, Nominal Group Technique, their modified versions or RAND/UCLA consensus approaches.

Articles lacking sufficient information, such as protocols, meeting abstracts, congress reports, or reviews.

Articles that did not include data on the consensus process.

Based on the literature, we predetermined consensus methodological criteria in the included original studies (Table 1), which were categorized into four phases: pre-registration, registration, consensus, and results. The pre-registration phase encompasses the selection of consensus methodology, rationale behind method selection, literature review, and synthesis procedure, sample size, and research team's role in consensus. The registration phase involves prospective and comprehensive registration of the analysis plan. The consensus phase evaluates the panelist selection process, number of experts, conflicts of interest, criteria for consensus attainment, data analysis, essential elements utilization, participant feedback in each round, and questionnaire modification. The results phase includes limitation discussion, expert final approval, temporality description, and expert anonymity management in qualification.[3,4,6,19–23] The 70% compliance

Box 1 Inclusion/Exclusion Criteria

Consensus study. Item reported, checklist, reporting guidelines. Completeness of reporting. Excellent reporting. Methodology: Systematic review of the literature
Study Design: Consensus Study
Population/Problem: Adult participants: health providers, stakeholders, or patients
Exposures: Qualitative research that aims to understand the how and why of certain individual and group decisions through a consensus study design
Outcomes: Sum of completed items and percentage of compliance
Inclusion criteria: Studies conducted among humans. No restriction on the year of publication. No restriction on the country of origin. Studies written in English
Exclusion criteria: Articles with incomplete information, meetings, or congress reports

Table 1 Review Checklist

Phase	Criteria
Pre-registration phase	The article exposes the reason for choosing the consensus methodology
	The article justifies the choice of the type of consensus (Delphi, modified Delphi, nominal, NIH, RAND-UCLA)
	The consensus method used was stated (Delphi, modified Delphi, nominal group, NIH, RAND-UCLA)
	The rationale for the choice of study was explained
	The review and synthesis of the literature before the construction of the items for consensus is described in detail.
	The rationality of the sample size is described
	The role of the investigator(s) in the consensus was described
Registration phase	The study has a prospective and complete record of the pre-analysis plan
Consensus Phase (Participation of panelists):	The panelist selection process is described
	The eligibility criteria for the panel of experts are as follows: invitation, acceptance, rejection, expertise, and others
	A conflict of interest was declared among the experts.
	The number of invited and included experts is described
	The number of experts who did not accept the invitation is described
	The percentage of participation in each round is described
	Reasons for the exclusion or abandonment of the experts are described
	The article described the number of experts in each of the rounds
	It was defined as how consensus was reached
	The article describes the data analysis process
	The article described the items and the questionnaire used.
	It described the feedback given to the participants after each round
Consensus Phase (Consensus rounds):	The number of rounds was specified a priori
	The number of rounds proposed a priori was met
	The entire consensus process is described
	The type and statistical analysis used to reach the final result are described
	The quantitative results of each round are described
	The qualitative results of each round are described
	The questionnaires for each round are attached
	The process was described when there was no final consensus on the items
	The modification of the questionnaires during the rounds has been described

(*Continued*)

Table 1 (Continued).

Phase	Criteria
Results phase:	Discussion of the study limitations
	The final questionnaire was presented to the experts for their approval
	The temporality of the consensus process is described
	The anonymity of the experts in voting or rating

threshold was established as part of a set of pre-defined criteria encompassing four critical phases: the Pre-registration phase, the Registration phase, the Consensus phase, and the Results phase. This threshold was designed to safeguard the integrity of the synthesis by minimizing the inclusion of studies with significant methodological deficiencies that could introduce bias or inconsistency. While stricter cutoffs may yield more rigorous studies, they risk excluding valuable insights that may have minor shortcomings. Consequently, we reviewers considered that the 70% threshold effectively balances inclusion with quality control.

To evaluate the use and reporting of consensus techniques, we developed a standardized data extraction form. The items included in the form were selected based on information gathered from a comprehensive literature search and refined through an iterative process involving both literature review and team discussions. These items addressed key aspects of the four phases of the consensus process, ensuring a structured and consistent evaluation framework aligned with the objectives of the study.

Search Strategy, Selection of Sources of Evidence

A team of investigators (MLAF, NP, CVM, and YM) conducted a comprehensive literature search in electronic databases until March 4, 2022. The search encompassed five electronic databases: Medline, Embase, LILACS, SciELO, and Scopus via Elsevier and included key terms such as Consensus Study, Item Reported, Checklist, Reporting Guidelines, Completeness of Reporting, Good Reporting, and Methodology. We cross-referenced the selected references to identify potential new publications. The complete literature of the database search is available in Table 2. Following the execution of the search strategy, references from online databases were imported to the Mendeley which was used for screening by the review team. To ensure consistency and accuracy in study selection, two pairs of reviewers (LA, CVM, NP, and YM) independently assessed the titles and abstracts of the identified studies based on the predetermined inclusion and exclusion criteria. For each

Table 2 Search Strategy

Database	Search Strategy
Embase	("consensus"/de OR consensus: ab, ti) AND ("checklist"/de OR checklist OR "reporting"/de OR reporting OR "guidelines"/de OR guidelines OR "completeness"/de OR completeness OR 'good: ab, ti) AND ("reporting"/de OR reporting: ab, ti) AND (methodology) [humans]/lim AND [embase]/lim
MEDLINE Ovid	("consensus"/de OR consensus: ab, ti) AND ("checklist"/de OR checklist OR "reporting"/de OR reporting OR "guidelines"/de OR guidelines OR "completeness"/de OR completeness OR 'good) AND ("reporting"/de OR reporting) AND [humans]/lim AND medline
SciELO	ab:(((consensus) OR (consensus development)) AND ((checklist) OR (reporting guidelines) OR (completeness of reporting) OR (good reporting)))
LILACS	(Consensus development OR development, consensus) AND (checklist OR reporting guidelines OR completeness of reporting OR good reporting) AND (db: ("LILACS") AND la: ("en" OR "es"))
Scopus - Elsevier	TITLE-ABS-KEY ((Consensus W/3 stud*) OR (Consensus) AND TITLE-ABS-KEY (Checklist OR (Reporting W/3 guidelines) OR (Completeness W/3 reporting) OR (Good W/3 reporting))

study deemed potentially eligible, two independent reviewers thoroughly evaluated the full text. In instances of discrepancies, consensus was reached between the reviewers, with involvement from a fifth reviewer (ES) if necessary.

Data Charting Process

The development of the evaluation criteria and checklist items followed a systematic, iterative process to ensure methodological rigor and broad applicability across all consensus studies. The process began with a comprehensive review of relevant literature on consensus methodologies, enabling the research team to identify essential elements critical to robust study design, execution, and reporting. This review provided the foundation for predefining the consensus methodological standards applied in the early phases, including information extraction, data analysis, and checklist item generation.

The research team conducted iterative discussions to refine these standards, ensuring through multiple rounds of deliberation that the criteria aligned with best practices in consensus methodology. To enhance the reliability of the matrix, each variable was assigned equal weight, minimizing the risk of bias across different phases of the evaluation.

A matrix was designed using Microsoft Excel® with pre-established consensus methodological criteria to evaluate the four phases of the consensus studies analyzed. This matrix was applied to the consensus studies that remained after applying inclusion criteria during screening, as well as after removing duplicates or excluding studies for other reasons, including those not retrieved or excluded. The matrix was applied independently to the full-text articles included in the qualitative synthesis by four researchers (YFM, CVM, NP, LA). A minimum compliance threshold of 70% was set as the inclusion criterion for the selection phase of the articles.

A pilot test was conducted using eight articles to identify potential issues and make necessary modifications. Certain concepts were unified to standardize the assessment across all researchers, such as the rationale behind the study design. In addition, items were added to the checklist (eg, description of the role of the researcher(s) in the consensus, description of eligibility criteria for the expert panel, and criteria for inclusion of experts: invitation, acceptance, rejection, expertise) and the reasons for exclusion or non-acceptance of experts in the consensus.

The information extraction matrix included relevant information for each study based on the four pre-established phases: Pre-registration (7 items), Registration (1 item), Consensus (20 items), and Results (4 items). This matrix underwent a pilot test by researchers (YFM, CVM, NP) with five articles. When a study reported including multiple consensus exercises, each method was considered separately. In cases of numerous publications from the same study, the most recent information was extracted. Pre-established criteria (Table 1) were independently evaluated for the pre-registration, consensus, and results phases, with a threshold of 70% considered adequate for inclusion.

Qualitative data were analyzed based on preestablished criteria using frequency measurements using frequency measurements to meet pre-established criteria, with 70% compliance considered adequate for all assessments.

Assessment of Risk of Bias

Although the risk of bias assessments is uncommon in scoping reviews, we opted to include one to provide a more comprehensive evaluation of the quality of the original consensus studies. We conducted a risk of bias assessment for each study using the CASP (Critical Appraisal Skills Program) rating checklist.[12] This additional step aimed to enhance the rigor of our synthesis by identifying potential methodological limitations across the selected studies.

The risk of bias in each study was assessed using the CASP (Critical Appraisal Skills Program) qualification checklist. Three researchers (YFM, CVM, NP) applied the CASP tool to evaluate the risk of bias within each study. This instrument consisted of 10 questions focusing on the validity of results, types of results, and impact of results in qualitative studies.[12,23] A pilot test with three studies was conducted to standardize and refine the classification of studies. The percentage of compliance for each item and article was calculated, beginning with the highest category (low risk of bias).

Results

Search Results and Study Selection

The search yielded a total of 8360 citations from various databases: Medline (1044), Embase (3044), SciELO (43), Lilacs (45), and Scopus - Elsevier (147). After removing 272 duplicate records and excluding 125 records for reasons such as

lack of data on consensus methodology, a total of 3783 references were screened based on title and abstract. Subsequently, 1633 references were selected for full-text analysis. Thirty-one studies could not be retrieved. The exclusions were attributed to language limitations (n = 100), lack of data (n = 138), protocol papers (n = 102), duplicates (n = 92), non-consensus methodology (n = 304), conference papers (n = 79), meeting abstracts (n = 98), and animal studies (n = 28).

A total of 644 full-text articles were included in the qualitative synthesis. Following the pilot test, consensus-driven discussions among the researchers, and the implementation of the 70% compliance threshold, 134 articles met the selection criteria for full-text evaluation. Each paper was assessed by a different reviewer for completeness in the pre-registration, registration, consensus, and results phases. Papers failing to achieve at least 70% compliance with these criteria were excluded. Finally, 20 papers were included for data extraction (Table 3). Among these included papers, 14 (70%) used the Delphi method and 6 (30%) employed the Modified Delphi method. None of the studies used nominal or RAND/UCLA methods. A PRISMA flowchart outlining the screening process is provided (Figure 1).

Phases and Checklist of Items for Evaluating Consensus Methods

The final checklist, comprising four predefined phases (pre-registration, registration, consensus, and results) and their corresponding items used to assess consensus methods, is presented in the Table 4. A total of 28 items were identified: 4 in the pre-registration phase, 1 in the registration phase, 19 in the consensus phase, and 4 in the results phase. The results of each method, delineated according to the criteria of the various phases, are presented:

Delphi Method

The studies employing the Delphi method successfully passed the initial screening and were included in the final analysis, resulting in a total of 14 studies that met the pre-established criteria threshold.

In the Pre-registration phase, fourteen (100%) studies applied the Delphi method.[24,25,29,31,32,35–40,43–45] All studies provided detailed descriptions of the literature review and synthesis. Rationalization of sample size was described in nine (64%) studies.[25,29,36–40,31,35] However, the role of the investigator in the consensus group was not specified in any study.

In the Consensus phase, variations were observed in the expert eligibility criteria. Reports ranged from the absence of eligibility criteria for the expert panel[35] to different methods of expert inclusion, such as inclusion by expertise,[22] invitation and acceptance (14%),[29,36] or based on the degree of knowledge, including invitation, approval, or rejection by the experts themselves.[25,31,38,39] Few articles (7%) established selection criteria beyond those mentioned above.[37]

Three studies (21%) described the invitation, acceptance, and level of expertise of the experts,[25,29,39] while three others (21%) described the invitation and the degree of expertise.[31,32,36] One study (7%) detailed experts' acceptance and degree of expertise.[43]

All studies reported experts' conflicts of interest, and 13 (93%) described the number of invited experts, whereas one study did not.[32] Furthermore, all 14 studies provided descriptions of the percentage of expert participation in each round, the complete process of the consensus method, and the qualitative results of each round.

In the Results phase, all studies discussed their limitations. However, in two studies (14%), the final questionnaire was not submitted to the experts for approval.[29,35] Six studies (43%) described the temporality of the consensus process,[29,37–39,32,45] whereas eight (53%) did not.[24–26,31,35,36,40,43] Nonetheless, all studies maintained the anonymity of the experts in the voting process.

Modified Delphi Method

The studies utilizing the Modified Delphi method successfully passed the initial screening and were included in the final analysis. A total of six studies met the threshold of the pre-established criteria.

In the pre-registration phase, six articles applied a modified Delphi method.[26,30,33,34,41,42] Of these, 17% justified the selection of this method to clarify conflicting evidence,[33] 50% cited the absence of a Clinical Practice Guideline[30,41,42] and the need to generate recommendations,[26,30,42] and 67% used it for standardization.[33,34,41,42] All articles provided detailed descriptions of the literature review and synthesis the research team's role during consensus, and 67% of the articles described the rationale for the sample size.[26,30,33,41]

Table 3 Summary of the Main Characteristics of the Included Studies

Code	Main Author	Publication Date	Country	Year of Study	Consensus Type	Total Study Participants	Sum of Completed Items (Total=13)	Percentage of Compliance (Total=100%)
4	Bishop D.V.M et al[24]	2016	Australia, Canada, Ireland, New Zealand, the United Kingdom, and the USA.	2015	Delphi	59	9	75
157	Ward L et al[25]	2014	Brazil, India, Sri Lanka, Turkey, United Kingdom, and USA	Not described	Delphi	41	10	77
217	Hanson CL et al[26]	2020	United Kingdom	2019	Modified Delphi	47	10	77
237	Beets-Tan RGH et al[27]	2013	Participants from the European Society of Gastrointestinal and Abdominal Radiology (ESGAR)	2012	Delphi	14	11	85
239	Mirabile A et al[28]	2018	USA, Netherlands, United Kingdom, Australia, Canada, Brazil, Italy, Spain, Finland, Norway, Switzerland, Denmark, and Germany	2016	Delphi	91	10	77
288	Dimairo M et al[29]	2018	The majority were from the UK, other European countries, and the USA	2017	Delphi	94	11	85
353	Mistry J et al[30]	2020	United Kingdom, Ireland, Australia, India, Switzerland, Norway, the Netherlands, USA, Italy, South Africa, Greece	2020	Modified Delphi	35	10	77
373	Van Hecke O et al[31]	2015	Participants: International Association for the Study of Pain Special Interest Group (SIG) on Neuropathic Pain and in collaboration with the IASP SIG on Genetics and Pain	2014	Delphi	28	10	77
377	Lynch TS et al[32]	2019	USA	2017	Delphi	55	10	77
387	Kelly SE et al[33]	2016	Canada, United Kingdom, USA, Australia, Spain, and New Zealand	2014	Modified Delphi	66	10	77
399	Sun B et al[34]	2013	Italy, USA, Canada, Netherlands, France, Japan, Switzerland, and United Kingdom	Not described	Modified Delphi	24	10	77
408	Cook C et al[35]	2010	Data were collected using DADOS-Survey, which allows online access to all survey questions for respondents from multiple countries	Not described	Delphi	9	10	77

https://doi.org/10.2147/JMDH.S484715 DovePress

429	Benstoem C et al[36]	2017	Argentina, Australia, Belgium, Brazil, Canada, Chile, China, Denmark, Italy, Macedonia, Mexico, Netherlands, Romania, Russia Federation, Saudi Arabia, Spain, Switzerland, Syria, United Kingdom, USA, Uruguay	2015	Delphi	86	10	77
561	Pandor A et al[37]	2019	Australia, Canada, France, Indonesia, Ireland, Italy, Mexico, South Africa, Taiwan, United Kingdom, and USA	2014	Delphi	80	11	85
569	Heuzenroeder L[38]	2019	Australia, Sweden, and the United Kingdom.	2018	Delphi	49	11	85
611	Diaz-Ledezma C et al[39]	2013	17 countries on five continents	2012	Delphi	159	11	85
640	Breimer GE et al[40]	2015	Australia, Brazil, Canada, France, Germany, Italy, Israel, Netherlands, Turkey, United Kingdom, and USA	Not described	Delphi	35	10	77
645	Jansen LAW et al[41]	2020	Norway, Scotland, and the United Kingdom	2013	Modified Delphi	125	10	77
679	Handler et al[42]	2008	A multidisciplinary expert panel of nursing home physicians, pharmacists, and advanced multidisciplinary professionals	Not described	Modified Delphi	36	10	77
681	Benhamou M et al[43]	2013	Rheumatologists and GPs were randomly selected from two French national databases	2009	Delphi	80	10	77

https://doi.org/10.2147/JMDH.S484715

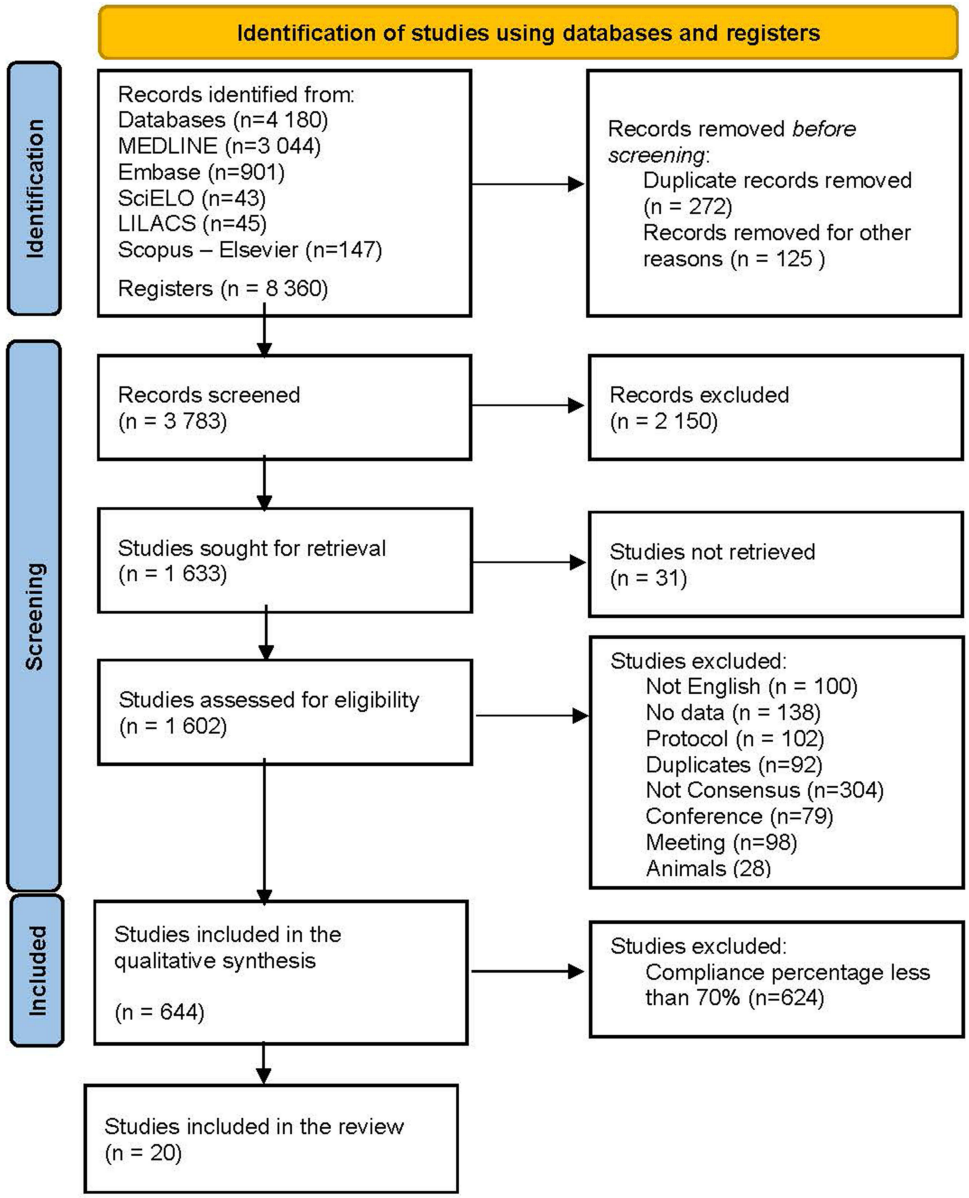

Figure 1 PRISMA flow diagram of the studies included in the review.

In the registration phase, only 33% of the articles[30,41] conducted a prospective and complete registration of the pre-analysis plan. Regarding the consensus phase, all articles described the invitation process for selecting the expert panel, the percentage of participation in each round and the entire consensus process. 67% of the articles described the number of experts who accepted,[26,30,34,42] their experience, and the approach taken to provide a qualitative description of each round. The qualitative synthesis of results was presented in 83% of the articles.[26,30,34,41,42] Similarly, a comparable percentage reported making modifications to the items,[26,30,34,41,42] with four indicating changes to the questionnaire[26,30,31,33,42] and two attaching the entire questionnaire.[34,41]

In the results phase, all articles described the limitations and maintained the anonymity of the experts during the evaluation rounds. However, 67% of the studies reported submitting the final questionnaire to the experts for approval,[33,34,41,42] and only 17% of the articles described the timing of the consensus process.[33]

Table 4 Final Checklist of Phases and Items Used to Evaluate Consensus Methods

Phase	Items		Options
Pre-registration	1. The manuscript exposes the reason for choosing the consensus methodology		YES/NO
	2. The article justifies the choice of the type of consensus method		Delphi; Modified Delphi; Nominal; RAND/UCLA
	3. The rationality of the sample size is described		YES/NO
	4. The role of the investigator(s) in the consensus was described		YES/NO
Registration phase	5. This article has a prospective and complete record of the pre-analysis plan		YES/NO
Consensus phase	**Participation of panelists**	6. The panelist selection process is described	YES/NO
		7. The process of eligibility criteria for the panel of experts is described	Invitation; Acceptance; Rejection; Expertise; Others
		8. A conflict of interest was declared among the experts	YES/NO
		9. The number of experts initially invited and the number of experts included in the study are described	YES/NO
		10. The number of experts who did not accept the invitation is described	YES/NO
		11. The percentage of participation in each round is described. Reasons for the exclusion or abandonment of the experts are described	YES/NO
		12. The article described the number of experts in each of the rounds	YES/NO
	Methodological process	13. The process by which consensus was reached is described	YES/NO
		14. The manuscript describes the data analysis process	YES/NO
		15. The manuscript describes the items and the questionnaire used	YES/NO
		16. It described the feedback given to the participants after each round	YES/NO
	Consensus rounds	17. The number of rounds was specified a priori	YES/NO
		18. The number of rounds proposed a priori was developed	YES/NO
		19. The entire consensus process is described	YES/NO
		20. The type and statistical analysis used to reach the final result are described	YES/NO
		21. The quantitative results of each round are described.	YES/NO
		22. The qualitative results of each round are described.	YES/NO
		23. The questionnaires from each round were included	YES/NO
		24. The modification of the questionnaires during the rounds has been described	YES/NO
Results phase	25. The final questionnaire was presented to the experts for their approval		YES/NO
	26. The temporality of the consensus process is described		YES/NO
	27. The anonymity of the experts in voting or rating each round was maintained		YES/NO
	28. Description of the study limitations		YES/NO

RAND/UCLA

Three studies that utilized the RAND/UCLA method passed the initial screening; however, they were ultimately excluded from the final analysis as they did not meet the predefined threshold criteria (< 70% of compliance). The compliance rate for the articles ranged from 31 to 62%.[27,46,47] These exclusions were attributed to several factors, including the absence

of rationality in sampling, reporting conflicts of interest among the experts, and inadequate documentation of the temporal aspects of the process.

Nominal

None of the five articles employing a nominal consensus approach were included in the final screening round because they did not meet the predefined criterion of at least 70% compliance.[48–52] These exclusions were primarily attributed to the absence of a systematic literature review description and inadequate documentation of the researchers' roles in the consensus process.

Risk of Bias Assessment

All articles included in the analysis demonstrated a clear statement of research objectives, employed an appropriate design to address these objectives, and adequately described the relevance of their findings. Additionally, 70% of the articles employed an appropriate qualitative methodology, collected data in alignment with the study's purpose, implemented an adequate recruitment strategy for participants, and provided a comprehensive explanation of their findings.[24,29,36–40,26,30,31,33,35,45] The most notable areas of concern in the risk of bias assessment included the thoroughness of analytical methods, lack of detailed descriptions regarding ethical considerations, and insufficient information regarding the relationship between participants and the consensus group. The most notable areas of concern in the risk of bias assessment included the thoroughness of analytical methods, lack of detailed descriptions regarding ethical considerations, and insufficient information regarding the relationship between participants and the consensus group (Table 5).

Discussion

This systematic scoping review provides a comprehensive evaluation of published scientific evidence about the development of consensus methodologies. The primary consensus methods examined included Delphi, modified Delphi, Nominal, and RAND/UCLA. Our objective was to explore the range of elements utilized across various consensus studies and compile a list of candidate items for the development of a structured checklist. This was achieved through predefined criteria common to the consensus methodologies included in this review and based on existing literature. It is important to clarify that the exclusion of studies does not imply poor methodological quality.

Considerable variability was observed across studies regarding the characteristics of the consensus procedure. Furthermore, study reports often lacked essential details necessary for interpreting the results. This review highlights significant inconsistencies in the methodology and reporting of the Consensus method used for guidelines.[53]

Despite attempts in the past to define items for inclusion in a consensus methodology checklist, there remains no universally accepted definition.[2,21–23,53]

Upon analysis of the included research studies, it became evident that most employed Delphi, modified Delphi, and RAND/UCLA methods. However, many of these studies lacked clear descriptions of the sampling process, conflicts of interest among panelists, and the overall decision-making. Specifically, the studies utilizing Delphi, modified Delphi, and RAND/UCLA methods often failed to adequately describe the rationale behind sampling, address conflicts of interest among panelists, or provide a comprehensive overview of the decision-making process. This observation is because they do not adequately describe these aspects or themes in the article. In addition, those employing the nominal method frequently omitted descriptions of the systematic literature review and the role of investigators in the consensus process.

The limited number of studies meeting the preestablished criteria underscores the necessity for more explicit quality standards and standardization in consensus study methodology. Although consensus methods are widely employed in research, their implementation varies significantly. There is substantial variation and inconsistency in how analyses are conducted across different rounds and how decisions are made to reach consensus. Ensuring transparency throughout the process remains a critical challenge.[54] The categorization of consensus group methods varies, with some considering them qualitative, others quantitative, and some incorporating elements of both methodologies. The fundamental issue lies in the lack of a robust philosophical foundation, leading to inadequate conceptualization of methods and methodological inconsistency.[55]

Table 5 Risk of Bias Assessment Using the Critical Appraisal Skills Program (CASP) Rating Checklist

Code	Article	1	2	3	4	5	6	7	8	9	10	Compliance*
		Was there a clear statement of the aims of the research?	Is a qualitative methods appropriate?	Was the research design appropriate to address the aims of the research?	Was the recruitment strategy appropriate to the aims of the research?	Was the data collected in a way that addressed the research issue?	Has a relationship between the researcher and participants been adequately considered?	Have ethical issues been taken into consideration?	Was the data analysis sufficiently rigorous?	Is there a clear statement of findings?	How valuable is the research?	
4	Bishop D.V.M et al[24]											90
157	Ward L et al[25]											90
217	Hanson CL et al[26]											90
238	Beets-Tan RGH et al[27]											70
239	Mirabile A et al[28]											80
288	Dimairo M et al[29]											90
353	Mistry J et al[30]											90
373	Van Hecke O et al[31]											90
377	Lynch TS et al[32]											60
387	Kelly SE et al[33]											70
399	Sun B et al[34]											70
408	Cook C et al[35]											90
429	Benstoem C et al[36]											80
561	Pandor A et al[37]											90
569	Heuzenroeder L[38]											90
611	Diaz-Ledezma C et al[39]											90
640	Breimer GE et al[40]											80
645	Jansen LAW et al[41]											80
679	Handler et al[42]											50
681	Benhamou M et al[43]											70
Compliance*		100	95	100	90	95	0	65	70	90	100	

Notes: Conventions: green, low risk of bias; yellow, unclear risk; red, high risk of bias. Compliance*: It was calculated for each item, starting from the highest category (low risk of bias).

https://doi.org/10.2147/JMDH.S484715

Typical features shared by consensus methods include anonymity, iteration, controlled feedback, statistical group response, and structured interaction.[56] Based on these features, we have established general criteria to evaluate studies, which are delineated into four phases: pre-registration, registration, consensus, and results. Each phase encompasses specific characteristics described in the literature, thus contributing to a comprehensive and transparent consensus exercise.

Additional items should be considered for inclusion in the proposed list. For instance, maintaining a comprehensive prospective record of the preanalysis plan for the consensus process during the preregistration phase is crucial.[22] In the consensus phase, it is imperative to ensure the involvement of a diverse range of stakeholders, such as patients or their relatives, administrators, etc. Involving multiple stakeholders can enhance a guideline's acceptability and feasibility for end users, ensure equity and human rights are considered, and support the integration of recommendations into policy and practice. This can ultimately improve adherence to the recommended treatments and practices.[57] Furthermore, an additional phase, the postresults phase, could be introduced to integrate communication and publication requirements.

We observed the absence of studies that utilized the nominal consensus methodology in the final analysis, likely due to its limited applicability. This methodology is constrained by its inability to address more than one issue at a time and its challenge in handling large audiences unless meticulously planned.[56] Numerous studies have explored the reporting of the Delphi method, developed quality indicators, discussed critical appraisal, and emphasized the need for methodological enhancements.[2,20–23] To enhance the critical assessment of the compiled evidence, we developed the CASP tool, which is widely used in assessing the quality of qualitative studies and is endorsed by Cochrane Qualitative and Implementation.[58] While our research identified areas for improvement in describing the researcher-participant relationship, overall, the aim and value of the study were well defined.

The proposed checklist in our literature review addresses aspects that are not covered by existing reporting guidelines for consensus methodology studies. This gap arises from the lack of consensus on evaluating a consensus procedure's applicability, resulting in no universally accepted requirements for utilizing consensus techniques.[53] Furthermore, there is considerable variation in how consensus methodology is implemented, with several modifications to the original Delphi and nominal methods described in the literature. However, standardized definitions for these modifications are not available. Thus, our proposed checklist would contribute to making reports more uniform, transparent and accurate reporting when using consensus methods for data collection.Moreover, the list of items proposed as quality indicators could facilitate implementation by other researchers applying consensus methodology.

At present, universal standards for evaluating the applicability of consensus procedures are lacking, leading to inconsistent use of these methods.[54] Applying this checklist will enhance completeness by promoting consistency in reporting. Given the substantial variation in how Delphi and nominal techniques are modified without standardized definitions, this tool ensures greater uniformity. The proposed checklist could be gradually integrated into consensus studies to ensure applicability across diverse research designs. Early testing in varied settings, paired with iterative feedback from experts, would refine the checklist, enhancing clarity and relevance. This validation process aims to establish the checklist as a robust tool for improving consistency and rigor in consensus-based research reporting. The proposed checklist aims to enhance the clarity, transparency, and completeness of consensus study reporting. By gradually integrating the tool into consensus research and refining it through expert feedback, the checklist seeks to standardize reporting and improve the consistency and rigor of consensus studies. Such guidelines benefit authors, reviewers, and editors, helping to increase the utility and impact of research while promoting more efficient use of resources and investments in health research.[20]

There is no consensus on how to evaluate the applicability of consensus procedures.[54] Accordingly, we identified common elements of applicability through a literature review, acknowledging that their relevance may vary. Several modifications to the original Delphi and Nominal Group methods have been described, although standardized definitions for these variations remain unavailable.[55] Our checklist captures shared features across consensus methodologies, including Delphi, Nominal, and RAND/UCLA. A comprehensive review revealed that Delphi and Modified Delphi studies are more prevalent than Nominal and RAND/UCLA studies. Although multiple consensus methods are employed, Delphi is the most widely used in healthcare research for several reasons.[54,55] It offers a rigorous framework for gathering expert opinions and provides clearer descriptions of agreement levels compared to other approaches.[5,59] Unlike the Nominal Group Technique, which requires in-person meetings, Delphi allows greater flexibility by enabling remote participation via email, minimizing logistical challenges and

travel costs.[55] Therefore, the final proposed checklist is grounded in the shared characteristics of the reviewed consensus methods. While it may appear heavily influenced by Delphi methodologies—potentially limiting its generalizability to other techniques—it remains a comprehensive, literature-informed tool applicable across various consensus approaches.

To the best of our knowledge, this research constitutes the initial endeavor to substantiate the stages and integrate the essential components required in the design and implementation of a consensus methodology. A significant advantage of this study lies in its incorporation of multiple investigations across various healthcare fields, rendering our findings applicable to a wide array of clinical settings.

Limitations

A limitation of our study is that the proposed checklist has not yet been validated, which may introduce biases in the measurement and identification of variables. However, it was developed through a rigorous process involving topical experts in consensus methodologies and clinical epidemiologists. Furthermore, the checklist is extensive, comprising 16 items, and includes prior knowledge in the critical appraisal of evidence.

A limitation of this study is the use of a 70% compliance threshold as part of the pre-established criteria across four phases. While this threshold ensures alignment with essential methodological benchmarks, it aims to balance inclusiveness and quality control rather than pursuing absolute rigor. Although stricter thresholds might improve methodological soundness, they could also exclude studies with minor flaws but valuable insights. To mitigate this risk, the 70% cutoff—agreed upon a priori—was selected to prevent bias while retaining a broad range of relevant data. Similar thresholds are applied by OMERACT (Outcome Measures in Rheumatology)[60] for consensus on core outcome domains and by CREDES (Consensus on the Evaluation and Monitoring of Systemic Diseases),[23] which uses 70–80% to manage variability and ensure coherence. However, the selection of any threshold introduces some subjectivity, which may influence the scope of included studies and the overall synthesis outcomes.

The proposed checklist has not yet undergone formal validation, which we recognize as a limitation. Rather than presenting it as a definitive solution, the checklist serves as a preliminary framework grounded in a comprehensive literature review to address gaps in the reporting of consensus methodologies. Validation efforts are essential to refine the checklist and ensure its ability to enhance the robustness and reliability of reporting. Future research should focus on evaluating and validating the checklist to confirm its effectiveness for consensus-based studies and clinical practice guidelines.

Future Approaches

Building on the findings of our review, we propose a preliminary checklist to enhance the quality of studies utilizing consensus methodologies. Although the checklist requires validation, it provides a structured framework to address current gaps in reporting practices. The insights from this scoping review can serve as a foundation for developing flow diagrams, explicit text, or checklists that guide authors in reporting consensus-based research more transparently and consistently, ultimately contributing to more reliable outcomes in clinical practice guidelines.

Conclusions

Significant flaws and heterogeneity exist in the reporting methodologies of consensus studies. Several authors emphasize the urgent need to enhance consensus methods, particularly in the areas of classification, diagnostic criteria, and guideline development. To address these shortcomings, we advocate the adoption of a standardized checklist in future consensus studies. Based on a thorough scoping review, we propose key items for a checklist that can serve as a foundation to ensure the rigor, transparency, and reliability of reporting methodologies in the development of clinical practice guidelines. This checklist is intended to serve as a starting point for further refinement and validation. Future studies will need to focus on validating the checklist to ensure it achieves the desired robustness and reliability in reporting.

What is Already Known

Consensus studies are essential in health research, particularly in areas like classification, diagnostic criteria, and guideline development. Consensus studies aid as a methodology for soliciting inputs from a panel of experts and aiming to achieve agreement on the topic under investigation.

Formal group consensus methods have been developed to organize subjective judgments and synthesize them with the available evidence.

There is no checklist of characteristics that consensus studies must follow and include, which affects their validity.

What This Paper Adds

This study identifies deficiencies and variability in reporting methodologies within consensus studies in classification, diagnostic criteria, and guideline development. It introduces initial criteria aimed at enhancing the methodological rigor and transparency of these studies. In addition, we design a checklist for the critical appraisal of future consensus studies.

Ethics Declarations

This study adhered to the ethical principles outlined in the Declaration of Helsinki. Because it involved the use of secondary data that cannot identify individuals, it was deemed to pose no risk. This research was conducted within the framework of Dr. Yimy F. Medina's doctoral project in Clinical Epidemiology (Reference code FM-CIE-0837-20).

Funding

This study was self-funded and did not receive external financial support.

Disclosure

Dr María Acosta Felquer reports personal fees from Janssen, support for attending international congress from AbbVie, outside the submitted work. Dr Enrique Soriano reports personal fees from AbbVie, Janssen, Novartis, Pfizer, and UCB, outside the submitted work. The authors report no other conflicts of interest in this work.

References

1. Jones J, Hunter D. Qualitative research: consensus methods for medical and health services research. *BMJ*. 1995;311(7001):376. doi:10.1136/bmj.311.7001.376
2. 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(4):401–409. doi:10.1016/j.jclinepi.2013.12.002
3. Bourrée F, Michel P, Salmi LR. Consensus methods: review of original methods and their main alternatives used in public health. *Rev Epidemiol Sante Publique*. 2008;56(6):e13–21. doi:10.1016/j.respe.2008.10.005
4. Black N, Murphy M, Lamping D, et al. Consensus development methods: a review of best practice in creating clinical guidelines. *J Heal Serv Res Policy*. 1999;4(4):236–248. doi:10.1177/135581969900400410
5. 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(10):1491–1498. doi:10.1097/ACM.0000000000001812
6. Nair R, Aggarwal R, Khanna D. Methods of formal consensus in classification/diagnostic criteria and guideline development. *Semin Arthritis Rheum*. 2011;41(2):95–105. doi:10.1016/j.semarthrit.2010.12.001.
7. Page MJ, McKenzie JE, Bossuyt P, Al E. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:1–9.
8. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*. 2017;358. doi:10.1136/bmj.j4008
9. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *PLoS Med*. 2010;7(3):e1000251. doi:10.1371/journal.pmed.1000251
10. Vandenbroucke JP, Von Elm E, Altman DG, Al E. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. *Ann Intern Med*. 2007;47(8):w–163–194. doi:10.7326/0003-4819-147-8-200710160-00010-w1
11. Meader N, King K, Llewellyn A, et al. A checklist designed to aid consistency and reproducibility of GRADE assessments: development and pilot validation. *Syst Rev*. 2014;3(1):1–9. doi:10.1186/2046-4053-3-82
12. Critical Appraisal Skills Programme (2018). CASP Qualitative Studies Checklist. Available from: https://casp-uk.net/casp-tools-checklists/qualitative-studies-checklist/. Accessed 2 November 2024.
13. Arakawa N, Bader LR. Consensus development methods: considerations for national and global frameworks and policy development. *Res Soc Adm Pharm*. 2022;18(1):2222–2229. doi:10.1016/j.sapharm.2021.06.024.
14. 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(1):1–10. doi:10.1186/s41073-022-00122-0
15. 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;124:50–57. doi:10.1016/j.jclinepi.2020.04.010
16. Peters MDJ, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. *Int J Evid Based Healthc*. 2015;13(3):141–146. doi:10.1097/XEB.0000000000000050
17. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. *JBI Evid Synth*. 2020;18(10):2119–2126. doi:10.11124/JBIES-20-00167

18. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med*. 2018;169(7):467–473. doi:10.7326/M18-0850
19. Veraar C, Hasler P, Schirmer M. A multidisciplinary Delphi consensus-based checklist to define clinical documentation tools for both routine and research purposes. *Heal Serv Res Manag Epidemiol*. 2018;5:233339281775416–233339281775416.
20. Moher D, Schulz KF, Simera I, Altman DG. Guidance for developers of health research reporting guidelines. *PLoS Med*. 2010;7(2):e1000217. doi:10.1371/journal.pmed.1000217
21. Waggoner J, Carline JD, Durning SJ. Is there a consensus on consensus methodology? Descriptions and recommendations for future consensus research. *Acad Med*. 2016;91(5):663–668. doi:10.1097/ACM.0000000000001092
22. 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
23. Jünger S, Payne SA, Brine J, Radbruch L, Brearley SG. Guidance on Conducting and REporting DElphi Studies (CREDES) in palliative care: recommendations based on a methodological systematic review. *Palliat Med*. 2017;31(8):684–706. doi:10.1177/0269216317690685
24. Bishop DVM, Snowling MJ, Thompson PA, et al. CATALISE: a multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. *PLoS One*. 2016;11(7):. doi:10.1371/journal.pone.0158753
25. Ward L, Stebbings S, Sherman KJ, Cherkin D, Baxter GD. Establishing key components of yoga interventions for musculoskeletal conditions: a Delphi survey. *BMC Complement Altern Med*. 2014;14:196. doi:10.1186/1472-6882-14-196
26. Hanson CL, Oliver EJ, Dodd-Reynolds CJ, Pearsons A, Kelly P. A modified Delphi study to gain consensus for a taxonomy to report and classify physical activity referral schemes (PARS). *Int J Behav Nutr Phys Act*. 2020;17(1). doi: 10.1186/s12966-020-01050-2
27. Beets-Tan RGH, Lambregts DMJ, Maas M, et al. Magnetic resonance imaging for the clinical management of rectal cancer patients: recommendations from the 2012 European society of gastrointestinal and abdominal radiology (ESGAR) consensus meeting. *Eur Radiol*. 2013;23(9):2522–2531. doi:10.1007/s00330-013-2864-4
28. Mirabile A, Numico G, Russi EG, et al. Sepsis in head and neck cancer patients treated with chemotherapy and radiation: Literature review and consensus. *Crit Rev Oncol Hematol*. 2015;95(2):191–213. doi: 10.1016/j.critrevonc.2015.03.003
29. Dimairo M, Coates E, Pallmann P, et al. Development process of a consensus-driven CONSORT extension for randomised trials using an adaptive design. *BMC Med*. 2018;16(1):210. doi:10.1186/s12916-018-1196-2
30. Mistry J, Falla D, Noblet T, Heneghan NR, Rushton A. Clinical indicators to identify neuropathic pain in low back related leg pain: a modified Delphi study. *BMC Musculoskelet Disord*. 2020;21(1).doi:10.1186/s12916-018-1196-2
31. Van Hecke O, Kamerman PR, Attal N, et al. Neuropathic pain phenotyping by international consensus (NeuroPPIC) for genetic studies: a NeuPSIG systematic review, Delphi survey, and expert panel recommendations. *Pain*. 2015;156(11):2337–2353. doi:10.1097/j.pain.0000000000000335
32. Lynch TS, Minkara A, Aoki S, et al. Best Practice Guidelines for Hip Arthroscopy in Femoroacetabular Impingement: results of a Delphi Process. *J Am Acad Orthop Surg*. 2020;28(2):81–89. doi:10.5435/JAAOS-D-18-00041
33. Kelly SE, Moher D, Clifford TJ. Defining Rapid Reviews: a Modified Delphi Consensus Approach. *Int J Technol Assess Health Care*. 2016;32(4):265–275. doi:10.1017/S0266462316000489
34. Sun BC, Thiruganasambandamoorthy V, Cruz JD. Standardized reporting guidelines for emergency department syncope risk stratification research. *Acad Emerg Med*. 2012;19(6):694–702. doi:10.1111/j.1553-2712.2012.01375.x
35. Cook C, Brismée JM, Pietrobon R, Sizer P, Hegedus E, Riddle DL. Development of a quality checklist using Delphi methods for prescriptive clinical prediction rules: the QUADCPR. *J Manipulative Physiol Ther*. 2010;33(1):29–41. doi:10.1016/j.jmpt.2009.11.010.
36. Benstoem C, Moza A, Meybohm P, et al. A core outcome set for adult cardiac surgery trials: a consensus study. *PLoS One*. 2017;12(11):e0186772. doi:10.1371/journal.pone.0186772
37. Pandor A, Kaltenthaler E, Martyn-St James M, et al. Delphi consensus reached to produce a decision tool for SelecTing approaches for rapid reviews (STARR). *J Clin Epidemiol*. 2019;114:22–29. doi:10.1016/j.jclinepi.2019.06.005.
38. Heuzenroeder L, Ibrahim F, Khadka J, Woodman R, Kitson A. A Delphi study to identify content for a new questionnaire based on the 10 Principles of dignity in care. *J Clin Nurs*. 2022;31(13–14):1960–1971. doi:10.1111/jocn.15462
39. Diaz-Ledezma C, Higuera CA, Parvizi J. Success after treatment of periprosthetic joint infection: a delphi-based international multidisciplinary consensus infection. *Clin Orthop Relat Res*. 2013;471(7):2374–2382. doi:10.1007/s11999-013-2866-1
40. Breimer GE, Haji FA, Hoving EW, Drake JM. Development and content validation of performance assessments for endoscopic third ventriculostomy. *Child's Nerv Syst*. 2015;31(8):1247–1259. doi:10.1007/s00381-015-2716-4
41. Jansen LAW, Koot MH, Van't Hooft J, et al. A core outcome set for hyperemesis gravidarum research: an international consensus study. *BJOG an Int J Obstet Gynaecol*. 2020;127(8):983–992. doi:10.1111/1471-0528.16172
42. Handler SM, Hanlon JT, Perera S, et al. Consensus list of signals to detect potential adverse drug reactions in nursing homes. *J Am Geriatr Soc*. 2008;56(5):808–815. doi:10.1111/j.1532-5415.2008.01665.x
43. Benhamou M, Boutron I, Dalichampt M, et al. Elaboration and validation of a questionnaire assessing patient expectations about management of knee osteoarthritis by their physicians: the Knee osteoarthritis expectations questionnaire. *Ann Rheum Dis*. 2013;72(4):552–559. doi:10.1136/annrheumdis-2011-201206
44. Richards BL, Whittle S, Buchbinder R, et al. Australian and New Zealand evidence-based recommendations for pain management by pharmacotherapy in adult patients with inflammatory arthritis. *Int J Rheum Dis*. 2014;17(7):738–748. doi:10.1111/1756-185X.12388
45. Chiarotto A, Boers M, Deyo RA, et al. Core outcome measurement instruments for clinical trials in nonspecific low back pain. *Pain*. 2018;159(3):481–495. doi:10.1097/j.pain.0000000000001117
46. Froud R, Eldridge S, Kovacs F, et al. Reporting outcomes of back pain trials: a modified Delphi study. *Eur J Pain*. 2011;15(10):1068–1074. doi:10.1016/j.ejpain.2011.04.015
47. Moore CM, Giganti F, Albertsen P, et al. Reporting magnetic resonance imaging in men on active surveillance for prostate cancer: the PRECISE recommendations—A report of a European school of oncology task force. *Eur Urol*. 2017;71(4):648–655. doi:10.1016/j.eururo.2016.06.011
48. Duffy JMN, Hirsch M, Vercoe M, et al. A core outcome set for future endometriosis research: an international consensus development study. *BJOG an Int J Obstet Gynaecol*. 2020;127(8):967–974. doi:10.1111/1471-0528.16157

49. Deslandes SF, Mendes CHF, de O PT, de S CD. Use of the nominal group technique and the delphi method to draw up evaluation indicators for strategies to deal with violence against children and adolescents in Brazil. *Rev Bras Saude Matern Infant*. 2010;10(SUPPL. 1):29–37. doi:10.1590/S1519-38292010000500003
50. Gaskins M, Dressler C, Werner RN, Nast A.Methods report: update of the German S3 guideline for the treatment of psoriasis vulgaris. JDDG. *J Ger Soc Dermatol*. 2018;16(5):1–92.
51. Duffy JMN, Cairns AE, Richards-Doran D, et al. A core outcome set for pre-eclampsia research: an international consensus development study. *BJOG an Int J Obstet Gynaecol*. 2020;127: 1516–26. Internet.
52. Duffy JMN, AlAhwany H, Bhattacharya S, et al. Developing a core outcome set for future infertility research: an international consensus development study. *Fertil Steril*. 2020; 35(12):2725–34.
53. Boulkedid R, Abdoul H, Loustau M, Sibony O, Alberti C. Using and reporting the Delphi method for selecting healthcare quality indicators, a systematic review. PloS one. 2011; 6(6):e20476
54. Tugwell P, Knottnerus JA. The need for consensus on consensus methods. *J Clin Epidemiol*. 2018;99:vi–viii. doi:10.1016/j.jclinepi.2018.06.001
55. Humphrey-Murto S, Varpio L, Gonsalves C, Wood TJ. Using consensus group methods such as Delphi and nominal group in medical education research. *Med Teach*. 2017;39(1):14–19. doi:10.1080/0142159X.2017.1245856
56. Oghenewiroro Odu G.Relationship between nominal group techniques and concurrent engineering: a review. *Int J Latest Res Eng Technol*. 2017;03 (03):47–62.
57. Petkovic J, Riddle A, Akl EA, et al. Protocol for the development of guidance for stakeholder engagement in health and healthcare guideline development and implementation. *Syst Rev*. 2020;9(1):1–11. doi:10.1186/s13643-020-1272-5
58. Long HA, French DP, Brooks JM. Optimising the value of the critical appraisal skills programme (CASP) tool for quality appraisal in qualitative evidence synthesis. *Res Methods Med Heal Sci*. 2020;1(1):31–42.
59. Okoli C, Pawlowski SD. The Delphi method as a research tool: an example, design considerations and applications. *Inf Manag*. 2004;42(1):15–29. doi:10.1016/j.im.2003.11.002
60. Establishing FOR, Core I. The OMERACT Handbook. OMERACT. 2021. 1–15. Available from: https://omeract.org/handbook/. Accessed November 22, 2024.