HHS Public Access

Author manuscript

Intern Emerg Med. Author manuscript; available in PMC 2015 April 21.

Published in final edited form as:

Intern Emerg Med. 2015 April ; 10(3): 373–383. doi:10.1007/s11739-014-1156-6.

Consensus development for healthcare professionals

Bory Kea, MD and

Assistant Professor, Center for Policy and Research in Emergency Medicine (CPR-EM), Department of Emergency Medicine, Oregon Health & Science University

Benjamin C. Sun, MD, MPH

Associate Professor, Center for Policy and Research in Emergency Medicine (CPR-EM), Department of Emergency Medicine, Oregon Health & Science University

Abstract

Consensus development sprang from a desire to synthesize clinician and expert opinions on clinical practice and research agendas in the 1950s. And since the American Institute of Medicine formally defined “guidelines” in 1990, there has been a proliferation of clinical practice guidelines (CPG) both formally and informally. This modern decision making tool used by both physicians and patients, requires extensive planning to meet the challenges of consensus development while reaping its rewards. Consensus allows for a group approach with multiple experts sharing ideas to form consensus on topics ranging from appropriateness of procedures to research agenda development. Disagreements can shed light on areas of controversy and launch further discussions. It has five main components: three inputs (defining the task, participant identification and recruitment, and information synthesis), the approach (consensus development by explicit or implicit means), and the output (dissemination of results). Each aspect requires extensive planning *a priori* as they influence the entire process, from how information will be interpreted, the interaction of participants, the resulting judgment, to whether there will be uptake of results. Implicit approaches utilize qualitative methods and/or a simple voting structure of majority wins, and are used in informal consensus development methods and consensus development conferences. Explicit approaches aggregate results or judgments using explicit rules set *a priori* with definitions of “agreement” or consensus. Because the implicit process can be more opaque, unforeseen challenges can emerge such as the undue influence of a minority. And yet, the logistics of explicit approaches may be more time consuming and not appropriate when speed is a priority. In determining which method to use, it is important to understand the pros and cons of the different approaches and how it will affect the overall input, approach, and outcome.

Keywords

Consensus development; clinical practice guidelines; research agendas

Corresponding author: Bory Kea, MD, 3181 SW Sam Jackson Park RD, MC CR 114, Portland, OR 97239, borykea@gmail.com, kea@ohsu.edu, Phone: (c) 650-353-6669, (w) 503-494-4430, (f) 503-494-8237.

INTRODUCTION

Clinicians make multiple clinical decisions every day, and ideally, these decisions would be based on rigorous evidenced-based guidelines. Much of health care practice however, has relied on expert opinion or been driven by clinician experience. Compared to other industries, consensus development in health care practice and research is a relatively young field with a growing body of data [1]. In the 1950s, in an attempt to synthesize clinician and 'expert' opinion for clinical practice guidelines and research agendas, the RAND Corporation (a nonprofit, non-partisan global public policy think tank) developed the first approach to consensus guideline development –the Delphi method, which uses iterative surveys without structured interaction to aggregate judgments. Studies using the Delphi method in healthcare and research have included topics on: research priorities for trauma nursing [2], clinical nursing research priorities [3], agreement on classification of electrocardiograms [4], and agreeing terminology for substance abuse [5]. However, this method has been phased out in favor of methods that include the benefits of structured interaction, such as Nominal Group Technique (NGT) and modified Delphi method.

The modified Delphi method, also known as the RAND/UCLA method, popularized by Brook et al of RAND [6, 7], convened three expert panels (made up of a multi-disciplinary set of physicians who were leaders in American medicine). After reviewing synthesized literature, the panels rated the appropriateness of six medical and surgical procedures for a set list of indications in cardiology (coronary angiography and coronary artery bypass surgery), cerebrovascular (carotid endarterectomy), and gastroenterology (colonoscopy, endoscopy and cholecystectomy), prior to meeting and then during a meeting. The study showed it was possible to produce appropriateness ratings of procedures for a large number of indications at a lower cost and logistically less intensive process than a randomized control trial. The predictive validity of such "softer science" methodology has come in the form of multiple randomized control trials (RCT). Shekelle et al 1998 analyzed the results of 7 subsequent RCTs on carotid endarterectomy (CEA) to assess the predictive validity of the appropriateness criteria from the RAND/UCLA study. The RCTs confirmed 44 of the indications for CEA (covering 30% of the operations performed in 1981) and refuted no ratings [8]. Thus, while not generating new information, there is added-value to consensus development until there is an RCT available.

Since the American Institute of Medicine's formal definition of "guidelines" in 1990, clinical practice guidelines (CPG) through consensus have become the modern decision-making tool for patient and clinicians for specific clinical scenarios [9]. Both fortunately and unfortunately, formal and informal guidelines have proliferated making it difficult to users to judge the quality or appropriateness of the guidelines for a specific population [10–12]. As a result, there has been a growth in guideline clearing houses to systematically identify, review and organize guidelines. From national guideline clearing houses (such as the United States (US) Agency for Healthcare Research and Quality (AHRQ) Guideline Clearing House, United Kingdom (UK) National Institutes for Health and Clinical Excellence), condition level clearing houses, to professional societies have been developed [1]. Since then the National Guideline Clearing House adopted the American Institute of Medicine primer on guideline development, Clinical Practice Guidelines We Can Trust [13].

This group approach to guidelines or research agenda development allows for an expanded range of knowledge and experience by multiple individuals, stimulates a deeper discussion with the inclusion of expert recommendations, challenges existing ideas, and increases the legitimacy of the output. Furthermore, the results are more likely to have both concurrent validity and internal logic, otherwise "closer to the truth."

However, there are also inherent challenges to the group approach as well. For example, the methods of a group interaction can influence how the group arrives at the truth and the resulting output [1]. Additionally, selection of the "right" participants may include research scientists, clinical decision-makers and patients who may all have different agendas. Thus, management of the interaction must be considered *a priori* in order to prevent dominance by a minority (e.g. high status members, individuals with strong opinions or personalities) while simultaneously ensuring contribution of introverted experts. Furthermore, if a group is too cohesive or there is undue influence of a minority or majority occur, the output is at risk of group think, group polarization, or minority influence [14]. And of course with larger groups, the logistics of organizing more people is a factor for both cost and resources.

GOAL

The objective of this article is to review the common approaches to consensus development and methodology, as well as the benefits and negatives of consensus. Specifically, this article will review examples in the health care literature, for which there is a slow growing body of literature, and on the development of quality indicators and future research agendas. Schunemann et al has published exhaustively on guideline development and systematic reviews, and refer the interested reader to these papers for details [15, 16]. The development of clinical practice guidelines have also been extensively covered by Shekelle et al [17–19], and their methodology will be briefly reviewed here.

OVERVIEW

The structure of consensus development can be categorized into five components: three inputs, the approach (consensus development), and the output [18]. In the input, tasks must be defined, participants are identified and recruited, and information synthesized (literature review and synthesis). The aggregation of judgments, the approach, can take on two methods: implicit and explicit. The implicit approach utilizes qualitative methods and/or a simple structure of majority vote, and is used in informal consensus development methods and consensus development conferences. The explicit aggregation of judgments involve more complex statistical methods (with or without weighting of judgment, means of individual judgments, mathematical integrations) to form consensus, and is exercised in the Delphi, NGT, and modified Delphi method. The output depends on the audience, but should use the active language to provide recommendations.

I. Defining the Task

Every group is tasked to address one or more questions proposed by the organizers. In some consensus conferences, these may be just a few questions, whereas with the Delphi method and NGT, they can include numerous questions that narrow after an initial survey. In clinical

practice guidelines (CPG), tasks are defined by a desired clinical behavior and specific outcomes of interest (intermediate, surrogate, and health outcomes) are addressed *a priori* to making recommendation, such as identifying "electrocardiogram changes as a surrogate for cardiac ischemia" [19]. The primary task should have an overarching outcome of interest, and may include hundreds of subtasks for specific clinical scenarios. These subtasks are rated for appropriateness of diagnostic or interventional actions, similar to the modified Delphi method.

This approach can be extended to research agendas as well where broad research questions are ranked and prioritized for impact, with subtasks for each question. This list is then refined with further input, or feedback from sponsoring groups, until a research agenda is formed with tens of agenda items. These tasks can be defined by sponsoring organizations such as the US National Institutes of Health (NIH), US Preventative Service and Task Force (USPSTF), or UK National Health Services, or by professional societies such as American College of Emergency Medicine or American Thoracic Society, or smaller interest groups such as the Pediatric Emergency Care Applied Research Network [18, 17, 19].

In an iterative approach, defining tasks can be first characterized by a literature review that includes synthesis and methodological grading. Subsequently, a conceptual model is developed to structure the task and focus participants [1]. The initial tasks are then defined based on participants' perspectives, sponsor requirements, and gaps in the literature. An advisory group, non-voting stakeholders not part of the guideline group, can provide feedback to further refine the tasks.

II. Identifying and Recruiting Participants [1]

The participants of a consensus development group play a major role in the recommendations or output of their interaction. Thus, the composition must be determined *a priori* where the expertise can range from clinicians (who implement the guidelines and are decision-makers), research scientists (investigating the questions methodologically) to patients (who have direct experience with the disease) [1]. For example, in one RAND/UCLA study [6, 7], experts rated 6 procedures in 3 categories. For cardiac procedures, experts included cardiologists, internists, family physicians, cardiothoracic surgeons, and a radiologist. For gastroenterology procedures, experts included a family physician, internists, gastroenterologists, general surgeons, and a radiologist. For cerebrovascular procedures, a family physician, internist, neurologists, vascular surgeons, a neurosurgeon, and a neuroradiologist were included on the panel. This heterogeneity can provide unique perspectives, alternative approaches in problem-solving, increased idea-sharing and buy-in from other fields [1].

However, if subgroups develop into cohesive groups with polarized views, there is a potential for conflict as well [1]. Constructive conflict resolution may be necessary when there is considerable controversy. And yet, if there are differing opinions among experts, participants are likely drift to the mean. Such diversity is also useful for the development of research agendas or exploration of areas of uncertainty.

A homogenous group in contrast, is more likely to reach agreement more quickly than a heterogeneous group, and will tend to have consistent results when compared across similarly composed groups. This increased validity is useful for finding common ground quickly but may lead to particular judgments due to polarized homogenous views.

Who—Most consensus development groups include experts in the field of interest who have credibility in the target audience as it may improve the validity [1]. Identification of such individuals may be based upon professional society recommendations, number of publications, administrative leadership, reputation or peer consensus. Individual with higher status however, are likely to dominate an interaction with a greater influence on the group. And while their input is important, it is equally important to include the input of other members. Therefore, methods to mitigate the unbalanced input of a few participants should be undergone (e.g. confidential voting) [1].

Another set of stakeholders, increasingly included in expert panels, are the patients. With the rise of the US Patient-Centered Outcomes Research Institute (PCORI), the focus of patient involvement is changing. PCORI was created as part of the Affordable Care Act (ACA) in the US, President Obama's universal health insurance initiative for Americans without health insurance, out of a need to address patient- and clinician-driven questions in real-world settings. It is financed through the US Treasury and a private health insurance fee. As an independent body from the government, it is solely a research funding institution that aims to generate information. Grant applications undergo a peer-review process that includes 3 patients or stakeholder reviewers. PCORI believes that patients and family members with direct experiences with a disease are able to contribute in a meaningful way as they are able address issues important to them or ask more patient-pertinent questions. All projects funded by PCORI have an explicit requirement to include all stakeholders (physicians, nurses, caregivers, family members, pharmaceutical industry, etc…) with a "stakeholder engagement plan" where there is meaningful involvement [20]. Patients are also on PCORI's grant review committees, thus, they have a prominent role in multiple aspects of research.

Optimal number—When there are more judges, the reliability of the composite judgment increases and is presumed to be more accurate [1]. Hogarth et al 1978 found that group validity was closer to the "truth" with more than five participants [21], thus most literature indicate at least six participants is more reliable. While the reliability will improve with greater numbers, there are diminishing returns with numbers greater than 12 as it is may be logistically more difficult to coordinate [1]. However, current technology can potentially incorporate many participants with the use of web-based surveys for different aspects of consensus development [22].

Conflict of interest—An important aspect gaining attention in healthcare, research and CPG, is the disclosure of potential conflicts of interest [23]. Many journals and conferences now require authors and presenters to disclose their conflicts of interest which may include testimony, stock holding, honoraria, in kind gifts, pain expert testimony, or acting board members. This conflict represents bias that could influence not only the question to be addressed but also the proceedings and final recommendations. Discovering such

undisclosed conflicts can adversely affect the reputation of those setting forth the guidelines, such as with the Surviving Sepsis Campaign [23].

Eichacker et al 2006 startling article uncovered the multiple conflicts of interest in this campaign [23]. In this campaign, Eli Lilly's Xigris (recombinant human activated protein C, or rhAPC, also known as drotrecogin alfa-activated) was included into performance bundles for sepsis treatment despite controversial evidence from a single phase 3 randomized, control trial—the Recombinant Activated Human Protein C Worldwide Evaluation in Severe Sepsis (PROWESS study) [24]. The US Food and Drug Administration (FDA) approved the drug contingent on prospective validation on subgroups due to a high risk of death and hemorrhage associated with the drug. As a result, sales of the drug declined and Eli Lily hired a public relations firm to improve sales. They launched a three pronged approach to improve sales: 1-target physicians and medical trade media, 2-because it was such an expensive drug, they created an environment of "rationing" such that physicians felt they were deciding who would survive if given the medication, and 3-establish the Surviving Sepsis Campaign to increase awareness about sepsis and develop guidelines on treatment. Eleven societies endorsed the guidelines, and held workshops on sepsis at their meetings supported by Lily. Lily also sponsored a popular quarterly periodical, Advances in Sepsis, that furthered their campaign and calls for new prospective studies of rhAPC.

The company supported multiple individuals in different societies and had a number of articles published where authors were board members of the company. Despite this undue sponsor influence, the guidelines were endorsed by multiple societies. The Infectious Disease Society of America was the only society that found methodological flaws in the guidelines and rejected the guidelines; however, it was not widely publicized. Two further studies, ADDRESS in 2005 [25] and RESOLVE in 2007 [26] warned of issues with rhAPC, but the campaign ignored these studies until the PROWESS-SHOCK study in 2011 [27] when Lily withdrew it from the market. Preventing the creation of conflicted guidelines requires declaration of potential participant conflicts of interest prior to their involvement in guideline development.

III. Information

Information, or evidence, given to participants will strongly influence their judgments [28]. Without information, participants rely on past experiences to form them [29]. As a result, organizers must determine the quantity of information to provide, the selection of evidence and how to present it as it will affect how participants assimilate, interpret, and use the information. Too much information will overwhelm participants, but too little will not provide enough data. By providing a literature review and/or library of articles and abstracts, participants will have equal access to all available data and create a common starting point for discussion [1].

Systematic reviews are a crucial component of many CPGs as they may summarize the current literature for the proposed CPG question. The American Institute of Medicine developed standards in 2011 for their development which is organized into four areas: (1) initiating a systematic review, (2) finding and assessing individual studies, (3) synthesizing the body of evidence, and (4) reporting of systematic review. However, how panelists

evaluate the literature, including systematic reviews, may impact the outcome of the CPG. Thus, methods assessing quality of a systematic review is essential to the CPG process. While there are multiple tools available, two popular methods include AMSTAR (Assessment of Multiple Systematic Reviews) [30, 31], and GRADE[32, 16, 33–52] (Grade of Recommendation, Assessment, Development and Evaluation), which aims to develop a common, transparent sensible system for grading the quality of evidence and the strength of evidence. This method has been taken up by multiple organizations including the World Health Organization, the Cochrane Collaboration, the National Institute of Clinical Excellence, UpToDate, and many other organizations. GRADE Profiler Software is freely available for downloading: http://tech.cochrane.org/gradepro.

While the systematic reviews, original articles and abstracts are less likely to be misinterpreted, presenting synthesized data in the forms of tables are easier for participants to assimilate and use the data [1]. Multiple studies have shown that how data and outcomes are presented will influence judgments [53–57]. In one Swiss study, physicians were asked to rate the effectiveness of drug treatment for hyperlipidemia. One questionnaire presented the data as absolute risk reduction for non-fatal and fatal myocardial infarction combined, fatal myocardial infarction alone, and total mortality. In a second questionnaire, the same data was presented as relative risk reduction. The effectiveness ratings for relative risk reductions were higher than absolute risk reduction. Thus, data syntheses and presentation must be carefully planned *a priori* such that participants are less likely to misinterpret the data.

Prior experience is also a form of information that is can influence how new information is interpreted. Experts are more likely to rely on experience but also may be better apt in reviewing literature for relevant articles in their area of expertise [1]. The involvement of methodologists during a literature review will likely improve the quality of the review. A methodologist is more likely to be better weight relevant, methodologically sound and appropriate literature during a synthesis. Subsequently, the judgments based on the reviews toward a more appropriate judgment [1]. Ultimately, judgments are influenced not only by the selection of literature reviewed, but also the impact of values and preferences for various outcomes by panelist.

Another aspect of information is feedback after voting. Feedback can be presented in different forms, but often includes the scores of individual participants, all participants anonymized and the group scores with means and medians. Such information is critical prior to another round of voting in the Delphi, modified Delphi and NGT methods as forms of structured communication. This iterative process is likely to decrease the dispersion of the voting and improve convergence and accuracy of a judgment, but not necessarily validity [1].

IV. Consensus Approaches

The aggregation of judgments can take on two approaches: implicit and explicit. The implicit approach utilizes qualitative methods and/or a simple voting structure of majority wins, and is used in informal consensus development methods and consensus development conferences. In contrast, the explicit aggregation of judgments involve more complex

statistical methods (with or without weighting of judgment, means of individual judgments, mathematical integrations) to form consensus, and is exercised in the Delphi, NGT, and modified Delphi.

Implicit approaches often begin with a literature review. Ideally, participants are given a selected methodology for grading the evidence, on quality, quantity and consistency across the literature, with consideration for publication bias [19]. This information is used in the group's discussion and then views are summarized by the facilitators. While easy to administer, the process may be opaque and dissent hidden. And yet, it may be the only available approach when certain questions do not lend themselves to a multiple item format for NGT, such as: What are the top research priorities for a certain clinical specialty? Most professional societies releasing CPG utilize this approach.

The explicit approach is similarly first characterized with a literature review that includes simple narrative summaries, evidence tables, meta-analyses, and modeling [19]. Using this information and their own perspectives, participants undergo several rounds of anonymized voting. Results are then aggregated using explicit rules set *a priori*. Furthermore, there are explicit definitions of "agreement" or "consensus," such as a certain percentage of concordance and lack of discordance equates agreement. Due to the multiple rounds of voting and explicit rules, it is more difficult to administer than the implicit approach.

The Delphi method developed in the 1950s by RAND attempted to reap the benefits of low-cost information (as members may have a wide geographic distribution) but limit the detriments of interaction [1]. Questionnaires are sent to individual participants and responses are summarized anonymously (also known as the forecast) and then returned back to members with the inclusion of their initial response. Members are encouraged to revise their response in light of the new information received. This iterative process can be continued multiple times. Votes can be statistically weighted for member expertise, and then responses are statistically aggregated. Critics however, remarked that the lack of interaction leads to less discourse about disagreements and does not provide an opportunity for explanations of perspectives. As a result, it has fallen out of favor.

In contrast to the Delphi method, the NGT utilizes structured interaction where members privately record their ideas, and then in a round robin style voice their idea until all unique ideas are recorded. The group then discusses each idea listed, and then members privately rank the ideas based on the new information. Individual rankings are statistically aggregated to derive a group judgment. Because there are two phases, the first phase of idea generation increases participation by more inhibited members and increases the number of ideas as all members must give an idea. Thus, this is an excellent method for idea generation but may be difficult to reach a valid consensus if there are multiple heterogeneous ideas without a particular direction and lack of iterative process to reach consensus.

Because of the criticism of the original Delphi method, a modified Delphi method was developed—also known as the Rand form of the NGT, an amalgamation of both the original Delphi and NGT [6]. A group of 9 experts are chosen and must agree upon a set of decision-making scenarios. Each expert ranks a scenario 1 through 9, where 1 is extremely

inappropriate (risk greatly exceed benefit) and 9 is extremely appropriate (benefits greatly exceed risk). These votes are aggregated and discussed as a group. In phase 2, each expert again casts their votes privately and then appropriateness scores are calculated based on the median. Agreements and disagreements are determined based on a statistical definition using binomial distribution. This approach allows for both structured interaction and explicit aggregation of votes.

Strategies for improving approach

Implicit: Based on prior experience of the author (BCS), challenges can arise when members dominate the discussion or if members do not commit to form consensus. Potential solutions include ground rules for discussion (round robin, parliamentary procedure, etc…) and *a priori* commitment to obtain consensus by all members.

Explicit: Strong disagreements can also be challenging, and determining a process to address it *a priori* can help alleviate downstream tension and conflict. In modified Delphi, disagreement is defined by four definitions where they note the degree of disagreement within a 3 point range and the entire 9 point range for both *all* participant ratings and then the ratings that remained after discarding the extreme high and low rating. While disagreement can be conflict provoking, useful information can be obtained with structured disagreement as it can uncover controversial topics.

V. Output

The dissemination of results has several approaches and must be considered *a priori* given the context of the audience. There may be a primary audience and a secondary audience. Not only is the scientific community impacted by new guidelines, but also the lay public with potential policy ramifications. For example, in 2009 the U.S. Preventative Services Task Force convened an independent group of experts (implicit approach) that deemed yearly mammograms were no longer necessary for all women younger than 50 (unless they had a family history of breast cancer or other risk factors) [58]. Despite using the most up-to-date literature, there was an immediate backlash by patient advocacy groups and professional societies [59]. Women saw it as a threat to their lives and their daughters lives and triggered an enormous amount of media coverage. Even the White House became involved as it affected President Obama's new health insurance policy. As a result, the USPTF modified their language such that decision to start regular, biennial mammograms before age 50 be an individual one and include the consideration of patient values [58]. The negative press however, increased the skepticism of women regarding their physicians' perspectives on mammography.

Such negative press and distrust from the public ought to be prevented. Methods to avoid such an incident may include peer review or consultation with the public to critically review drafts and provide feedback prior to final publication. This group however does not have voting capabilities but allows the opportunity for increased validity, alternative viewpoints, and "buy-in" by a heterogeneous group. Explicit rules, and/or transparent methods, for inclusion or ignoring criticism must be created in order to justifiably invoke and ignore feedback [18].

Guidelines and agendas are often published in scientific journals, but strategies to support guideline uptake may include development of educational resources, using multiple formats and channels for guideline dissemination based on preferences of the target group, use of behaviorally specific language in the guideline, pre-emptive identification of potential barriers of recommendations, and *a priori* generation of solutions to address them by the guideline development group, use of data collection tools, and identification of resource implications of recommendations [18, 60].

EXAMPLES OF CONSENSUS APPROACHES IN THE HEALTH CARE LITERATURE

In the emergency medicine literature, consensus development on quality measures and guidelines has primarily been approached through the modified Delphi method. This approach is best used where a pre-identified set of items are easily rated, and structured interaction is conducive to further discussion, followed by feedback and re-voting to bring about consensus on a more narrowed scope. In research agenda setting, both structured and unstructured approaches have been used. If speed is priority, the unstructured approach is best, however, a structured approach will likely have more transparency and increased consensus.

NGT/Modified Delphi

1. Penciner et al 2011 conducted a study to establish consensus on emergency medicine clerkship competencies in Canada using modified Delphi process [61]. They convened an expert panel of thirty persons from nine different medical schools. An initial list of 152 competencies were rated in the round 1, and reduced to 62 after a second round of the Delphi process.
2. In Lindsay et al 2002 on ED quality indicators were developed using the modified-Delphi approach [62]. After a literature review to identify conditions frequently treated in the ED and the outcomes associated with the condition by the authors, an expert panel was convened to rate specific clinical condition-outcome pairs where quality of care could improve the outcome for the condition through a questionnaire. As a result, another literature review was performed to identify measurable clinical indicators of performance for these pairs. Participants then rated these indicators for appropriateness using a Likert-type scale (1-“strongly disagree” to 9 “strongly agree”) in another questionnaire. In a structured interaction, panel participants were given both individual and summarized results and then allowed to discuss each indicator in turn similar to the NGT technique. Panelists then individually rated the indicators in light of the new information. Core indicators were selected based on strong agreement between panelists.
3. Solberg et al 2003 study on ED crowding measures [63], used both the modified-Delphi and NGT techniques to develop ED quality measures using structured interaction and web-based anonymous voting. A core set of investigators and 74 expert participants developed 6 rating categories for measures, and elicited ED quality measures outlined by their conceptual model from all experts in an

anonymous fashion. Duplicate measures and those not meeting criteria were eliminated. In a structured interaction, core investigators further refined the measures (by combining, rewording and eliminating those of little value). These measures were then rated by anonymized expert participants via a website using magnitude estimation technique. Summarized ratings and comments underwent another review with core investigators during another meeting and added 8 more measures which were then rated on a website by expert participants as well, and resulted in a final set of ED crowding measures

Internet-based survey

1. In Sun et al 2012 study on ED syncope risk-stratification reporting guidelines [22], guidelines were developed using an expert panel modified Delphi process. Using internet-based surveys in an iterative approach, experts anonymously rated tasks in two rounds. After the first round, both summarized and anonymized individual ratings were given back to participants for review. Further, structured interaction was encouraged through moderated conference calls and e-mail forums. And then a second round of internet-based surveys took place privately. Those items achieving greater than 80% consensus were included in the final guidelines.

Unstructured research agenda setting

In 2010, research consensus conferences were held with US National Institutes of Health (NIH) Task Force members and academicians, to prioritize opportunities for facilitating and conducting emergency care research to advise the new Office in Emergency Research at NIH [64-66]. With the multi-disciplinary reach of Emergency Medicine both clinically and within research, it was critical for NIH to have clear research agendas established to justify targeted research funding. These round table discussions advised the new NIH Office of Emergency Care on emergency trauma, medical-surgical, and neurologic and psychiatric emergency research.

1. In emergency trauma and the medical-surgical emergency roundtables, experts within the field were invited to list challenges and research priorities in a pre-conference survey. This list was revised during and after the conference until consensus was reached [64, 66].
2. In the neurologic and psychiatric emergencies roundtable, a multidisciplinary planning group developed an initial agenda, selected participants and presenters, and based on the state of the science, burden of disease, and request for information responses, narrowed the final research agenda by consensus voting during a roundtable conference [65].

Structured research agenda setting

1. In the Italian syncope project, an international multi-disciplinary panel of syncope experts identified research priorities in a preconference survey on syncope, developed a conceptual model of ED decision-making for syncope, and then refined during and after the conference through an iterative approach until priorities for research were published [67].

CONCLUSION

Consensus development can be used for both clinical practice guidelines and research agendas. ‘Expert Consensus’ is important when no evidence exists for most clinical/research scenarios. In determining which approach to use, it is important to understand the pros and cons of the different approaches and how it will affect the overall input, process, and outcome. Because the implicit process can be more opaque, unforeseen challenges can emerge such as the undue influence of a minority. Also, lists developed without a limitation to the number of items or prioritization can make these lists potentially useless. Careful planning, however, can mitigate both of these situation using structured interaction as well as clear aims and priorities set prior to meetings. Its primary strength compared to explicit approaches is the rapid assembly of “experts” and dissemination through journals. While explicit approaches may result in more valid and transparent results, it logistically requires significantly more effort and resources than implicit approaches.

References

1. Murphy MK, Black NA, Lamping DL, McKee CM, Sanderson CF, Askham J, Marteau T. Consensus development methods, and their use in clinical guideline development. Health technology assessment. 1998; 2(3):i–iv. 1–88. [PubMed: 9561895]
2. Bayley EW, Richmond T, Noroian EL, Allen LR. A Delphi study on research priorities for trauma nursing. American journal of critical care: an official publication. American Association of Critical-Care Nurses. 1994; 3(3):208–216.
3. Bond S, Bond J. A Delphi survey of clinical nursing research priorities. Journal of advanced nursing. 1982; 7(6):565–575. [PubMed: 6759554]
4. Kors JA, van Herpen G, van Bemmel JH. Variability in ECG computer interpretation. Analysis of individual complexes vs analysis of a representative complex. Journal of electrocardiology. 1992; 25(4):263–271. [PubMed: 1402511]
5. Rinaldi RC, Steindler EM, Wilford BB, Goodwin D. Clarification and standardization of substance abuse terminology. JAMA: the journal of the American Medical Association. 1988; 259(4):555–557.
6. Brook RH, Chassin MR, Fink A, Solomon DH, Kosecoff J, Park RE. A method for the detailed assessment of the appropriateness of medical technologies. International journal of technology assessment in health care. 1986; 2(1):53–63. [PubMed: 10300718]
7. Park RE, Fink A, Brook RH, Chassin MR, Kahn KL, Merrick NJ, Kosecoff J, Solomon DH. Physician ratings of appropriate indications for six medical and surgical procedures. American journal of public health. 1986; 76(7):766–772. [PubMed: 3521341]
8. Shekelle PG, Chassin MR, Park RE. Assessing the predictive validity of the RAND/UCLA appropriateness method criteria for performing carotid endarterectomy. International journal of technology assessment in health care. 1998; 14(4):707–727. [PubMed: 9885461]
9. Field, M.; Lohr, K. Committe to Advise the Public Health Service on clinical practice guidelines IoM: Clinical practice guidelines: directions for a new program. Institute of Medicine; Wshington, D.C: 1990.
10. Weisz G, Cambrosio A, Keating P, Knaapen L, Schlich T, Tournay VJ. The emergence of clinical practice guidelines. The Milbank quarterly. 2007; 85(4):691–727.10.1111/j.1468-0009.2007.00505.x [PubMed: 18070334]
11. Hurwitz B. Clinical guidelines: proliferation and medicolegal significance. Quality in health care: QHC. 1994; 3(1):37–44. [PubMed: 10136259]
12. HIaQ Authority. National Quality Assurance Criteria for Clinical Guidelines. Health Information and Quality Authority, dublin. 2011

13. Greenfield, S.; Steinberg, EP.; Auerbach, A.; Avorn, JL.; Galvin, RS.; Gibbons, R.; Lau, J.; Morrow, M.; Mulrow, CD.; Rosoff, AJ.; Santa, J.; Shiffman, RN.; smitt, WR.; Stewart, WR.; Stovall, E.; Storm, BL.; Titler, MG. Clinical Practice Guidelines We Can Trust. Institute of Medicine; Washington, DC: 2011.

14. Pagliari C, Grimshaw J. Impact of group structure and process on multidisciplinary evidence-based guideline development: an observational study. Journal of evaluation in clinical practice. 2002; 8(2):145–153. [PubMed: 12180363]

15. Schunemann HJ, Wiercioch W, Etxeandia I, Falavigna M, Santesso N, Mustafa R, Ventresca M, Brignardello-Petersen R, Laisaar KT, Kowalski S, Baldeh T, Zhang Y, Raid U, Neumann I, Norris SL, Thornton J, Harbour R, Treweek S, Guyatt G, Alonso-Coello P, Reinap M, Brozek J, Oxman A, Akl EA. Guidelines 2.0: systematic development of a comprehensive checklist for a successful guideline enterprise. CMAJ: Canadian Medical Association journal = journal de l'Association medicale canadienne. 2014; 186(3):E123–142.10.1503/cmaj.131237

16. Guyatt G, Vandvik PO. Creating clinical practice guidelines: problems and solutions. Chest. 2013; 144(2):365–367.10.1378/chest.13-0463 [PubMed: 23918097]

17. Shekelle P, Woolf S, Grimshaw JM, Schunemann HJ, Eccles MP. Developing clinical practice guidelines: reviewing, reporting, and publishing guidelines; updating guidelines; and the emerging issues of enhancing guideline implementability and accounting for comorbid conditions in guideline development. Implementation science: IS. 2012; 7:62.10.1186/1748-5908-7-62 [PubMed: 22762242]

18. Eccles MP, Grimshaw JM, Shekelle P, Schunemann HJ, Woolf S. Developing clinical practice guidelines: target audiences, identifying topics for guidelines, guideline group composition and functioning and conflicts of interest. Implementation science: IS. 2012; 7:60.10.1186/1748-5908-7-60 [PubMed: 22762776]

19. Woolf S, Schunemann HJ, Eccles MP, Grimshaw JM, Shekelle P. Developing clinical practice guidelines: types of evidence and outcomes; values and economics, synthesis, grading, and presentation and deriving recommendations. Implementation science: IS. 2012; 7:61.10.1186/1748-5908-7-61 [PubMed: 22762158]

20. PCORI. Engagement in Research. Patient-Centered Outcomes Research Institute; 2014. http://www.pcori.org/get-involved/what-is-engagement-in-research/ [Accessed 23 June 2014 2014]

21. Hogarth R. A note on aggregating opinions. Organizational Behavior and Human Performance. 1978; 21:40–46.

22. Sun BC, Thiruganasambandamoorthy V, Cruz JD. EDSRSR Consortium to Standardize. Standardized reporting guidelines for emergency department syncope risk-stratification research. Academic emergency medicine: official journal of the Society for Academic Emergency Medicine. 2012; 19(6):694–702.10.1111/j.1553-2712.2012.01375.x [PubMed: 22687184]

23. Eichacker PQ, Natanson C, Danner RL. Surviving sepsis--practice guidelines, marketing campaigns, and Eli Lilly. The New England journal of medicine. 2006; 355(16):1640–1642.10.1056/NEJMp068197 [PubMed: 17050887]

24. Bernard GR, Vincent JL, Laterre PF, LaRosa SP, Dhainaut JF, Lopez-Rodriguez A, Steingrub JS, Garber GE, Helterbrand JD, Ely EW, Fisher CJ Jr. CWEiSSsg Recombinant human protein . Efficacy and safety of recombinant human activated protein C for severe sepsis. The New England journal of medicine. 2001; 344(10):699–709.10.1056/NEJM200103083441001 [PubMed: 11236773]

25. Abraham E. Effects of recombinant human activated protein C in human models of endotoxin administration. Proceedings of the American Thoracic Society. 2005; 2(3):243–247.10.1513/pats.200501-004AC [PubMed: 16222045]

26. Nadel S, Goldstein B, Williams MD, Dalton H, Peters M, Macias WL, Abd-Allah SA, Levy H, Angle R, Wang D, Sundin DP, Giroir B. REs Sepsis, g Organ dysfunction in children: a gLobal perspective study . Drotrecogin alfa (activated) in children with severe sepsis: a multicentre phase III randomised controlled trial. Lancet. 2007; 369(9564):836–843.10.1016/S0140-6736(07)60411-5 [PubMed: 17350452]

27. Ranieri VM, Thompson BT, Barie PS, Dhainaut JF, Douglas IS, Finfer S, Gardlund B, Marshall JC, Rhodes A, Artigas A, Payen D, Tenhunen J, Al-Khalidi HR, Thompson V, Janes J, Macias WL, Vangerow B, Williams MD. P-SS Group . Drotrecogin alfa (activated) in adults with septic

shock. The New England journal of medicine. 2012; 366(22):2055–2064.10.1056/NEJMoa1202290 [PubMed: 22616830]

28. Jacoby I. Evidence and consensus. JAMA: the journal of the American Medical Association. 1988; 259(20):3039.
29. Fink A, Kosecoff J, Chassin M, Brook RH. Consensus methods: characteristics and guidelines for use. American journal of public health. 1984; 74(9):979–983. [PubMed: 6380323]
30. Shea BJ, Bouter LM, Peterson J, Boers M, Andersson N, Ortiz Z, Ramsay T, Bai A, Shukla VK, Grimshaw JM. External validation of a measurement tool to assess systematic reviews (AMSTAR). PloS one. 2007; 2(12):e1350.10.1371/journal.pone.0001350 [PubMed: 18159233]
31. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, Porter AC, Tugwell P, Moher D, Bouter LM. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC medical research methodology. 2007; 7:10.10.1186/1471-2288-7-10 [PubMed: 17302989]
32. Guyatt G, Oxman AD, Sultan S, Brozek J, Glasziou P, Alonso-Coello P, Atkins D, Kunz R, Montori V, Jaeschke R, Rind D, Dahm P, Akl EA, Meerpohl J, Vist G, Berliner E, Norris S, Falck-Ytter Y, Schunemann HJ. GRADE guidelines: 11. Making an overall rating of confidence in effect estimates for a single outcome and for all outcomes. Journal of clinical epidemiology. 2013; 66(2):151–157.10.1016/j.jclinepi.2012.01.006 [PubMed: 22542023]
33. Guyatt GH, Oxman AD, Santesso N, Helfand M, Vist G, Kunz R, Brozek J, Norris S, Meerpohl J, Djulbegovic B, Alonso-Coello P, Post PN, Busse JW, Glasziou P, Christensen R, Schunemann HJ. GRADE guidelines: 12. Preparing summary of findings tables-binary outcomes. Journal of clinical epidemiology. 2013; 66(2):158–172.10.1016/j.jclinepi.2012.01.012 [PubMed: 22609141]
34. Guyatt GH, Thorlund K, Oxman AD, Walter SD, Patrick D, Furukawa TA, Johnston BC, Karanicolas P, Akl EA, Vist G, Kunz R, Brozek J, Kupper LL, Martin SL, Meerpohl JJ, Alonso-Coello P, Christensen R, Schunemann HJ. GRADE guidelines: 13. Preparing summary of findings tables and evidence profiles-continuous outcomes. Journal of clinical epidemiology. 2013; 66(2): 173–183.10.1016/j.jclinepi.2012.08.001 [PubMed: 23116689]
35. Guyatt G, Eikelboom JW, Akl EA, Crowther M, Gutterman D, Kahn SR, Schunemann H, Hirsh J. A guide to GRADE guidelines for the readers of JTH. Journal of thrombosis and haemostasis: JTH. 2013; 11(8):1603–1608.10.1111/jth.12320 [PubMed: 23773710]
36. Andrews JC, Schunemann HJ, Oxman AD, Pottie K, Meerpohl JJ, Coello PA, Rind D, Montori VM, Brito JP, Norris S, Elbarbary M, Post P, Nasser M, Shukla V, Jaeschke R, Brozek J, Djulbegovic B, Guyatt G. GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. Journal of clinical epidemiology. 2013; 66(7):726–735.10.1016/j.jclinepi.2013.02.003 [PubMed: 23570745]
37. Andrews J, Guyatt G, Oxman AD, Alderson P, Dahm P, Falck-Ytter Y, Nasser M, Meerpohl J, Post PN, Kunz R, Brozek J, Vist G, Rind D, Akl EA, Schunemann HJ. GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. Journal of clinical epidemiology. 2013; 66(7):719–725.10.1016/j.jclinepi.2012.03.013 [PubMed: 23312392]
38. Guyatt GH, Oxman AD, Schunemann HJ. GRADE guidelines-an introduction to the 10th–13th articles in the series. Journal of clinical epidemiology. 2013; 66(2):121–123.10.1016/j.jclinepi.2012.05.011 [PubMed: 22968177]
39. Brunetti M, Shemilt I, Pregno S, Vale L, Oxman AD, Lord J, Sisk J, Ruiz F, Hill S, Guyatt GH, Jaeschke R, Helfand M, Harbour R, Davoli M, Amato L, Liberati A, Schunemann HJ. GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence. Journal of clinical epidemiology. 2013; 66(2):140–150.10.1016/j.jclinepi.2012.04.012 [PubMed: 22863410]
40. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, Devereaux PJ, Montori VM, Freyschuss B, Vist G, Jaeschke R, Williams JW Jr, Murad MH, Sinclair D, Falck-Ytter Y, Meerpohl J, Whittington C, Thorlund K, Andrews J, Schunemann HJ. GRADE guidelines 6. Rating the quality of evidence--imprecision. Journal of clinical epidemiology. 2011; 64(12):1283–1293.10.1016/j.jclinepi.2011.01.012 [PubMed: 21839614]
41. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, Alonso-Coello P, Glasziou P, Jaeschke R, Akl EA, Norris S, Vist G, Dahm P, Shukla VK, Higgins J, Falck-Ytter Y, Schunemann HJ, Group GW. GRADE guidelines: 7. Rating the quality of evidence--

inconsistency. Journal of clinical epidemiology. 2011; 64(12):1294–1302.10.1016/j.jclinepi.2011.03.017 [PubMed: 21803546]

42. Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J, Alonso-Coello P, Djulbegovic B, Atkins D, Falck-Ytter Y, Williams JW Jr, Meerpohl J, Norris SL, Akl EA, Schunemann HJ. GRADE guidelines: 5. Rating the quality of evidence--publication bias. Journal of clinical epidemiology. 2011; 64(12):1277–1282.10.1016/j.jclinepi.2011.01.011 [PubMed: 21802904]
43. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, Alonso-Coello P, Falck-Ytter Y, Jaeschke R, Vist G, Akl EA, Post PN, Norris S, Meerpohl J, Shukla VK, Nasser M, Schunemann HJ, Group GW. GRADE guidelines: 8. Rating the quality of evidence--indirectness. Journal of clinical epidemiology. 2011; 64(12):1303–1310.10.1016/j.jclinepi.2011.04.014 [PubMed: 21802903]
44. Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, Alonso-Coello P, Atkins D, Kunz R, Brozek J, Montori V, Jaeschke R, Rind D, Dahm P, Meerpohl J, Vist G, Berliner E, Norris S, Falck-Ytter Y, Murad MH, Schunemann HJ, Group GW. GRADE guidelines: 9. Rating up the quality of evidence. Journal of clinical epidemiology. 2011; 64(12):1311–1316.10.1016/j.jclinepi.2011.06.004 [PubMed: 21802902]
45. Scott IA, Guyatt GH. Clinical practice guidelines: the need for greater transparency in formulating recommendations. The Medical journal of Australia. 2011; 195(1):29–33. [PubMed: 21728938]
46. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, Montori V, Akl EA, Djulbegovic B, Falck-Ytter Y, Norris SL, Williams JW Jr, Atkins D, Meerpohl J, Schunemann HJ. GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). Journal of clinical epidemiology. 2011; 64(4):407–415.10.1016/j.jclinepi.2010.07.017 [PubMed: 21247734]
47. Brozek JL, Akl EA, Compalati E, Kreis J, Terracciano L, Fiocchi A, Ueffing E, Andrews J, Alonso-Coello P, Meerpohl JJ, Lang DM, Jaeschke R, Williams JW Jr, Phillips B, Lethaby A, Bossuyt P, Glasziou P, Helfand M, Watine J, Afilalo M, Welch V, Montedori A, Abraha I, Horvath AR, Bousquet J, Guyatt GH, Schunemann HJ, Group GW. Grading quality of evidence and strength of recommendations in clinical practice guidelines part 3 of 3. The GRADE approach to developing recommendations. Allergy. 2011; 66(5):588–595.10.1111/j.1398-9995.2010.02530.x [PubMed: 21241318]
48. Balshem H, Helfand M, Schunemann HJ, Oxman AD, Kunz R, Brozek J, Vist GE, Falck-Ytter Y, Meerpohl J, Norris S, Guyatt GH. GRADE guidelines: 3. Rating the quality of evidence. Journal of clinical epidemiology. 2011; 64(4):401–406.10.1016/j.jclinepi.2010.07.015 [PubMed: 21208779]
49. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, Norris S, Falck-Ytter Y, Glasziou P, DeBeer H, Jaeschke R, Rind D, Meerpohl J, Dahm P, Schunemann HJ. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. Journal of clinical epidemiology. 2011; 64(4):383–394.10.1016/j.jclinepi.2010.04.026 [PubMed: 21195583]
50. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, Alderson P, Glasziou P, Falck-Ytter Y, Schunemann HJ. GRADE guidelines: 2. Framing the question and deciding on important outcomes. Journal of clinical epidemiology. 2011; 64(4):395–400.10.1016/j.jclinepi.2010.09.012 [PubMed: 21194891]
51. Guyatt GH, Oxman AD, Schunemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. Journal of clinical epidemiology. 2011; 64(4):380–382.10.1016/j.jclinepi.2010.09.011 [PubMed: 21185693]
52. Guyatt G, Akl EA, Hirsh J, Kearon C, Crowther M, Gutterman D, Lewis SZ, Nathanson I, Jaeschke R, Schunemann H. The vexing problem of guidelines and conflict of interest: a potential solution. Annals of internal medicine. 2010; 152(11):738–741.10.7326/0003-4819-152-11-201006010-00254 [PubMed: 20479011]
53. Fahey T, Griffiths S, Peters TJ. Evidence based purchasing: understanding results of clinical trials and systematic reviews. BMJ (Clinical research ed). 1995; 311(7012):1056–1059. discussion 1059-1060.
54. Bucher HC, Weinbacher M, Gyr K. Influence of method of reporting study results on decision of physicians to prescribe drugs to lower cholesterol concentration. BMJ (Clinical research ed). 1994; 309(6957):761–764.

55. Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. The New England journal of medicine. 1988; 318(26):1728–1733.10.1056/NEJM198806303182605 [PubMed: 3374545]
56. Bettman JW Jr, Bettman JW Sr. One-day hospitalization for extracapsular cataract surgery without phacoemulsification. Ophthalmic surgery. 1977; 8(4):81–86. [PubMed: 331177]
57. Jarvenpaa SL. Graphic displays in decision making-the visual salience effect. Behavioral Decision Making. 1990; 3(4):247–262.10.1002/bdm.3960030403
58. Summaries for patients. Screening for breast cancer: U.S. Preventive Services Task Force recommendations. Annals of internal medicine. 2009; 151(10):I44.10.7326/0003-4819-151-10-200911170-00002 [PubMed: 19920254]
59. Goodman, A. [Accessed 18 June 2014 2014] Attitudes Toward Mammography Guidelines Influenced by Media. Medscape. 2011. http://www.medscape.com/viewarticle/744107
60. Gagliardi AR, Brouwers MC, Palda VA, Lemieux-Charles L, Grimshaw JM. How can we improve guideline use? A conceptual framework of implementability. Implementation science: IS. 2011; 6:26.10.1186/1748-5908-6-26 [PubMed: 21426574]
61. Penciner R, Langhan T, Lee R, McEwen J, Woods RA, Bandiera G. Using a Delphi process to establish consensus on emergency medicine clerkship competencies. Medical teacher. 2011; 33(6):e333–339.10.3109/0142159X.2011.575903 [PubMed: 21609170]
62. Lindsay P, Schull M, Bronskill S, Anderson G. The development of indicators to measure the quality of clinical care in emergency departments following a modified-delphi approach. Academic emergency medicine: official journal of the Society for Academic Emergency Medicine. 2002; 9(11):1131–1139. [PubMed: 12414461]
63. Solberg LI, Asplin BR, Weinick RM, Magid DJ. Emergency department crowding: consensus development of potential measures. Annals of emergency medicine. 2003; 42(6):824–834.10.1016/S0196064403008163 [PubMed: 14634610]
64. Cairns CB, Maier RV, Adeoye O, Baptiste D, Barsan WG, Blackbourne L, Burd R, Carpenter C, Chang D, Cioffi W, Cornwell E, Dean JM, Dyer C, Jaffe D, Manley G, Meurer WJ, Neumar R, Silbergleit R, Stevens M, Wang M, Weiner D, Wright D. P Roundtable External, C Roundtable Steering, P Federal. NIH Roundtable on Emergency Trauma Research. Annals of emergency medicine. 2010; 56(5):538–550.10.1016/j.annemergmed.2010.05.029 [PubMed: 21036294]
65. D'Onofrio G, Jauch E, Jagoda A, Allen MH, Anglin D, Barsan WG, Berger RP, Bobrow BJ, Boudreaux ED, Bushnell C, Chan YF, Currier G, Eggly S, Ichord R, Larkin GL, Laskowitz D, Neumar RW, Newman-Toker DE, Quinn J, Shear K, Todd KH, Zatzick D. P Roundtable External, C Roundtable Steering, P Federal. NIH Roundtable on Opportunities to Advance Research on Neurologic and Psychiatric Emergencies. Annals of emergency medicine. 2010; 56(5):551–564.10.1016/j.annemergmed.2010.06.562 [PubMed: 21036295]
66. Kaji AH, Lewis RJ, Beavers-May T, Berg R, Bulger E, Cairns C, Callaway C, Camargo CA Jr, Carcillo J, DeBiasi R, Diaz T, Ducharme F, Glickman S, Heilpern K, Hickey R, Hoek TV, Hollander J, Janson S, Jurkovich G, Kellermann A, Kingsmore S, Kline J, Kuppermann N, Lowe R, McLario D, Nathanson L, Nichol G, Peitzman A, Richardson L, Sanders A, Shah M, Shapiro N, Silverman R, Than M, Wilber S, Yealy DM. P Roundtable External, C Roundtable Steering, P Federal. Summary of NIH Medical-Surgical Emergency Research Roundtable held on April 30 to May 1, 2009. Annals of emergency medicine. 2010; 56(5):522–537.10.1016/j.annemergmed.2010.03.014 [PubMed: 21036293]
67. Sun BC, Costantino G, Barbic F, Bossi I, Casazza G, Dipaola F, McDermott D, Quinn J, Reed M, Sheldon RS, Solbiati M, Thiruganasambandamoorthy V, Krahn AD, Beach D, Bodemer N, Brignole M, Casagranda I, Duca P, Falavigna G, Ippoliti R, Montano N, Olshansky B, Raj SR, Ruwald MH, Shen WK, Stiell I, Ungar A, van Dijk JG, van Dijk N, Wieling W, Furlan R. Priorities for Emergency Department Syncope Research. Annals of emergency medicine. 201410.1016/j.annemergmed.2014.04.014

Table 1

Summary of consensus development methods.

	Implicit Methods	Explicit Methods
Approaches	Informal consensus development, consensus development conferences, unstructured research agenda setting	Delphi, NGT, modified Delphi (or RAND/UCLA), structured research agenda setting
General methodology	Qualitative approach and/or simple structure of majority vote, usually of experts	Structured interaction (for NGT and modified Delphi) with explicit rules for aggregation of judgments with complex statistical methods
Pros	Easy to administer, rapid assembly of experts and dissemination	Results are likely more valid, transparency in methods
Cons	Opaque process, hidden dissent, unclear structure of interaction, potential greater influence or dominance of interaction due to participant status, long lists without prioritization	Logistically more difficult to administer, time consuming
Examples	Most clinical practice guidelines released by professional societies; roundtable US NIH discussions to set research priorities in EM	Consensus on ED quality indicators, ED crowding measures, research priorities in syncope
When to use	Research agenda setting where there are multiple research priorities, speed is a priority, logistically infeasible to do explicit methods, lack evidence for most clinical/research scenarios	When not limited by time, logistics, and evidence

Abbreviations: NGT, nominal group technique; ED, emergency department; EM, emergency medicine; US NIH, United States National Institutes of Health