
Available from: http://dx.doi.org/10.7748/nr.22.3.35.e1297

Accessed from: http://hdl.handle.net/1959.13/1333938
Title:
Research methods for formal consensus development in health.

Authors:
Daphne J James¹, Helen M Warren-Forward¹

Author Affiliations:
1. School of Health Sciences, University of Newcastle, Australia.

Corresponding author contact details:
Name: Daphne James
Address: School of Health Sciences, University of Newcastle, University Drive, Callaghan. N.S.W. 2308 Australia
Telephone: +61 2 49215596
Fax: +61 2 49217053
Email: Daphne.James@newcastle.edu.au

Research Higher Degree student – PhD (MRS), University of Newcastle
KEYWORDS
Research design, consensus guidelines

ABSTRACT

Aim: This paper provides a review of three research methods for consensus development.

Background: Consensus statements and guidelines are increasingly used to clarify and standardise practice, and inform health policy when relevant and rigorous evidence is lacking. Clinicians need to evaluate the quality of practice guidelines to determine whether to incorporate them into clinical practice or reject them. Formal methods of consensus development provide a scientific method, using expert panel members, to evaluate current evidence and expert opinions to produce consensus statements for clinical problems.

Data Sources: Online search for relevant literature was conducted in Medline and CINAHL.

Review Methods: A literature review on consensus, consensus development and research methods papers published in peer-reviewed journals and written in English.

Discussion: The three methods of consensus development discussed are the Delphi technique, nominal group technique and the consensus development conference. The techniques and their respective advantages are described, and examples from the literature are provided. The three methods are compared and a flowchart to assist researchers selecting an appropriate method is included. Online resources with information on the development and evaluation of clinical guidelines are reviewed.

Conclusion: This paper will assist researchers to select an appropriate research method for development of consensus statements and guidelines.

Implications for research/practice: When developing consensus guidelines for clinical practice, researchers should use a formal research method to ensure rigour and credibility.
INTRODUCTION

Consensus statements and guidelines are frequently used in health care to inform practice and ensure appropriate practice policies for specific patient conditions. Clinical practice guidelines are used to assist clinicians in decision making and to provide a consistent approach across health departments (Gabel & Shipan, 2004). Ideally these guidelines should be based on sound scientific evidence, however in practice most are derived from the opinions and experiences of clinicians or an expert in the area at the time. In areas where there is a lack of evidence or where evidence is contradictory, practice can vary widely (Murphy et al., 1998). These variations in practice patterns are a significant concern in the health professions as they may result in inconsistencies in patient care and treatment. Expert consensus panels are increasingly used as a decision-making tool to develop practice guidelines and treatment policies (Gabel & Shipan, 2004).

Consensus development methods have been used in health since the 1950’s (Black, 2006). This type of research method uses a quantitative approach to organise the opinions and judgements of a group of people, ie qualitative data. Formal methods bring together a group of experts to evaluate evidence, comment on statements and ideas, and ultimately come to a consensus opinion on a clinical problem (Vakil, 2011). They attempt to identify all relevant issues and frame these into a series of explicit statements which the group participants rank as to their level of agreement with each statement (Black, 2006). Formal methods attempt to overcome the disadvantages found in informal group decision making processes such as domination of the discussion by a particular individual or pressure to agree to a majority or powerful person’s opinion (Murphy et al., 1998).

Consensus development methods are not a method of creating new scientific knowledge, rather they serve as a process to improve clinical decision-making and assist in the development of health policy (Halcomb et al., 2008). Their objective is to determine a central tendency and grade the level of agreement reached (Nair, Aggarwal, & Khanna, 2011). Consensus is not necessarily defined as complete agreement between participants. Instead
consensus may be defined by a final vote with a pre-determined percentage of agreement (e.g. 80%) (Falzarano & Pinto Zipp, 2013) or by a rating scale where a specified mean rating is achieved for each topic. (Nair et al., 2011)

All methods of formal consensus development consist of several key features:

1. Experts are provided with an independent summary of all scientific and research evidence pertaining to the issue.
2. Experts provide their views privately so other members are unaware of their judgements.
3. Experts are given the opportunity to change their initial opinions after seeing the group views.
4. Statistical analysis is used to derive a group decision (Black, 2006).

Several factors attribute to the success of a consensus method. Selection of appropriate participants is important in determining the outcomes of the group process. Participants should be considered experts in the field, either by virtue of clinical experience or a thorough knowledge of the literature (Baker, Lovell, & Harris, 2006). Often patients or other lay persons are incorporated into the group because they have personal experience of the impact of the disease or intervention in question. A diverse group may be able to consider all aspects of the topic, however this may lead to increased levels of disagreement (Nair et al., 2011). The size of the group should be selected carefully. Larger groups can make the process difficult to manage but may result in an increased reliability of the final decisions. Once recruited, participants should be provided with a summary of current literature to ensure all participants begin with a common level of understanding and that the process remains evidence based.(Murphy et al., 1998).

Health care is increasingly influenced by economics, politics, and social and cultural factors (Halcomb et al., 2008). The use of research methods designed to achieve consensus across a range of stakeholders are frequently being used. This paper provides a review of three
formal consensus development methods used in health care: the Delphi technique, consensus development conference, and nominal group technique (NGT); and offers guidance on which method to use for particular situations before providing information on how to evaluate the quality of guidelines prior to incorporation into clinical practice.

**FORMAL CONSENSUS METHODS**

**THE DELPHI TECHNIQUE**

The Delphi technique was first developed by the Rand Corporation in the 1950’s and named after the oracle at Delphi (Vernon, 2009). It was originally used in technological forecasting and to synthesise expert opinion on new technology (Murphy et al., 1998). Since the 1970’s it has been used extensively in health, in particular in the nursing profession. There are a number of modified versions of this technique and over the years it has often been criticised for a lack of methodological rigour (Hasson, Keeney, & McKenna, 2000).

The Delphi technique is characterised by the following factors: expert panel, iteration with feedback, statistical group response, and anonymity (Vernon, 2009). It utilises a series of questionnaires; each followed by analysis and feedback. The Delphi can be conducted via email with online surveys or via post. Therefore, it can be applied to groups with large numbers of participants from different geographical areas when it is not practical to bring them together (Nair et al., 2011).

Prior to the first round the goal of the Delphi and a definition of consensus should be defined by the research team. A thorough literature search should be completed to evaluate any existing evidence and a summary should be provided to each participant. The participants, or expert panel, are selected based on their clinical and/or research expertise (Vernon, 2009). Typically 3-4 survey rounds are completed with iterative analysis and feedback. Some areas of consensus may emerge from each round and any areas not reaching consensus are developed into subsequent rounds. When an acceptable level of consensus
The advantages of the Delphi technique are:

- the ability to gain the opinions of large numbers of experts,
- participants are able to express their opinions freely due to anonymity,
- reduction in the potential for moderator bias or dominance by an individual,
- cost effectiveness and convenience,
- application to a diverse range of topics, and
- it can be used preceding a nominal group technique in a modified Delphi version.

Some disadvantages of the method include: reliance on questionnaire design and selection of “expert” panel, no personal contact between experts, possible lack of generalizability or scientific validation of findings, and difficulties coordinating large groups (Falzarano & Pinto Zipp, 2013; Nair et al., 2011). Also due to the required number of rounds and their iterative nature, it can be a lengthy process.

The Delphi technique has been utilised in a wide variety of applications in health care to establish consensus opinion, identify research priorities, and develop clinical guidelines. A table of examples of recent studies that have used the Delphi technique is provided (Table 1) showing the diversity of its application. The Delphi should be the research method of choice when there is little scientific evidence or conflicting evidence on the topic, and when the cost and practicalities of bringing the participants together is prohibitive.
<table>
<thead>
<tr>
<th>Study</th>
<th>Purpose</th>
<th>Rounds &amp; completion time</th>
<th>Panel</th>
<th>Outcomes</th>
<th>Consensus definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA (Goligher, Ferguson, &amp; Kenny, 2012)</td>
<td>Identify and standardise the core clinical knowledge and skills required to care for patients receiving mechanical ventilation</td>
<td>4 rounds, 26 months</td>
<td>14 participants, Content experts and educators</td>
<td>List of learning objectives in 8 categories</td>
<td>Not reported</td>
</tr>
<tr>
<td>South Africa (Versteeg, du Toit, &amp; Couper, 2013)</td>
<td>Obtain consensus on the biggest challenges and important priorities for rural health care delivery in South Africa</td>
<td>3 rounds, Time frame not reported</td>
<td>53 participants, Health workers, academics</td>
<td>List of top 5 priorities and challenges</td>
<td>Not reported</td>
</tr>
<tr>
<td>Australia (Walmsley, Rivett, &amp; Osmotherly, 2009)</td>
<td>Establish consensus regarding clinical identifiers for early stage primary adhesive capsulitis</td>
<td>3 rounds, Time frame not reported</td>
<td>70 participants, Musculoskeletal experts</td>
<td>List of 8 clinical identifiers established</td>
<td>Not reported</td>
</tr>
<tr>
<td>USA (Goodrich et al., 2013)</td>
<td>Establish consensus on case history questions and eye examinations for patients with mild traumatic brain injury</td>
<td>2 rounds, Time frame not reported</td>
<td>16 participants, Optometrists</td>
<td>17 history questions and 7 examination procedures</td>
<td>80%</td>
</tr>
<tr>
<td>Sweden (Lindberg, Lundström-Landegren, Johansson, Lidén, &amp; Holm, 2012)</td>
<td>Describe core competencies for nursing practice in renal care in Sweden</td>
<td>4 rounds, 6 months</td>
<td>17 participants, Renal care nurses</td>
<td>List of 43 core competencies</td>
<td>75%</td>
</tr>
</tbody>
</table>

**CONSENSUS DEVELOPMENT CONFERENCE (CDC)**

This method was devised by the US National Institutes of Health and differs from other consensus approaches by providing a public forum for the discussion of issues (Vernon, 2009). A decision-making group of participants (usually about 10 people) are chosen for their methodological expertise rather than expertise in the area of concern (Vakil, 2011). They are presented with evidence from a small group of experts in the topic who are not involved in the decision making process. In this method any type of evidence, including research evidence, clinical expertise and consumer experiences can be presented to the decision-making panel (Halcomb et al., 2008). The meeting is chaired and the panel discuss the evidence and attempt to reach consensus through questioning and discussion. Similar to a legal trial, the group, like a jury, hear evidence and later deliberate, make judgements and produce a definitive consensus statement by the conclusion of the conference proceedings (Duncan, 2006). However, unlike a jury, the panel are able to ask questions to clarify ideas.
and perceptions, and any audience members may also contribute to the discussion. The chairperson, or facilitator, controls the proceedings, directing discussion and delegating tasks (Murphy et al., 1998; Vakil, 2011). It is important that the facilitator moderating the discussion is independent and experienced (Halcomb et al., 2008). The facilitator should ensure that all panel members are given an opportunity to contribute to the discussion and that any potential conflicts are managed appropriately. The optimal panel size is reported to be between six and twelve participants, as reliability declines with less than six, and more than twelve becomes difficult to manage (Fink, Kosecoff, Chassin, & Brook, 1984; Murphy et al., 1998).

The main advantage of the CDC is that it “fosters dialogue, debate and discussion”. It allows for interaction between participants which is important when multiple perspectives are being considered (Halcomb et al., 2008). Another advantage of this method is that bias is reduced as the decision-making panel are not involved in research in the topic of concern, and that all panel members have an equal opportunity to influence outcomes (Halcomb et al., 2008). A disadvantage of not using experts to make the decisions is that there is a possibility that some meaning of data may be lost, and as the topic experts only have a limited time to present their evidence, not all evidence may be delivered. Also as the panel members meet, this method does not have the anonymity of the Delphi technique (Black, 2006).

Table 2 provides examples of applications of CDC.

Table 2: Examples of applications of CDC in Health

<table>
<thead>
<tr>
<th>Study</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sorrel et al 2009 USA (Sorrell et al., 2009)</td>
<td>Management of hepatitis B</td>
</tr>
<tr>
<td>Signore and Spong 2010 USA (Signore &amp; Spong, 2010)</td>
<td>Vaginal birth after Cesarean: New insights</td>
</tr>
<tr>
<td>Daviglus et al 2010 USA (Daviglus et al., 2010)</td>
<td>Preventing Alzheimer disease and cognitive decline</td>
</tr>
<tr>
<td>Berry et al 2011 Australia (Berry, Davidson, Nicholson, Pasqualotto, &amp; Rolls, 2011)</td>
<td>Oral hygiene in the critically ill</td>
</tr>
<tr>
<td>Wolff et al 2011 Germany (Wolff et al., 2011)</td>
<td>Chronic graft-versus-host disease</td>
</tr>
</tbody>
</table>
NOMINAL GROUP TECHNIQUE (NGT)

Delbecq and Van de Ven devised this technique in 1971 for committee decision making (Vernon, 2009). In this method the group consists of a small number of members, typically 6-9 people (Duncan, 2006), and the final views are an aggregation of the members’ views rather than a communal viewpoint (Black, 2006). The NGT method is conducted in several iterative stages over one session. The first stage consists of each panel member suggesting any relevant issues surrounding the topic. These suggestions are collated and used to develop a questionnaire covering all identified issues. The questionnaire is circulated and members are asked to rate their agreement on each suggestion using a Likert scale. Finally, the aggregated responses are distributed and members engage in a structured group discussion facilitated by an independent researcher. Each suggestion is discussed by the group and members record their judgements or level of agreement. Further discussion and voting may ensue until a group judgement is decided on (Black, 2006; Murphy et al., 1998).

Membership of the NGT group should include representation from the full range of people to which the guidelines will apply. This gives this technique the advantage of including patient opinions in the development of clinical guidelines (Rycroft-Malone, 2001). The technique may be used as part of a “modified Delphi” technique where the first rounds are completed by email and then the panel are bought together for a face-to-face discussion (Vakil, 2011). An advantage of the NGT is that each member is given an equal opportunity to generate and present suggestions, preventing individual members from either dominating the idea generation or leaving it to the rest of the group. Also as the generation of ideas and the discussion and evaluation phases are separated, a greater number of ideas may be potentially developed (Murphy et al., 1998). Limitations of the NGT are the small number of participants involved and the practicalities and cost of arranging at least one face-to-face meeting for all participants (Duncan, 2006). Due to the relatively small number of participants contributing their views, this technique has been criticised for its ability to be representative (Black, 2006). A major concern regarding the NGT is that it does not specifically allow for
integration of evidence from literature and thus it has been criticised for a lack of rigour (Vakil, 2011). To ensure greater scientific validity for this technique, any clinical recommendations should be developed using systematic reviews or meta-analysis and the expertise of key stakeholders to whom the guidelines may apply including clinicians, academics and patients (Rycroft-Malone, 2001). Table 3 provides examples of the applications of NGT.

**Table 3: Examples of applications of NGT.**

<table>
<thead>
<tr>
<th>Study</th>
<th>Purpose</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Netherlands</td>
<td>Selection of attributes for discrete choice experiments – drug treatment choice in osteoporosis</td>
<td>4-8 patients</td>
</tr>
<tr>
<td>(Hiligsmann et al., 2013)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canada</td>
<td>Design of an oral mucositis assessment instrument for use in children</td>
<td>9 health care professionals</td>
</tr>
<tr>
<td>(Tomlinson et al., 2009)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>Characteristics of dialysis important to patients and family caregivers</td>
<td>6 groups - 17 patients and 17 caregivers</td>
</tr>
<tr>
<td>(Morton, Tong, Webster, Snelling, &amp; Howard, 2011)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canada</td>
<td>Defining Quality Criteria for Online Continuing Medical Education Modules</td>
<td>9 clinical educators</td>
</tr>
<tr>
<td>(Shortt, Guillemette, Duncan, &amp; Kirby, 2010)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SELECTING AN APPROPRIATE METHOD**

The decision to use a particular research method should be based on the purpose of the study, the availability of scientific evidence in the field, time and cost factors, and the number of participants and the model of participant interaction (Halcomb et al., 2008). Initially a search of the current literature should be completed and all evidence summarised. The level of available evidence will help select the most appropriate method. The CDC is more appropriate for areas where higher, and more varied, levels of evidence are found (Halcomb et al., 2008). The Delphi or NGT are used when low, or conflicting, levels of evidence are available (Black, 2006). Selection of the “experts” is the next key step which has a direct impact on the credibility and reliability of the research findings (Baker et al., 2006). There is little consensus in the literature about what constitutes an “expert” however they are usually people who have considerable knowledge or experience in the specific field of study. There is potential for bias when selecting experts who are known to the researcher, although this is
sometimes unavoidable when studying small, very specific topics. Researchers should define what “expert” means in the context of their research and thus be able to justify their decision for the selection and rejection of panel members (Baker et al., 2006). The number of participants and whether they will meet face-to-face is another consideration. This will impact on the costs involved. The Delphi is the most cost effective method as it can be conducted completely via email and online.

A comparison of some of the features of these three research methods is given in Table 4 and a flowchart for selecting a method is provided (Figure 1).

Table 4: Comparison of consensus methods

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Delphi Technique</th>
<th>Consensus Development Conference</th>
<th>Nominal Group Technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>Postal or email surveys to assist in prioritisation of issues relating to policy and practice</td>
<td>Presentation of current evidence and subsequent discussion of issues relating to policy and practice</td>
<td>Generation and collation of ideas with subsequent discussion and voting on priorities</td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td>Distance</td>
<td>Local</td>
<td>Local</td>
</tr>
<tr>
<td>Time Frame</td>
<td>Several rounds conducted over months</td>
<td>One to three days</td>
<td>One day</td>
</tr>
<tr>
<td>Anonymity</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Panel size</td>
<td>Variable</td>
<td>6-12 panel members</td>
<td>6-9 panel members</td>
</tr>
<tr>
<td>Analysis</td>
<td>Variable Statistical and descriptive</td>
<td>Variable Majority voting and levels of agreement</td>
<td>Statistical Ranking</td>
</tr>
</tbody>
</table>
EVALUATION OF CONSENSUS GUIDELINES

In all areas of health, clinical consensus guidelines are increasingly being developed to inform and guide practice to ensure consistent, quality care. Clinicians must be able to evaluate these guidelines to decide whether to incorporate them into their day to day clinical practice (Vakil, 2011). The method used to develop clinical guidelines should be explicitly detailed and a formal research method should be used to ensure the guidelines reflect the available evidence and the views of experts in the area. A rigorous method helps reduce bias and increases guideline credibility (Halcomb et al., 2008). Several national government agencies have created online resources to assist clinicians on how to develop and evaluate clinical guidelines. Some also serve as an open access repository for clinical practice guidelines.
In Australia, the National Health and Medical Research Council (NHMRC) launched the Clinical Practice Guidelines Portal (National Health and Medical Research Council (NHMRC), 2010) in February 2010, and it now has over 300 guidelines registered. The NHMRC website includes information on how to develop clinical guidelines to the NHMRC standard (National Health and Medical Research Council (NHMRC), 2013). The website contains a link to the NHMRC policy on clinical guideline development and conflict of interest published in 2012, which aims to provide guidance and transparency in the declaration of interests. A lack of information on managing conflict of interest in the NHMRC principles of guideline development had been reported (Williams, Kevat, & Loff, 2011) and so this policy was developed in recognition that many experts involved in clinical guideline development have interests which may be conflicting and therefore should be appropriately managed and declared.

The U.S. Department of Health and Human Services, Agency for Healthcare Research and Quality (AHRQ), runs the National Guideline Clearinghouse which is a “public resource for evidence-based clinical practice guidelines” (Agency for Healthcare Research and Quality (AHQR), 2013). It has over 2,700 guideline summaries available and also enables the comparison of guidelines on similar topics (Agency for Healthcare Research and Quality (AHQR), 2013). In the UK, the National Institute for Health and Care Excellence (NICE) provides guidance and advice to improve health care and their website contains access to over 180 clinical guidelines and information on guideline implementation (National Institute for Health and Care Excellence (NICE)). The Canadian Institutes of Health Research funds the AGREE Enterprise website (Appraisal of Guidelines Research and Evaluation) (Canadian Institutes of Health Research, 2013). The original AGREE instrument was developed in 2003 to assess the quality of clinical guidelines using six quality domains: scope and purpose, stakeholder involvement, rigour of development, clarity of presentation, applicability, and editorial independence. The tool was updated in 2010 and the new AGREE
Il can be used to evaluate the process of practice guideline development, components of final recommendations and the quality of reporting (Brouwers et al., 2010).

CONCLUSION

Formal consensus methods are used widely in health to assist in the development of clinical practice guidelines and health policy. There are several methodologies available; each with its own advantages and disadvantages. The Delphi technique is the method of choice when participant anonymity is required and cost is a concern. Consensus development conferences are useful when there is a large, but conflicting amount of evidence in the literature. The nominal group technique is best for small groups of participants and when patient opinions are desirable. Researchers should choose the research method and the group participants carefully to ensure credibility and any outcomes should remain closely tied to evidence based literature to ensure rigour and credibility. Once consensus statements or guidelines have been developed, clinicians should carefully evaluate not only their outcomes, but also their method of development prior to incorporating into clinical practice.

REFERENCES


Versteeg, M., du Toit, L., & Couper, I. (2013). Building consensus on key priorities for rural health care in South Africa using the Delphi technique. *Glob Health Action, 6*, 19522. doi: [http://dx.doi.org/10.3402/gha.v6i0.19522](http://dx.doi.org/10.3402/gha.v6i0.19522)

