Consensus Procedures and Their Role in Pediatric Rheumatology

Nicolino Ruperto, MD, MPH, Silvia Meiorin, MD, Silvia Mirela Iusan, MD, Angelo Ravelli, MD, Angela Pistorio, MD, PhD, and Alberto Martini, MD

Corresponding author
Nicolino Ruperto, MD, MPH
IRCCS G. Gaslini, Pediatria II-PRINTO, University of Genoa, Largo Gaslini, 5, 16147 Genova, Italy.
E-mail: nicolaruperto@ospedale-gaslini.ge.it

Current Rheumatology Reports 2008, **10:**142–146
Current Medicine Group LLC ISSN 1523-3774

The Delphi Technique and Nominal Group Technique are two well-recognized consensus-formation methodologies specifically designed to combine judgments from a group of experts. The Delphi Technique utilizes a series of well-defined questionnaire-based surveys, whereas Nominal Group Technique is a structured face-to-face meeting designed to facilitate consensus. Consensus-formation techniques require that each step build on the results of the previous steps. In this review, we describe these techniques, how they work, and their practical application in pediatric rheumatology, where they have been widely used to develop the outcome measures of several chronic rheumatic diseases, including juvenile idiopathic arthritis, rheumatoid arthritis, systemic lupus erythematosus, and idiopathic inflammatory myopathies, as well as the classification criteria for juvenile systemic sclerosis and juvenile vasculitides.

Introduction

Medical research often involves situations in which varied quantitative or qualitative information should be pooled to summarize results and make decisions. Although quantitative information usually is summarized effectively with statistical techniques such as meta-analysis, qualitative research relies mainly on consensus methods to synthesize information. Indeed, given the diversity of opinion that any group of people may express on a given topic, methods to organize subjective judgments are needed. Group-based techniques to generate ideas and offer solutions were employed in social sciences first and have become commonplace in many other disciplines. In medicine, consensus methods can be particularly useful when published information is inadequate or nonexistent or when input from experts is necessary.

The two most widely used consensus-formation methodologies in the medical literature are the Delphi Technique (DT) and the Nominal Group Technique (NGT) [1]. Both techniques are designed to combine judgments from literature and/or a group of experts in a particular field. These techniques have been used widely to develop the outcome measures of several chronic rheumatic diseases, including juvenile idiopathic arthritis (JIA), rheumatoid arthritis, systemic lupus erythematosus (SLE), and idiopathic inflammatory myopathies, as well as the classification criteria for juvenile systemic sclerosis (SSc) and juvenile vasculitides. In this review, we describe these methods, how they work, and their practical applications in pediatric rheumatology, including examples from the literature.

Consensus Methods

The DT and NGT are two methods commonly used to synthesize information from conflicting evidence. Consensus methods are primarily concerned with deriving quantitative estimates through qualitative approaches. This means greater flexibility, because it allows for a wider range of study types to be considered than is usual in statistical reviews. In general, the DT concentrates on measuring consensus via mail questionnaire, whereas the NGT can be used to develop consensus within a group of people around a table.

The Delphi Technique

The DT takes its name from the Delphic oracle's skills of interpretation and foresight and was developed in the 1970s [2,3]. The methodology enables consensus among a group of people (expert and lay people) who cannot meet together physically. Delphi is an iterative multistage method that allows the use of written responses without the need for face-to-face meetings. It is usually conducted anonymously, avoiding problems associated with hostile or intimidating

interpersonal interactions. Therefore, consensus is gathered from a selected panel regarding a topic about which there is little available evidence, without the need for panelists to meet face to face. The main advantage of the DT is that participation from different geographic areas requires relatively little expense because there is no need for face-to-face meetings. Participants are generally contacted by mail, but the Internet is used increasingly often, which also helps facilitate international research.

Delphi is essentially a series of questionnaires involving several steps, each of which is based on the results of the previous step; usually, panel members are fed information through a series of rounds of questionnaires. The process stops when consensus is reached or when sufficient information exists to proceed with the project.

The Delphi process can be divided into several simplified steps. First, develop the Delphi question. The responders must understand the problem at hand, and the first questionnaire usually poses a broad question to elicit different opinions from the panelists. Questions are initially open-ended and seek individual responses. In an example described in more detail later, physicians were asked to list the variables they used in current clinical practice to evaluate response to therapy in several pediatric rheumatic diseases [4,5]. Before sending out the final version of the questionnaire, it is wise to test one or more versions. In general, the survey should be as simple as possible, ideally no longer than 1 page.

Second, select the responders. The panel of responders must be adequate in number and expertise. Ideally, they are motivated and interested in the problem to be solved, which will help ensure adequate response rates to the survey.

Third, analyze responses and develop subsequent questionnaire(s). Once received, the replies from the panelist should be analyzed with simple descriptive statistics (eg, median with interquartile ranges) to generate a series of statements, which are compiled into a second questionnaire and sent back to the individual participants. In the second survey, the panelists should be provided with the overall results and possibly his/her previous reply. Often, in the later questionnaires, responders are asked to rank the items identified with the first survey by importance (eg, on a scale of 1 to 10, with 10 being the most important) in order to limit the possible scenarios or determine if they agree/disagree with a given statement. The rankings are then summed and items ranked highest are considered for the following questionnaires.

Lastly, develop a final report summarizing the goals, process, and results. The DT has been used widely in pediatric rheumatology (as discussed in later examples) and other medical specialties [6–8].

The Nominal Group Technique

The NGT combines quantitative and qualitative data collection in a group setting. It avoids some of the problems of other group methods (eg, brainstorming) but does not have the limitations of a more informal meeting, which might be dominated by powerful or influential individuals.

The NGT uses a highly structured meeting to gather information from relevant experts about a given issue. It consists of two rounds, in which panelists rate, discuss, and then re-rate a series of items or questions focusing on a single goal. The method was developed in the United States in the 1960s and has been applied often in pediatric rheumatology (as discussed in more detail later) and other fields [6]. A nominal group meeting is facilitated by an expert on the topic or a credible nonexpert.

Preparation for the NGT meeting includes setting up one or more meeting rooms to accommodate about six or seven people plus the moderator around a "U"-shaped table. If more than one table is used, participants can be allocated to the tables at random. Tables should be separated to avoid the influence of activity at one table on the activity at others. Facilitators should provide a flip chart (or a computer with a wide-screen monitor) situated close to moderators, cards for voting exercises, and paper and pencil for each participant. An introductory lecture should be prepared to describe the goals of the meeting, how and why the group was formed, and the NGT process.

The NGT can be divided into simple steps. First, the process requires the silent generation of ideas. Participants spend several minutes writing down their views about the topic in question. In this phase, participants should work autonomously in a quiet environment, and the moderator should sanction anyone who interrupts the activities.

Second, there is the round-robin recording of ideas. The moderator records the ideas of each participant by asking each person around the table for their input. Ideas are recorded on the flip chart and each participant contributes to the generation list. The written record stimulates further thought among the participants and offers the benefit of showing people how the problem is going to be solved. Similar suggestions can be grouped together, where appropriate. This recording can proceed for several rounds, and the process requires that a new person speak first at each round (ie, the first round begins with an idea from the first person to the left of the moderator, the second round with the second person to the left of the moderator, etc.). In this way, everyone gets the opportunity to speak first, avoiding the undue influence of strong personalities.

The third step involves serial discussion and voting. Each idea written on the flip chart should be discussed, following the round-robin method previously described. At the end of the discussion, participants vote on the importance of each item, usually ranked on a scale from 0 to 5, with 5 being the most important item (or alternatively, for a more fine-tuned result, from 0 to 10). Voting should use the cards rather than a show of hands, open discussion, or other forms in order to avoid the problem of social pressure. Ranking for each item should be summed and presented to the panel using descriptive

statistics (eg, medians and interquartile ranges or other graphical representation); then, a second round of discussion on the top items can be implemented. Feeding back the group's response enables participants to consider their initial ranking in relation to their colleagues' assessments. Consensus is reached when 70% to 80% of the participants agree on given items. Items for which consensus is not achieved can be discarded or reformulated. When the meeting is held at more than one table discussing the same items, a convergent consensus might be required from all tables. In this case, items for which consensus is not achieved should be discussed in a plenary session with the participants from every table.

The method can be adapted. For example, it has been conducted as a single meeting or by mail for the first stage and face-to-face for the discussion and re-rating steps. Some nominal group meetings have incorporated a detailed review of literature as background material for the topic under discussion. Alongside the consensus process, there may be a nonparticipant observer collecting qualitative data on the nominal group.

Of course, the existence of a consensus does not mean that the "correct" answer has been found, because the NGT and DT are not replacements for scientific reviews of literature. Rather, they are methods for identifying opinions and areas of disagreement.

Who should participate

Usually the people who designed the Delphi questionnaires or set up an NGT meeting do not participate in the panel discussion. The panel members should be contacted beforehand to evaluate their interest to participate and to inform them of the incoming surveys or upcoming meeting. In theory, the sample can be larger with the DT than with NGT, but this initial number of participants might decrease progressively with a decrease in the response rate in each additional step. Interest in participation is essential for successful implementation; however, this should be balanced with relative impartiality. Experts in the field are ideal participants, but the inclusion of lay people (eg, general practitioners) could help provide a composite mix of opinions. Clearly, the potential for bias exists in participant selection, because the exact composition of the panel can affect the results obtained [9].

Examples from the Pediatric Rheumatology Literature

American College of Rheumatology Pediatric 30 criteria to evaluate response to JIA therapy

In 1993, to standardize the assessment of clinical response in JIA, a Delphi questionnaire was mailed to the 16 members of an ad hoc Advisory Committee composed of members of the Rheumatology Section of the American Academy of Pediatrics, the Pediatric Sections of the American College of Rheumatology (ACR) and the Arthritis Foundation, Outcomes Measures in Rheumatology (OMERACT) participants, and private and academic practitioners. The questionnaire asked participants to select and rank the variables to evaluate response to therapy in JIA, from a list of 25 variables derived from the literature. Next, in 1994, a consensus conference was convened in North America to revise the performance characteristics of each candidate variable, and then, using NGT, attendees developed a preliminary core set of response variables that included six end points [4].

Following this first conference, organizers felt it was necessary to hold a broader international consensus about the core set and its use to define improvement. A second Delphi questionnaire was mailed to a larger, more international sample of 198 practitioners to obtain their reaction to the core set and its proposed use. The questionnaire was returned by 140 (71%) physicians (88 European, 52 North American). This second Delphi questionnaire was designed to provide preliminary data for the organization of a second International Consensus Conference. The 1996 NGT-based conference in Italy was attended by 21 pediatric rheumatologists from 14 different countries. The goal of the meeting was to decide on a preliminary definition of improvement employing the core set of end points using a combination of statistical and consensus-formation techniques.

This project demonstrates how the combination of DT, NGT, and statistics inferences can be used to solve a complex problem such as the definition of criteria to evaluate response to therapy in JIA. Of note, this definition was endorsed by the ACR and is now called the ACR Pediatric 30 JIA criteria. It has been adopted by the US Food and Drug Administration and the European Medicines Agency and applied and validated in several clinical trials [10–13].

Criteria for clinical remission in JIA

A Delphi serial questionnaire consensus-formation approach gathered criteria used by pediatric rheumatologists to define clinical remission in oligoarticular (persistent and extended), rheumatoid factor–positive and negative polyarticular, and systemic JIA. Results from sequential questionnaires provided an agenda for an NGT conference to reach consensus on unresolved questions [14]. The consensus-based criteria were further validated with an evidence-based approach on patient data [15], and an ongoing Paediatric Rheumatology International Trials Organisation (PRINTO) clinical trial to establish the best time to stop methotrexate in patients with clinical remission of JIA. This project shows how DT and NGT can be used to solve a complex problem such as the definition of criteria for clinical remission in JIA.

The juvenile SLE and juvenile dermatomyositis project

Using an approach similar to the previously described project, PRINTO [16] and the North America–based Pediatric Rheumatology Collaborative Study Group (PRCSG) [10,17] committees set up a project to establish a core set of outcome measures and a definition of improvement for the evaluation

of response to therapy, as well as a core set for the evaluation of disease damage for juvenile SLE (JSLE) and juvenile dermatomyositis (JDM).

This multinational effort was divided into three phases. First, two sequential Delphi e-mail surveys were conducted to select and rank the variables used in routine clinical practice to assess treatment response in JSLE or JDM. The surveys involved 277 practitioners with an 80% response rate. They were designed to obtain information on the variables used in daily clinical practice to sensitively and practically measure disease activity and damage.

Following the e-mail surveys, organizers held a 2-day consensus conference, which was attended by 40 experienced pediatric rheumatologists from 34 different countries [18]. Through the NGT, meeting participants sought to reach a consensus on the appropriate domains and variables to be included in the preliminary JSLE and JDM disease activity and damage core sets. Before the meeting, participants received a booklet containing relevant articles about the potential outcome measures for JSLE and copies of the instruments to be analyzed during the consensus conference. At the meeting, after introductory lectures concerning the characteristics and scoring systems of the outcome measures most commonly used in the assessment of activity and damage, attendees were randomly assigned to three working groups. Each group worked in a separate room with a moderator. For every exercise, attendees were first asked to work individually and then to express their opinion in a guided discussion. At the end of the work, the results were pooled. An 80% consensus from all 40 attendees was required to consider each problem as solved. During the meeting, five NGT exercises were conducted [5].

The second phase of the project attempted to formally validate the preliminary JSLE disease activity core set for the evaluation of response to therapy through prospective, large-scale data collection among the members of the PRINTO/PRCSG networks [19,20]. The third phase was dedicated to defining improvement in order to evaluate response to therapy in JSLE and JDM using a procedure similar to that previously described for JIA [21••].

This project demonstrated how the DT and NGT can be combined with evidence-based development of criteria through a large-scale data collection and classic inferential statistics. Of note, this project has been endorsed by the ACR [20,21••].

Similar approaches have been used to derive a core set of outcome measures to develop preliminary definitions of improvement for adult and juvenile myositis as composite end points for therapeutic trials [22], as well as for the international consensus guidelines for trials of therapies in the idiopathic inflammatory myopathies [23].

Disease activity measures for systemic JIA

A Delphi survey process in two steps was used to reach consensus on disease activity measures for children with systemic JIA using an international pool of 187 pediatric rheumatologists [24]. From an initial list of 2607 items, the DT helped identify 29 items as the most important and most frequently seen indicators of active disease. The authors state in their paper that the next step will be the testing of the measurement properties of these items in order to develop a disease activity tool for clinical trials.

Classification criteria for juvenile systemic sclerosis and juvenile vasculitides

In order to develop criteria for the classification of juvenile SSc, a three-phase study was designed. In the first phase, data were collected regarding actual juvenile SSc patients' signs and symptoms that are useful to define particular organ involvement. In the second phase, two Delphi surveys were employed to select the essential parameters for the classification of juvenile SSc and for preparation of a set of provisional classification criteria. In the third phase, organizers implemented a consensus conference with NGT and statistical evaluations to derive the provisional classification criteria for juvenile SSc [25••].

A similar approach has been followed for the new classification criteria for childhood vasculitides. In this case, the initial classification criteria have been prepared with an NGT-based consensus conference [26]. Data collection for final validation with a second NGT consensus conference is ongoing.

Conclusions

The DT and NGT are two well-recognized consensus formation methodologies that have been instrumental in conducting several international research projects. A more widespread use of the techniques will help further improve the pediatric rheumatology field.

Acknowledgment

Drs. Meiorin and Iusan are recipients of scholarships from the European League Against Rheumatism. The authors have written this paper on behalf of PRINTO.

Disclosures

The authors have reported no potential conflicts of interest relevant to this article.

References and Recommended Reading

Papers of particular interest, published recently, have been highlighted as:

- • Of importance
- •• Of major importance

1. Delbecq AL, Van de Ven AH, Gustafson DH: *Group Techniques for Program Planning: A Guide to Nominal Group and Delphi Processes,* edn 1. Glenview, IL: Scott, Foresman and Company; 1975.

2. Dalkey NC: *Delphi*. Santa Monica, CA: Rand Corporation; 1967.
3. Helmer O: *Analysis of the Future: the Delphi Method*. Santa Monica, CA: Rand Corporation; 1967.
4. Giannini EH, Ruperto N, Ravelli A, et al.: **Preliminary definition of improvement in juvenile arthritis.** *Arthritis Rheum* 1997, **40:**1202–1209.
5. Ruperto N, Ravelli A, Murray KJ, et al.: **Preliminary core sets of measures for disease activity and damage assessment in juvenile systemic lupus erythematosus and juvenile dermatomyositis.** *Rheumatology (Oxford)* 2003, **42:**1452–1459.
6. Jones J, Hunter D: **Consensus methods for medical and health services research.** *BMJ* 1995, **311:**376–380.
7. van Teijlingen E, Pitchforth E, Bishop C, Russell E: **Delphi method and nominal group technique in family planning and reproductive health research.** *J Fam Plann Reprod Health Care* 2006, **32:**249–252.
8. Keeney S, Hasson F, McKenna HP: **A critical review of the Delphi technique as a research methodology for nursing.** *Int J Nurs Stud* 2001, **38:**195–200.
9. McKee M, Priest P, Ginzler M, Black N: **How representative are members of expert panels?** *Qual Assur Health Care* 1991, **3:**89–94.
10. Lovell DJ, Giannini EH, Reiff A, et al.: **Etanercept in children with polyarticular juvenile rheumatoid arthritis. Pediatric Rheumatology Collaborative Study Group.** *N Engl J Med* 2000, **342:**763–769.
11. Ruperto N, Lovell DJ, Cuttica R, et al.: **A randomized, placebo-controlled trial of infliximab plus methotrexate for the treatment of polyarticular-course juvenile rheumatoid arthritis.** *Arthritis Rheum* 2007, **56:**3096–3106.
12. Ruperto N, Lovell DJ, Prieur AM, et al.: **Efficacy and safety of abatacept in children and adolescents with active juvenile idiopathic arthritis (JIA).** *Ann Rheum Dis* 2006, **65(Suppl 11):**248.
13. Ruperto N, Lovell DJ, Goodman S, et al.: **48-week data from the study of adalimumab in children with juvenile rheumatoid arthritis (JRA).** *Ann Rheum Dis* 2006, **65 (Suppl 11):**56.
14. Wallace CA, Ruperto N, Giannini E, et al.: **Preliminary criteria for clinical remission for select categories of juvenile idiopathic arthritis.** *J Rheumatol* 2004, **31:**2290–2294.
15. Wallace CA, Huang B, Bandeira M, et al.: **Patterns of clinical remission in select categories of juvenile idiopathic arthritis.** *Arthritis Rheum* 2005, **52:**3554–3562.
16. Ruperto N, Martini A: **International research networks in pediatric rheumatology: the PRINTO perspective.** *Curr Opin Rheumatol* 2004, **16:**566–570.
17. Giannini EH, Brewer EJ, Kuzmina N, et al.: **Methotrexate in resistant juvenile rheumatoid arthritis. Results of the U.S.A.-U.S.S.R. double-blind, placebo-controlled trial. The Pediatric Rheumatology Collaborative Study Group and the Cooperative Children's Study Group.** *N Engl J Med* 1992, **326:**1043–1049.
18. Martini A, Ruperto N, Lovell DJ, et al.: *An international consensus conference to define core sets of outcome measures for juvenile systemic lupus erythematosus and juvenile dermatomyositis. Proceedings of the Conference, March 31–April 3, 2001.* Pavia, Italy: Paediatric Rheumatology International Trials Organization (PRINTO); 2001.
19. Ruperto N, Ravelli A, Cuttica R, et al.: **The Pediatric Rheumatology International Trials Organization criteria for the evaluation of response to therapy in juvenile systemic lupus erythematosus: prospective validation of the disease activity core set.** *Arthritis Rheum* 2005, **52:**2854–2864.
20. Ruperto N, Ravelli A, Pistorio A, et al.: **The provisional Paediatric Rheumatology International Trials Organisation/ American College of Rheumatology/European League Against Rheumatism Disease activity core set for the evaluation of response to therapy in juvenile dermatomyositis: a prospective validation study.** *Arthritis Rheum* 2007, **59:**4–13.
21.•• Ruperto N, Ravelli A, Oliveira S, et al.: **The Pediatric Rheumatology International Trials Organization/American College of Rheumatology provisional criteria for the evaluation of response to therapy in juvenile systemic lupus erythematosus: prospective validation of the definition of improvement.** *Arthritis Rheum* 2006, **55:**355–363.
This paper describes how the combination of statistical inferences and NGT can be used to solve a complex problem like the evaluation of response to therapy in JSLE.
22. Rider LG, Giannini EH, Brunner HI, et al.: **International consensus on preliminary definitions of improvement in adult and juvenile myositis.** *Arthritis Rheum* 2004, **50:**2281–2290.
23. Oddis CV, Rider LG, Reed AM, et al.: **International consensus guidelines for trials of therapies in the idiopathic inflammatory myopathies.** *Arthritis Rheum* 2005, **52:**2607–2615.
24. Ramanan AV, Schneider R, Batthish M, et al.: **Developing a disease activity tool for systemic-onset juvenile idiopathic arthritis by international consensus using the Delphi approach.** *Rheumatology (Oxford)* 2005, **44:**1574–1578.
25.•• Zulian F, Woo P, Athreya BH, et al.: **The Pediatric Rheumatology European Society/American College of Rheumatology/European League against Rheumatism provisional classification criteria for juvenile systemic sclerosis.** *Arthritis Rheum* 2007, **57:**203–212.
This paper describes how the combination of statistical inferences, DT, and NGT can be used to solve a complex problem like the classification criteria for juvenile SSc.
26. Ozen S, Ruperto N, Dillon MJ, et al.: **EULAR/PReS endorsed consensus criteria for the classification of childhood vasculitides.** *Ann Rheum Dis* 2006, **65:**936–941.