

The Delphi method: an efficient procedure to generate knowledge

Johann Steurer

Published online: 11 June 2011
© ISS 2011

Physicians base the recommendations and decisions they make in daily practice on the results of clinical research findings. Results from randomized trials and systematic reviews inform physicians about the effects of treatments, and cross-sectional studies inform them about the accuracy of diagnostic tests. Researchers and clinical experts strive to increase the knowledge and to generate a solid *knowledge base of medicine* to support physicians in caring for patients. However, not all relevant questions in medicine can be answered by conventional research methods based on patient data, and in some instances there are more effective methods for creating knowledge that are pertinent to patient care. For example, conventional methods may not be useful when there is conflicting scientific evidence, when accurate information is not available, and when judgmental information is required. An alternative for such situations is to ask experts what they think and collect their knowledge in a structured and systematic way.

During the last two decades, “evidence-based medicine” has become the new paradigm in medicine. In parallel, expertise has been discredited as a source of knowledge, ranking low in the hierarchy of evidence. Certainly some ‘experts’ fail to resist temptations that influence what they say and write. However, this does not justify mistrust of expert knowledge in general. Expert knowledge gleaned from years of experience in the field is a valuable and underrated source of knowledge that should be made available to other, less experienced physicians. There are

several methods for garnering the knowledge of experts, one of which is the Delphi method.

The term ‘Delphi method’ goes back to the oracle of Delphi in ancient Greece, which was consulted on matters that ranged from public policy to personal affairs. In the 1950s, the deity Apollo, who was thought to speak through this oracle, was replaced by secular experts after a series of studies that strove to reach the most reliable consensus among a group of experts was conducted by the RAND cooperation [1]. “Project Delphi,” financed by the U.S. Air Force, was the name given to the first project for forecasting technological developments. Similar methods include the nominal group technique and the consensus conference. The nominal technique and the consensus conference use highly structured meetings of experts to collect information using a format that includes a discussion among the experts.

The Delphi method consists of “a series of structured group processes, each referred to as a round, to survey expert opinion and reach a group response,” [2]. Opinions, beliefs, and judgments are collected and organized in a systematic fashion that focuses primarily on consensus but also on dissenting views. The core principles include individual feedback regarding a given topic or questionnaire; assessment of the group’s judgment or views; opportunities for individuals to revise their views; and anonymity for individual responses.

The Delphi method has two important advantages compared to the nominal technique or the traditional consensus conference. One advantage is that experts from all over the world today can participate via electronic communication and the second is the anonymous response format, which allows experts to express opinions or beliefs without being influenced or governed by other experts. This anonymity reduces the effects of group interaction that lead

J. Steurer (✉)
Horten Centre for Patient-Oriented Research and Knowledge
Transfer,
Raemistrasse 100,
CH 8091, Zurich, Switzerland
e-mail: johann.steurer@usz.ch

some to follow the opinion of the majority or that of the greatest assumed authority. Anonymity also protects experts from losing face if their opinions are dissenting and allows them to change their positions during the survey, even if they had held those views for decades.

The Delphi method is a survey technique that has been applied to health research within the fields of health technology assessment, health education, diagnostic criteria development, indicator selection for quantifying medical care quality, opinion surveys for developing appropriateness criteria for medical interventions and establishment of research agendas and research priorities for various disciplines [3–6].

A Delphi survey has three main tasks: First, defining and describing the topic and preparing one or more questions to send to the experts; second, selection of a panel of participating experts; and third, organizing and running the survey, which involves two or more rounds. After the survey topic is clarified, the next step is to formulate the questions and determine the information the researchers would like to obtain. For example, diagnosis of lumbar spinal stenosis is a challenging task—it is unclear which radiological criteria should be used to describe and diagnose this condition. The first round of the survey can be designed at least two ways. One way is to ask experts which parameters and points they use for measuring the distances or square areas in the spinal canal. The other way is to prepare a list of published parameters and then ask the experts if the list is comprehensive and whether they would include additional parameters.

One of the most critical aspects of the Delphi survey is the selection of qualified experts. It is common to contact one or two experts in the field and ask them to suggest other experts. This is a convenient approach, but can lead to selection bias, as the experts will most likely nominate colleagues who share their views. Another more elaborate approach is to include only experts with more than, for example, five publications on the chosen topic in peer-reviewed journals during the last 3 years. Note that a Delphi survey does not depend *per se* on a sample that is representative of any particular population. That is, there are no specified rules regarding the selection of participants or the number of participants. Rather, the Delphi method depends on the information the researchers want to gather. For example, 20 panelists may be adequate for the development of diagnostic indicators, whereas the development of a research agenda for colorectal surgery can and should involve inviting all members of the corresponding medical society to participate. If the results of a survey will be used to shape public policy regarding, for example,

priorities in the health care system, a representative sample of the population should participate as well.

The third task is collection of participants' opinions and beliefs. This can either be done using paper and pencil response or via the Internet. There are several Web-based applications (e.g., www.surveymonkey.com) for conducting surveys. In general, the survey is accomplished in two or three rounds. In the first round, a single question, a questionnaire, or, in the lumbar spinal stenosis example, a list of radiological parameters to describe stenosis is sent to the participants. An answer, or completion of a list, is requested from each participant. The feedback from the participants is collected and, depending on the topic, question(s) are reformulated, new questions are added or a list of items is updated and adapted.

In the second round, participants rank their agreement with statements or weigh the relevance of information. In the Delphi survey on radiological signs of lumbar spinal stenosis, for example, participants may get an updated list in the second round and rank the diagnostic relevance of each parameter on a visual analog scale that ranges from 0 (completely irrelevant) to 10 (very relevant).

Whether a third round is necessary depends on the topic and design of the survey and on the desired result. If the aim is to reach consensus among all of the experts, it is likely that more than three rounds will be necessary; however, most of the time, full consensus is not the aim of the survey. Delphi surveys can be used to measure the degree of consensus or degree of dissent. The results of a Delphi survey are reported in terms of medians or means and interquartile ranges or standard deviations. Statistical measures, like Cronbach's α , can be used to quantify the homogeneity or consistency of opinion among the experts [3].

Critics may argue that the results of such surveys are collections of personal beliefs and opinions. This may be a limitation of such a survey; therefore careful selection of participants is of greatest importance to obtain sensible and trustworthy results. Another limitation may be that the results themselves cannot be discussed within the group of experts. Sometimes such a discussion might lead to different results. However, as noted above, group discussions have their own shortcomings.

Purists, always believing only in the results of empirical studies, will distrust the results of Delphi surveys. In general, the Delphi method is used only when scientific evidence is either absent or contradictory, and judgmental information is necessary. For such situations, and there are many of them in medicine, collective expert opinions and beliefs are helpful and necessary to inform both patients and physicians. One premise underlying the Delphi method

is that such collective beliefs are generally more trustworthy than the beliefs of a single person.

References

1. Dalkey N, Helmer O. An experimental application of the Delphi method to the use of experts. *Manage Sci.* 1963;9:458–67.
2. Burt CG, Cima RR, Koltun WA, Littlejohn CE, Ricciardi R, Temple LK, et al. Developing a research agenda for the American Society of Colon and Rectal Surgeons: results of a Delphi approach. *Dis Colon Rectum.* 2009;52:898–905.
3. Graham B, Regehr G, Wright JG. Delphi as a method to establish consensus for diagnostic criteria. *J Clin Epidemiol.* 2003;56:1150–6.
4. Gagliardi AR, Fleshner N, Langer B, Stern H, Brown AD. Development of prostate cancer quality indicators: a modified Delphi approach. *Can J Urol.* 2005;12:2808–15.
5. Park RE, Fink A, Brook RH, Chassin MR, Kahn KL, Merrick NJ, et al. Physician ratings of appropriate indications for six medical and surgical procedures. *Am J Public Health.* 1986;76:766–72.
6. Phitayakorn R, Delaney CP, Reynolds HL, Champagne BJ, Heriot AG, Neary P, et al. Standardized algorithms for management of anastomotic leaks and related abdominal and pelvic abscesses after colorectal surgery. *World J Surg.* 2008;32:1147–56.