

wherever he or she is via a briefcase sized laptop computer. In this particular case, however, such equipment probably would not have helped.

The emergency medical kit provided by British Airways includes 88 items and is suitable for most medical emergencies—with an emphasis on cardiac drugs and delivering babies.⁴ Unfortunately its surgical equipment is pretty sparse. This is justifiable as most people suffering a surgical emergency do not come to harm if treatment is delayed for one or two hours. A suitable local anaesthetic would, however, have been a helpful addition. This has now been addressed by the providers of the M5 medical emergency kit, and a 20 ml multidose vial of lignocaine 2% will be included in future (personal communication, Aeromedic Innovations, London).

Traumatic pneumothorax has not previously been reported as presenting during a commercial flight. Several cases are on record of patients with such a condition being evacuated by air after receiving treatment for their thoracic trauma in conditions of warfare

or civil disastro.⁵ Survival is good as long as thoracic drainage is established before flight.⁶

This case shows the need for doctors to be adaptable to work in very strange environments dealing with conditions they do not normally treat and with unfamiliar equipment. Innovation in the use of the materials to hand to enable adequate chest drainage may well have saved the patient's life.

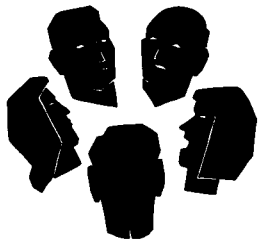
We thank Dr Sandra Mooney and Dr Michael Davies of British Airways health services for their help in the preparation of this paper.

- 1 Skjenna OW, Evans JF, Moore M, Thibeault C, Tucker AG. Helping patients travel by air. *Can Med Ass J* 1991;144:287-93.
- 2 Cummings RO, Schubach JA. Frequency and types of medical emergencies among commercial air travelers. *JAMA* 1989;261:1295-9.
- 3 Cottrell JJ. *Aviation auscultation*. A 1990;263:233.
- 4 Davies DM. Medical kits on airliners. *Lancet* 1991;337:1353.
- 5 Haid MM, Paladini P, Maccherini M, Di Bisceglie M, Biagi G, Gotti G. Air transport and the fate of pneumothorax in pleural adhesions. *Thorax* 1992;47:833-4.
- 6 Demartines N, Kiener A, Scheidegger D, Harder F. Drainage thoracique sur les lieux d'accidents. *Helv Chir Acta* 1990;57:273-7.

Qualitative Research

Consensus methods for medical and health services research

Jeremy Jones, Duncan Hunter



This is the sixth in a series of seven articles describing non-quantitative techniques and showing their value in health research

Health providers face the problem of trying to make decisions in situations where there is insufficient information and also where there is an overload of (often contradictory) information. Statistical methods such as meta-analysis have been developed to summarise and to resolve inconsistencies in study findings—where information is available in an appropriate form. Consensus methods provide another means of synthesising information, but are liable to use a wider range of information than is common in statistical methods, and where published information is inadequate or non-existent these methods provide a means of harnessing the insights of appropriate experts to enable decisions to be made. Two consensus methods commonly adopted in medical, nursing, and health services research—the Delphi process and the nominal group technique (also known as the expert panel)—are described, together with the most appropriate situations for using them; an outline of the process involved in undertaking a study using each method is supplemented by illustrations of the authors' work. Key methodological issues in using the methods are discussed, along with the distinct contribution of consensus methods as aids to decision making, both in clinical practice and in health service development.

Defining consensus and consensus methods

Quantitative methods such as meta-analysis have been developed to provide statistical overviews of the results of clinical trials and to resolve inconsistencies in the results of published studies. Consensus methods are another means of dealing with conflicting scientific evidence. They allow a wider range of study types to be considered than is usual in statistical reviews. In addition they allow a greater role for the qualitative assessment of evidence (box 1). These methods, unlike those described in the other papers in this series, are primarily concerned with deriving quantitative estimates through qualitative approaches.

The aim of consensus methods is to determine the

Box 1—Features of consensus methods

Anonymity	To avoid dominance; achieved by use of a questionnaire in Delphi and private ranking in nominal group
Iteration	Processes occur in "rounds", allowing individuals to change their opinions
Controlled feedback	Showing the distribution of the group's response (indicating to each individual their own previous response in Delphi)
Statistical group response	Expressing judgment using summary measures of the full group response, giving more information than just a consensus statement

Adapted from Pill¹ and Rowe²

extent to which experts or lay people agree about a given issue. They seek to overcome some of the disadvantages normally found with decision making in groups or committees, which are commonly dominated by one individual or by coalitions representing vested interests. In open committees individuals are often not ready to retract long held and publicly stated opinions, even when these have been proved to be false.

The term "agreement" takes two forms, which need to be distinguished: firstly, the extent to which each respondent agrees with the issue under consideration (typically rated on a numerical or categorical scale) and, secondly, the extent to which respondents agree with each other, the consensus element of these studies (typically assessed by statistical measures of average and dispersion).

Application

The focus of consensus methods lies where unanimity of opinion does not exist owing to a lack of

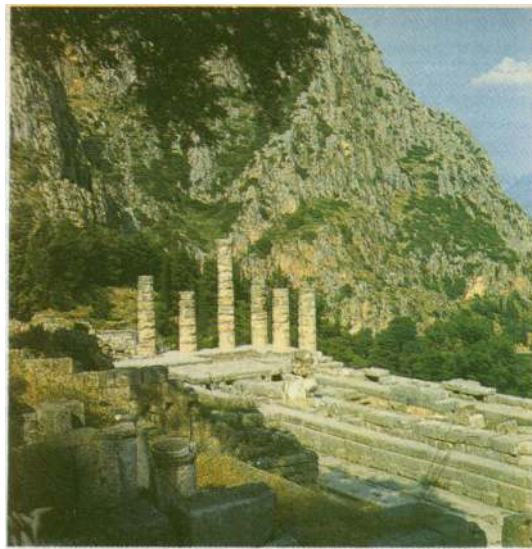
Nuffield Community Care Studies Unit, Department of Epidemiology and Public Health, University of Leicester, Leicester LE1 7RH

Jeremy Jones, lecturer in health economics

Health Services Research Unit, Department of Public Health and Policy, London School of Hygiene and Tropical Medicine, London WC1E 7HT
Duncan Hunter, research fellow

Correspondence to: Dr Jones.

BMJ 1995;311:376-80



MICHAEL HOLFORD

The Delphi process takes its name from the Delphic oracle's skills of interpretation and foresight

scientific evidence or where there is contradictory evidence on an issue. The methods attempt to assess the extent of agreement (consensus measurement) and to resolve disagreement (consensus development).

The three best known consensus methods are the Delphi process, the nominal group technique (also known as the expert panel), and the consensus development conference. Each of these methods involves measuring consensus, and the last two methods are also concerned with developing consensus. The consensus development conference will not be covered in this paper because it requires resources beyond those at the disposal of most researchers (unlike the other two methods), is commonly organised within defined programmes (for example, by the King's Fund in Britain and the National Institutes of Health in the United States), and has been discussed at length elsewhere.³⁴

The methods described

THE DELPHI PROCESS

The Delphi process takes its name from the Delphic oracle's skills of interpretation and foresight and proceeds in a series of rounds as follows:

- Round 1: Either the relevant individuals are invited to provide opinions on a specific matter, based on their knowledge and experience, or the team undertaking the Delphi expresses opinions on a specific matter and selects suitable experts to participate in subsequent questionnaire rounds;
- These opinions are grouped together under a limited number of headings and statements drafted for circulation to all participants on a questionnaire;
- Round 2: Participants rank their agreement with each statement in the questionnaire;
- The rankings are summarised and included in a repeat version of the questionnaire;
- Round 3: Participants rerank their agreement with each statement in the questionnaire, with the opportunity to change their score in view of the group's response;
- The rerankings are summarised and assessed for degree of consensus: if an acceptable degree of consensus is obtained the process may cease, with final results fed back to participants; if not, the third round is repeated.

Figure 1 shows an example of this process for a Delphi study undertaken by one of the authors (JJ). In addition to scoring agreement with statements, respondents are commonly asked to rate the confidence or certainty with which they express their opinions.

The Delphi technique has been used widely in health research within the fields of technology assessment,⁷⁻¹⁰ education and training¹¹⁻¹⁴ and priorities and information,¹⁵⁻¹⁷ and in developing nursing and clinical practice.¹⁹⁻²¹ It enables a large group of experts to be contacted cheaply, usually by mail with a self administered questionnaire (though computer communications have also been used), with few geographical limitations on the sample. Some situations have included a round in which the participants meet to discuss the process and resolve uncertainty or any ambiguities in the wording of the questionnaire.

THE NOMINAL GROUP TECHNIQUE

The nominal group technique uses a highly structured meeting to gather information from relevant experts (usually 9-12 in number) about a given issue. It consists of two rounds in which panellists rate, discuss, and then rerate a series of items or questions. The method was developed in the United States in the 1960s and has been applied to problems in social services, education, government, and industry.²² In the context of health care the method has most commonly been used to examine the appropriateness of clinical interventions²³⁻²⁷ but has also been applied in education and training,²⁸⁻³⁰ in practice development,³¹⁻³³ and for identifying measures for clinical trials.³⁴⁻³⁶

A nominal group meeting is facilitated either by an expert on the topic³⁷ or a credible non-expert³⁸ and is structured as follows:

- Participants spend several minutes writing down their views about the topic in question;
- Each participant, in turn, contributes one idea to the facilitator, who records it on a flip chart;
- Similar suggestions are grouped together, where appropriate. There is a group discussion to clarify and evaluate each idea;
- Each participant privately ranks each idea (round 1);
- The ranking is tabulated and presented;
- The overall ranking is discussed and reranked (round 2);

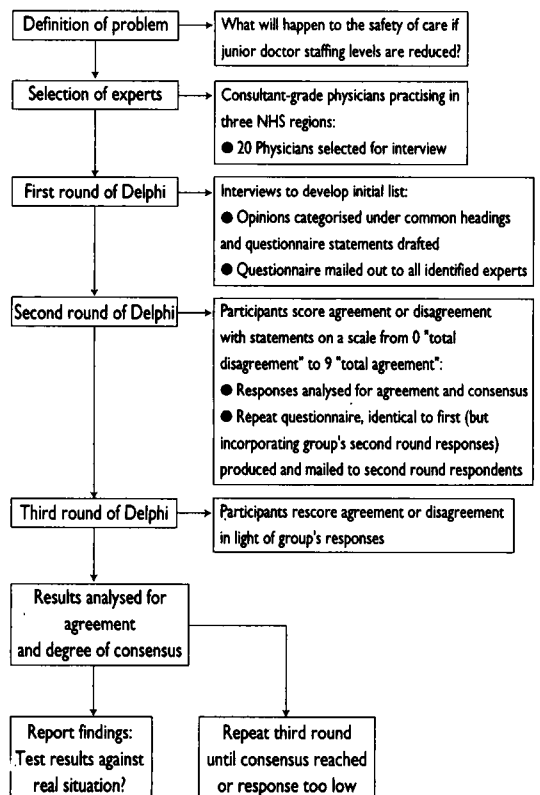


FIG 1—Example of Delphi process used in study by JJ

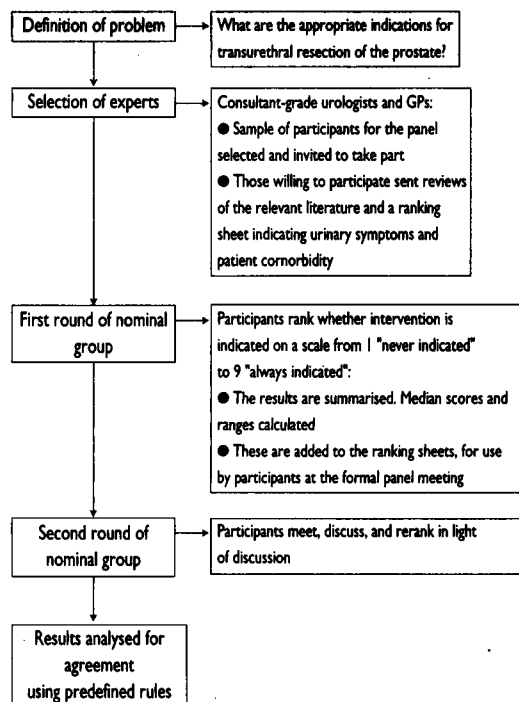


FIG 2—Example of modified nominal group undertaken by DH

- The final rankings are tabulated and the results fed back to participants.

Figure 2 shows an example of a modified nominal group undertaken by one of the authors (DH).

The method can be adapted and has been conducted as a single meeting or with the first stage conducted by post followed by a discussion and rerating at a face to face meeting. 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 non-participant observer collecting qualitative data on the nominal group. This approach has some features in common with focus groups (see article by Kitzinger³⁹). However, the nominal group technique focuses on a single goal (for example, the definition of criteria to assess the appropriateness of a surgical intervention) and is less concerned with eliciting a range of ideas or the qualitative analysis of the group process per se than is the case in focus groups.

Methodological issues

WHO TO INCLUDE AS PARTICIPANTS

There can be few hard and fast rules about who to include as participants, except that each must be justifiable as in some way "expert" on the matter under discussion. Clearly, for studies concerned with defining criteria for clinical intervention, the most appropriate experts will be clinicians practising in the field under consideration. However, the inclusion of other clinicians such as general practitioners may be appropriate to provide an alternative clinical view, particularly when the study is expected to have an impact beyond a particular specialist field. When the discussion concerns matters of general interest, such as health service priorities, participants should include non-clinical health professionals and the expression of lay opinions should also be allowed for.

There is clearly a potential for bias in the selection of participants. Although it has been shown that doctors who are willing to participate in expert panels are representative of their colleagues,⁴⁰ the exact composition of the panel can affect the results obtained.²⁴ The results will also be affected by any "random" variation in panel behaviour. These problems can be overcome

by using a different mixture of participants in further panels.

HOW TO MEASURE THE ACCURACY OF THE ANSWER OBTAINED

The existence of a consensus does not mean that the "correct" answer has been found—there is the danger of deriving collective ignorance rather than wisdom. The nominal group is not a replacement for rigorous scientific reviews of published reports or for original research, but rather a means of identifying current medical opinion and areas of disagreement. For Delphi surveys, Pill recommends that the results should, when possible, be matched to observable events.¹ Observers of the accuracy of opinion polls before the 1992 general election in Britain might well agree with this conclusion.

HOW TO FEED BACK THE RESULTS OF EACH ROUND

Agreement with statements is usually summarised by using the median and consensus assessed by using interquartile ranges for continuous numerical scales. These summary statistics may be fed back to participants at each round along with fuller indications of the distribution of responses to each statement in the form of tables of the proportions ranking at each point on the scale (see box 2), histograms, or other graphical representations of the range (see box 3). Feeding back the group's response enables participants to consider their initial ranking in relation to their colleagues' assessments. It should be made clear to each participant that they need not conform to the group view—though, in the nominal group technique, those with atypical opinions (compared with the rest of the group) may face critical questioning of their view from other panel members. In a Delphi exercise, the researcher undertaking the study may ask participants who they have defined as outliers (for example, those in the lower and upper quartiles) to provide written justification for their responses.

For nominal groups, rules have been developed to assess agreement when statements have been ranked on a 9 point scale (see box 3). In this example, the scale can be broken down so that scores 1-3 represent a region where participants feel intervention is not indicated; 4-6, a region where participants are equivocal; and 7-9, a region where participants feel intervention is indicated. The first rule is based on where the

Box 2—Example of feedback of second round results in a Delphi⁴⁰

The following are possible adverse effects of lowering the number of junior medical staff in general medicine and its associated specialties. The star indicates the number you selected to indicate the extent to which you agreed or disagreed with each statement in response to the previous questionnaire. Each of the numbers below the scale represents the percentage of those responding to the questionnaire who selected that particular value. We would be grateful if you would read through the questionnaire and consider whether, in the light of your colleagues' assessments, you would like to alter your response. Please indicate the extent to which you agree or disagree with each statement by circling the appropriate number (0 indicates total disagreement and 9 total agreement): if your choice remains unchanged please circle the same number you selected on the previous questionnaire.

i) Mortality rates in hospital will rise

	*										
disagree	0	1	2	3	4	5	6	7	8	9	agree
	5	3	9	4	3	18	18	12	16	12	

Box 3—Example of feedback of first round results in a nominal group²³

A—Chronic Retention: TURP is indicated for patients with chronic retention (but not acute) and there are:

IRRITATIVE SYMPTOMS

	Comorbidity (years of life expectancy)			
	High	Medium	Low	
None	○ 2 3 4 5 6 7 8 9	1 2 ○ 4 5 6 7 8 9	1 2 3 ○ 5 6 7 8 9	1—never indicated
Mild	1 ○ 2 3 4 5 6 7 8 9	1 2 ○ 4 5 6 7 8 9	1 2 3 4 ○ 6 7 8 9	9—always indicated
Moderate	1 2 3 ○ 5 6 7 8 9	1 2 3 4 5 ○ 7 8 9	1 2 3 4 5 6 ○ 7 8 9	○ Median
Severe	1 2 3 4 5 ○ 7 8 9	1 2 3 4 5 6 7 ○ 8 9	1 2 3 4 5 6 7 8 ○ 9	— Range

scores fall on the ranking scale (box 4): if all ratings fall within one of these predefined regions there is said to be strict agreement (in the example, all participants agreed that transurethral resection of the prostate was not indicated). An alternative relaxed definition for agreement is that all ratings fall within any 3 point region. This may be treated as agreement, in that all ratings are within an acceptable range, but the group opinion is ambiguous as to whether intervention is indicated or not.

The second rule tests whether extreme rankings are having an undue influence on the final results and consists of assessing the strict and relaxed definitions by including all ratings for each statement and then by excluding one extreme high and one extreme low rating for each statement. The ranges indicated in box 3 include all ratings, and it is noticeable that several of these ranges are from 1 to 9. It may be that these ranges exaggerate the dispersion of the group's response.

Validity and applicability

There has been an active debate on the validity of the Delphi method. For example, Harold Sackman argued that the Delphi method fails to meet the standards normally set for scientific methods.⁴¹ Many of his criticisms were aimed at past studies of poor quality rather than fundamental critiques of the method itself; he particularly criticised poor questionnaire design, inadequate testing of reliability and validity of methods, and the methods of defining and selecting experts. He also argued that the method forces consensus and is weakened by not allowing participants to discuss issues.

Reviews by Pill¹ and by Gerth and Smith (personal communication) showed no clear evidence in favour of meeting based methods over Delphi. Rowe *et al*, though, concluded that the Delphi technique is generally inferior to the nominal group technique, but state that the degree of inferiority is small, arising more from practical than from theoretical difficulties; they argue for further research aiming to improve the practice of Delphi studies—particularly a careful consideration of what constitutes expertise.²

Consensus methods, in particular Delphi, have been described as methods of "last resort,"⁴² with defenders warning against "overselling" the methods⁴³ and suggesting that they should be regarded more as methods for structuring group communication than as a means for providing answers. There is clearly a

danger that since these approaches have a prescribed method and are often used to generate quantitative estimates, they may lead the casual observer to place greater reliance on their results than might be warranted. As we stated earlier, unless the findings can be tested against observed data, we can never be sure that the methods have produced the "correct" answer. This should be made clear in reporting study results.

The structures of Delphi and nominal groups (shown in box 1) aim to maximise the benefits from having informed panels consider a problem (often termed "process gain") while minimising the disadvantages associated with collective decision making ("process loss"), particularly domination by individuals or professional interests. The extent to which these are realised depends on the ability of those running the studies to use the advantages of the methods. An important role of the facilitator in the nominal group is to ensure that all participants are able to express their views and to keep particular personal or professional views from dominating the discussion; participants in both Delphi and nominal group panels should be selected as to ensure that no particular interest or preconceived opinion is likely to dominate.

Uses

Consensus methods provide a useful way of identifying and measuring uncertainty in medical and health services research. Delphi and nominal group techniques have been used to clarify particular issues in health service organisation: to define professional roles, to aid design of educational programmes, to enable long term projections of need for care for particular client groups where there has been considerable uncertainty (for example, for cases of HIV and AIDS⁴), and to develop criteria for appropriateness of interventions as part of technology assessment. In addition to forming studies in their own right, these techniques have been widely used as component parts of larger projects.^{8,31} The two pieces of research from which materials have been presented in this paper each formed part of larger projects: the Delphi exercise⁴⁴ was concerned with defining possible adverse effects of reducing junior doctor staffing levels as part of a study of the adequacy of hospital medical staffing levels; the nominal group²³ was concerned with defining appropriate indications for surgical intervention as part of a population based assessment of need for prostate surgery within an NHS region.

Conclusions

The emphasis, when the findings of Delphi and nominal group studies are presented, should be on the justification in using such methods, the use of sound methodology (including selection of experts and the clear definition of target "acceptable" levels of consensus), appropriate presentation of findings (where proposed standards for presentation—as for clinical practice guidelines⁴⁵—should be considered),

Box 4—Examples of strict and relaxed rules for agreement in a nominal group

○ 2 3 4 5 6 7 8 9	Strict:	all ratings fall within a predefined 3 point region (agreement that intervention is "not indicated")
1 ○ 2 3 4 5 6 7 8 9	Relaxed:	all ratings fall within a 3 point region, but not one of the predefined regions (crossing "not indicated" (1-3) and "equivocal" (4-6) regions)

and on the relevance and systematic use of the results. The output from consensus approaches (including consensus development conferences) is rarely an end in itself. Dissemination and implementation of such findings is the ultimate aim of consensus activities—for example, the publication of consensus statements intended to guide health policy, clinical practice, and research, such as the consensus statement on cancer of the colon and rectum.⁴⁶

- 1 Pill J. The Delphi method: substance, context, a critique and an annotated bibliography. *Socio-Economic Planning Science* 1971;5:57-71.
- 2 Rowe G, Wright G, Bolger F. Delphi: a re-evaluation of research and theory. *Technological Forecasting and Social Change* 1991;39:235-51.
- 3 McGlynn EA, Koseoff J, Brook RH. Format and conduct of consensus development conferences. *International Journal of Technology Assessment in Health Care* 1990;6:450-69.
- 4 Perry S. The NIH consensus development program: a decade later. *N Engl J Med* 1987;317:485-8.
- 5 Perry S, Kalberer JT. The NIH consensus development program and the assessment of health care technologies: the first two years. *N Engl J Med* 1980;303:169-72.
- 6 Stocking B, Jenett B, Spiby J. Criteria for change. In: *The history and impact of consensus development conferences in the UK*. London: King's Fund Centre, 1991.
- 7 Beers MH, Ouslander JG, Rollinger I, Reuben DB, Brooks J, Beck JC. Explicit criteria for determining inappropriate use in nursing home residents. *Arch Intern Med* 1991;151:1825-32.
- 8 Bellamy N, Anastassiades TP, Buchanan WW, Davis L, Lee P, McCain GA, et al. Rheumatoid arthritis anti-rheumatic trials. III. Setting the delta for clinical trials of anti-rheumatic drugs—results of a consensus development (Delphi) exercise. *J Rheumatol* 1991;18:1908-15.
- 9 Chin J, Sato PA, Mann JM. Projections of HIV infections and AIDS cases to the year 2000. *Bull World Health Organ* 1990;68:1-11.
- 10 Khan JP, Bernstein SJ, Leape LL, Hilborne LH, Park RE, Parker L, et al. Measuring the necessity of medical procedures. *Med Care* 1994;32:357-65.
- 11 Crotty M. The emerging role of the British nurse teacher in Project 2000 programmes: a Delphi survey. *J Adv Nurs* 1993;18:150-7.
- 12 Fraser CE, Smith QW, Luchi RJ. Geriatric fellows' perceptions of the quality of their research training. *Acad Med* 1992;67:696-8.
- 13 Elder OC Jr, Andrew ME. Important curriculum content for baccalaureate allied health programs: a survey of deans. *J Allied Health* 1992;21:105-15.
- 14 Rizzolo MA. Factors influencing the development and use of interactive video in nursing education. A Delphi study. *Comput Nurs* 1990;8:151-9.
- 15 Attala JM, Gresley RS, McSweeney N, Jobe MA. Health needs of school-age children in two midwestern counties. *Issues in Comprehensive Paediatric Nursing* 1993;16:51-60.
- 16 Moscovice I, Armstrong P, Shortell S. Health services research for decision-makers: the use of the Delphi technique to determine health priorities. *J Health Polit Policy Law* 1988;2:388-410.
- 17 Oranga HM, Nordberg E. The Delphi panel method for generating health information. *Health Policy and Planning*. 1993;8:405-12.
- 18 Rainhorn JD, Brudon-Jakobowicz P, Reich MR. Priorities for pharmaceutical policies in developing countries: results of a Delphi survey. *Bull World Health Organ* 1994;72:257-64.
- 19 Mobily PR, Herr KA, Kelley LS. Cognitive-behavioural techniques to reduce pain: a validation study. *Int J Nurs Stud* 1993;30:537-48.
- 20 Passannante MR, Restifo RA, Reichman LB. Preventive therapy for the patient with both universal indication and contraindication for isoniazid. *Chest* 1993;103:825-31.
- 21 Smith DM, Murphy WM. Histological changes in prostate carcinomas treated with leuprolide (luteinizing hormone-releasing hormone effect). Distinction from tumor differentiation. *Cancer* 1994;73:1472-7.
- 22 Fink A, Koseoff J, Chassin M, Brook RH. Consensus methods: characteristics and guidelines for use. *Am J Public Health* 1984;74:979-83.
- 23 Hunter DJW, McKee CM, Sanderson CFB, Black NA. Appropriateness of prostatectomy: a consensus panel approach. *J Epidemiol Community Health* 1994;48:58-64.
- 24 Scott EA, Black NA. Appropriateness of cholecystectomy—a consensus panel approach. *Gut* 1991;9:1066-70.
- 25 Technology Subcommittee of the Working Group on Critical Care, Ontario Ministry of Health. Hemodynamic monitoring: a technology assessment. *Can Med Assoc J* 1991;145:114-21.
- 26 Technology Subcommittee of the Working Group on Critical Care, Ontario Ministry of Health. Guidelines for medical technology in critical care. *Can Med Assoc J* 1991;144:1617-22.
- 27 Ziemba SE, Hubbell FA, Fine MJ, Burns MJ. Resting electrocardiograms as baseline tests: impact on the management of elderly patients. *Am J Med* 1991;91:576-83.
- 28 Bartles JB, Kirk LM, Dowell DL, Frnka S. The health sciences communicator as faculty developer. *J Biocommun* 1989;16:2-8.
- 29 Buchan MS, Hegge MJ, Stenvig TE. A tiger by the tail: tackling barriers to differentiated practice. *J Contin Educ Nurse* 1991;22:109-12.
- 30 McKee M, Priest P, Ginzler M, Black N. Which tasks performed by pre-registration house officers out of hours are appropriate? *Med Educ* 1992;26:51-7.
- 31 Justice J, Jang R. Tapping employee insights with the Nominal Group Technique. *Am Pharm* 1990;NS30:43-5.
- 32 Manning D, Balson PM, Barenberg N, Moore TM. Susceptibility to AIDS: what college students do and don't believe. *J Am Coll Health* 1989;38:67-73.
- 33 Neidlinger SH, Ljubic PM, Moed AB, Chin PP, Kowalczyk ME, Anton RS. Defining the major nursing administration issue: doing more with less. *J Nurs Admin* 1993;23:9-10.
- 34 Felson DT. Choosing a core set of disease activity measures for rheumatoid arthritis clinical trials. *J Rheumatol* 1993;20:531-4.
- 35 Gallagher M, Hares T, Spencer J, Bradshaw C, Webb I. The nominal group technique: a research tool for general practice? *Fam Pract* 1993;10:76-81.
- 36 Liang MH, Katz JN, Phillips C, Sledge C, Cats-Baril W. The total hip arthroplasty outcome evaluation form of the American Academy of Orthopaedic Surgeons. Results of a nominal group process. The American Academy of Orthopaedic Surgeons Task Force on Outcome Studies. *J Bone Joint Surg [Am]* 1991;73:639-46.
- 37 Delbecq A, Van de Ven A. A group process model for problem identification and program planning. *Journal of Applied Behavioural Science* 1971;7:467-92.
- 38 Glaser EM. Using behavioural science strategies for defining the state-of-the-art. *Journal of Applied and Behavioural Sciences* 1980;16:79-92.
- 39 Kitzinger J. Introducing focus groups. *BMJ* 1995;311:299-302.
- 40 McKee M, Priest P, Ginzler M, Black N. How representative are members of expert panels? *Quality Assurance in Health Care* 1991;3:89-94.
- 41 Sackman H. *Delphi critique*. Lexington, MA: Lexington Books, 1975.
- 42 Coates JF. In defense of Delphi: a review of *Delphi assessment, expert opinion, forecasting and group process* by H Sackman. *Technological Forecasting and Social Change* 1975;7:193-4.
- 43 Linstone HA. The Delphi technique. In: Fowles RB, ed. *Handbook of futures research*. Westport, CT: Greenwood, 1978.
- 44 Jones JMG, Sanderson CFB, Black NA. What will happen to the quality of care with fewer junior doctors? A Delphi study of consultant physicians' views. *J R Coll Physicians Lond* 1992;26:36-40.
- 45 Hayward RSA, Wilson MC, Tunis SR, Bass EB, Rubin HR, Haynes RB. More informative abstracts of articles describing clinical practice guidelines. *Ann Intern Med* 1993;118:731-7.
- 46 King's Fund Forum. Consensus statement on cancer of the colon and rectum. *Int J Technol Assess Health Care* 1991;7:236-9.

Research meets clinical practice

I am a psychiatrist in a service for adolescents with serious psychiatric illness. I was recently invited to take part in a double blind, multicentre trial looking at the use of a relatively new antidepressant. My involvement would be to identify patients with major depressive illness who would then be screened for their suitability for the study. The study seemed well designed and would of course be presented to the relevant ethical committees.

It remains to be shown that antidepressant drug treatment has a therapeutic effect greater than that of placebo in adolescents. Some doctors continue to prescribe in the belief that the problem is one of study design and that in the meantime patients should not be deprived of what might be an effective treatment. Obviously this is an area of major importance and I agreed to be involved in the study.

But I subsequently became aware of several difficulties:

(1) The drug in question is not my usual drug of choice. In order to use it most effectively I would need to familiarise myself rapidly with it and with its side effects.

(2) Patients with serious suicidal ideation are excluded from the study. Although this is a relief from a clinical viewpoint, I have concerns about the usefulness of a study into major depression which excludes patients with

serious suicidal ideation. A further difficulty is that determining whether suicidal ideation is serious is a subjective judgment which is likely to vary with different centres and clinicians.

(3) Patients with serious depressive illness will receive placebo. A research perspective would suggest that as the medication has not been shown to be more effective than placebo this is acceptable. On the other hand, given that the medication may be effective, as a clinician I find its being withheld a source of concern.

(4) The research makes demands on patients who are already having to cope with depressive illness. They have to cope not only with discussion of their condition and its treatments but also of the research project. If they refuse consent or give it reluctantly this may put a strain on our relationship. I also wonder about the effect of telling patients that they may not be receiving an active treatment.

I am reminded that research and clinical perspectives are different and that accommodating both is not always easy. Clinicians are best involved in research at an early stage and not just as providers of patients. I still intend to cooperate with the study.—TONY JAFFA is a consultant child and adolescent psychiatrist in Cambridge