

LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE



Raine, R; Sanderson, C; Black, N (2005) Developing clinical guidelines: a challenge to current methods. *BMJ*, 331 (7517). pp. 631-3. ISSN 1468-5833 DOI: <https://doi.org/10.1136/bmj.331.7517.631>

Downloaded from: <http://researchonline.lshtm.ac.uk/13061/>

DOI: [10.1136/bmj.331.7517.631](https://doi.org/10.1136/bmj.331.7517.631)

Usage Guidelines

Please refer to usage guidelines at <http://researchonline.lshtm.ac.uk/policies.html> or alternatively contact researchonline@lshtm.ac.uk.

Available under license: Creative Commons Attribution Non-commercial
<http://creativecommons.org/licenses/by-nc/3.0/>

Developing clinical guidelines: a challenge to current methods

Rosalind Raine, Colin Sanderson, Nick Black

Current methods for producing clinical guidelines are cumbersome and not always reliable. Could a more streamlined approach improve coverage and make decisions more transparent?

Clinical guidelines are rarely based solely on research evidence. In most cases they also incorporate the consensus views of experts. Despite recognition of the need for rigour in developing a consensus, current approaches often lack sufficient transparency, fail to make clear what influence the level of resources in the health system has, lack sufficient reliability, and will never achieve comprehensive and timely coverage of the whole range of health care. We propose a new approach that we believe will be more cost effective and that could meet these challenges.

Need for consensus

Most professional societies and national agencies in North America, Australia, and Europe recognise that guidelines cannot be based on research evidence alone. To paraphrase the philosopher David Hume: "ought" statements such as guidelines cannot be constructed from "is" statements such as research evidence.¹ The conversion from is to ought inevitably introduces value judgments about underlying goals and, in this context, somewhat subjective assessments of the quality of the research evidence and its relevance for particular patients and settings.

The recognised need to go beyond the often limited research evidence has led to the use of expert consensus in developing guidelines. The assumption is that the views of a group have greater validity and reliability than the judgment of an individual. In addition, formal or structured methods for developing a consensus have advantages over informal committees: they should offer more transparent ways of synthesising individual judgments, can reduce the influence of dominating personalities and "group think," and can provide valuable information on the extent and reasons for differences of opinion.²

Since there can be no standard for checking the validity of a consensus based guideline, we must ensure that the way consensus is developed, in particular who to involve and how to structure the process, is as rigorous as possible. Despite rapid growth in the use of formal methods for developing consensus throughout the world,³ several challenges remain.^{4 5}

Current approaches to developing guidelines

The three commonly used methods for developing guidelines are the nominal group technique, the Delphi survey, and a hybrid of the two. In the nominal group technique about 10 people are usually selected to identify the questions to be covered in the guideline, express their views in private, and then to discuss areas of disagreement. After the discussion the participants again provide their views in private, which the organis-

ers then analyse to derive the group view. These meetings reduce the risk of misunderstandings and expose the reasons for differences of opinion.⁶ In contrast, the Delphi survey involves two or more rounds of postal questionnaires. This allows larger and more geographically dispersed groups of participants and avoids the risk of some individuals exercising undue influence.⁷ However, the opportunities for clarification and resolution of differences of opinion are more limited. A hybrid approach, pioneered by RAND, uses a postal questionnaire for the first round of ratings and a meeting for the second round.⁸

Formal, structured methods are used in England by the National Institute for Health and Clinical Excellence, which has a statutory role in producing advice on best practice.⁹ These methods are also widely used in other countries, with many of the guidelines appearing on the website of the National Guideline Clearinghouse (www.guideline.gov), an initiative of the US Agency for Healthcare Research and Quality.

Methodological and practical concerns

These methods give rise to three main methodological concerns and one practical concern. Firstly, the format, although structured, is often not sufficient to allow the reasons for judgments to be fully transparent. And only some aspects of the process that influence the outcome are made explicit (and even then not always). These influences include the composition of the group (for example, specialists tend to overstate the appropriateness of the interventions they perform compared with generalists¹⁰⁻¹³), whether a literature review was used, and the procedure for aggregating judgments.

Our second concern is how the resource constraints faced in group members' clinical work affect their views. International differences in guidelines sug-

Health Services Research Unit, London School of Hygiene and Tropical Medicine, London WC1E 7HT

Rosalind Raine
MRC clinician scientist

Colin Sanderson
reader in health services research

Nick Black
professor in health services research

Correspondence to:
R Raine
rosalind.raine@lshtm.ac.uk

BMJ 2005;331:631-3



STONEGEM

gest that the level of resources available in a healthcare system has an influence, along with cultural and organisational factors.¹⁴ The few attempts that have been made to examine such influences directly have produced conflicting results.^{4 15} Until this issue has been clarified, it is unwise to assume that contextual issues, such as availability of resources, are ignored by groups when making decisions or that a group's view about value and affordability of treatments coincides with that of the wider community.^{16 17}

The third concern is reliability, particularly for nominal groups. Their strength lies in providing a forum for detailed discussion, but it can also be their weakness because it can lead to unrepresentative, and therefore unreliable, judgments. Delphi surveys using larger groups show greater within and between group reliability (A Hutchings et al, unpublished data). A trade-off is therefore needed between keeping a group small enough to allow face to face discussions that reveal people's reasoning and large enough to ensure reliable guidelines are produced.

Finally, we have a practical concern about the sustainability of existing programmes to develop guidelines given the time and cost involved. For example, in England the National Institute for Health and Clinical Excellence takes at least 18 months and convenes as many as 15 meetings of the guideline development groups to produce guidance that may need to be reviewed every couple of years. The producers of one guideline in Spain had to draft 13 versions of a data collection form and establish 10 working groups (www.guideline.gov). These processes mean that only a small proportion of health care will ever be covered and it may not be feasible to update guidelines frequently enough.

Suggestions for a way forward

With these concerns in mind, we suggest some improvements in the process for developing guidelines. Transparency could be enhanced by:

- Making the goals of each guideline explicit—for example, effectiveness, cost effectiveness, equity
- Providing information on the reasons for disagreements within a guideline development group—such as, differences in interpretation of the research literature, differences in personal experience, different perceptions of or responses to costs
- Publishing information on the closeness of agreement about a recommendation as well as the strength of support for a recommendation.

The influence of the resource constraints in the local health system should be considered at an early stage. Two useful steps would be to provide the members of the group with authoritative information about the costs of the treatments being considered and to carry out an outcome valuation exercise to ensure that the costs of any recommended treatments are not greatly out of line with those in other guidelines and broader societal views.

Reliability should be ensured by checking the views of nominal groups against those of the wider community. This could also help improve ownership of the guidelines. Including clinicians in typical practice, research methodologists, and patients in the process might also improve reliability.

The practicality of creating and maintaining guidelines across the whole range of health care requires the use of a more efficient means of guideline development, with fewer group meetings and a shorter time frame.

New approach

We propose three meetings of a guideline development group comprising relevant practitioners and other stakeholders. At the first meeting the group would identify the specific issues to be examined. Methodologists would then review and synthesise the research evidence and other relevant material on effectiveness, cost, and cost effectiveness, taking care to document any judgments about conflicting evidence and methodological limitations.¹⁸ The evidence and stakeholders views would be used to develop and pilot a questionnaire consisting of the relevant clinical scenarios to be considered. Each scenario would be accompanied by a Likert scale for participants to provide a rating from 1 (strongly disagree) to 9 (strongly agree), with a separate box for don't know.

Members of the guideline development group would complete these questionnaires privately, and the groups' view would be presented at the second meeting. The group would explore the extent of, and reasons for, disagreements and clarify ambiguities. Participants would then have an opportunity to revise their ratings privately. The meeting would be observed and audiotaped to enable a thematic analysis of the extent to which issues such as cost, effectiveness, priority, feasibility, and acceptability influence ratings. For example, it may be important to understand whether the development group judged an intervention to be appropriate in a particular instance because it accepted a limited amount of evidence of a large benefit or a wealth of evidence for a small benefit.

The second meeting should also consider the value of the expected health outcomes. It is not yet clear what approach would be feasible for routine use, but a first step would be to report group members' valuations and how these affect their judgments. The representativeness of the groups' ratings should be checked by posting a random sample of the results to a large, similarly composed group, who would be invited to rate the results and to comment on the overall guidelines. The development group would then meet for a third time to turn their appropriateness ratings into recommendations, having considered the results of the larger survey.

The published guidelines would include an indication of the underlying assumptions of the group and the strength of support for each recommendation together with the extent of agreement about it. Surprising or controversial recommendations would be explained.

The research evidence informing the guidelines should be reviewed using the previously agreed search strategy every two years. The guideline development group should then meet to decide whether the guidelines should be reconsidered in the light of new research evidence. In addition, the guideline development group would meet if major new research evidence is published in the interim.

Summary points

The translation of research evidence into guidelines has barely been considered

Formal consensus methods used to develop guidelines lack sufficient transparency and reliability, and the process is too cumbersome to be sustainable

A new approach is suggested which makes the goals, reasons for disagreement, and degree of consensus explicit

Inclusion of a survey stage enhances reliability

Meetings of the development group are limited to three to ensure sustainability

We believe such an approach would meet the practical challenges outlined above. We also contend that the resulting guidelines would prove to be as valid and reliable as those emanating from the existing production methods and would be more cost effective. This contention should be tested through a rigorous comparison with currently used methods.

We thank John Cairns and Andrew Hutchings for advice and comments on the drafts of this article.

Funding: RR was funded by a Medical Research Council clinician scientist fellowship.

Contributors and sources: RR led a four year research programme into the methodological basis of developing guidelines. NB and CS are members of the steering group for this research programme. The proposals outlined are the outcome of lengthy discussion between the authors about the implications of the findings of the research RR wrote the first

draft and all authors discussed ideas and collaborated on subsequent drafts. RR is the guarantor.

Competing interests: None declared.

- Hume D. *A treatise of human nature*. Oxford: Clarendon Press, 1978.
- Janis I. *Groupthink*. 2nd ed. Boston: Houghton-Mifflin, 1982.
- Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. *BMJ* 1999;318:527-30.
- Murphy MK, Black NA, Lamping DL, McKee CM, Sanderson CFB, Ashkam J, et al. Consensus development methods, and their use in clinical guideline development. *Health Technol Assess* 1998;2:1-88.
- Raine R, Sanderson C, Hutchings A, Carter S, Larkin K, Black N. An experimental study of determinants of group judgments in clinical guideline development. *Lancet* 2004;354:429-37.
- Delbecq, Van de Ven A. A group process model for problem identification and program planning. *J Appl Behav Sci* 1971;7:467-92.
- Bayley EW, Richmond T, Noroian EL, Allen LR. A Delphi study on research priorities for trauma nursing. *Am J Crit Care* 1994;3:208-16.
- Brook RH. The RAND/UCLA appropriateness method. In: McCormick KA, Moore SR, Siegel RA, eds. *Clinical practice guidelines development: methodology perspectives*. Rockville, MD: Public Health Service, US Department of Health and Human Services, 1994:59-70. (AHCPR publication No 95-0009).
- Rawlins M. NICE work: providing guidance to the British National Health Service. *N Engl J Med* 2004;351:1383-5.
- Scott EA, Black N. When does consensus exist in expert panels? *J Pub Health Med* 1991;13:35-9.
- Herrin J, Etchason JA, Kahan JP, Brook RH, Ballard DJ. Effect of panel composition on physician ratings of appropriateness of abdominal aortic aneurysm surgery: elucidating differences between multispecialty panel results and specialty society recommendations. *Health Policy* 1997;42:67-81.
- Fitch K, Lázaro P, Aguilar MD, Martín Y, Bernstein SJ. Physician recommendations for coronary revascularization: variations by clinical specialty. *Eur J Pub Health* 1999;9:181-7.
- Bernstein SJ, Lazaro P, Fitch K, Aguilar MD, Kahan JP. Effect of specialty and nationality on panel judgments of the appropriateness of coronary revascularization: a pilot study. *Med Care* 2001;39:513-20.
- Savoie I, Kazanjian A, Bassett K. Do clinical practice guidelines reflect research evidence? *J Health Serv Res Policy* 2000;5:76-82.
- Taffé P, Burmand B, Wietlisbach V, Vader JP. Influence of clinical and economical factors on the expert rating of appropriateness of preoperative use of recombinant erythropoietin in elective orthopedic surgery patients. *Med Decis Making* 2004;24:122-30.
- Rawlins MD, Culyer AJ. National Institute for Clinical Excellence and its value judgments. *BMJ* 2004;329:224-7.
- Shekelle P. The appropriateness method. *Med Decis Making* 2004;24:228-31.
- Burgers JS, van Everdingen JJ. Beyond the evidence in clinical guidelines. *Lancet* 2004;364:392-3.

(Accepted 27 July 2005)

Chocolate poisoning

Readers may be intrigued to know that “man’s best friend” may be at risk from one of our favourite treats—chocolate. The potential hazards to humans of eating too much chocolate are well known (obesity and dental caries to name but two), but you may be unaware that chocolate is potentially lethal to dogs.

Chocolate is derived from the roasted seeds of the plant *Theobroma cacao*, and the main toxic components are the methylxanthine alkaloids theobromine and caffeine. Humans can easily digest and excrete methylxanthines, the half life of theobromine being 2-3 hours. However absorption in dogs is slow, with metabolism in the liver and extrahepatic recirculation before excretion in the urine. The half life of theobromine in dogs is about 18 hours.

Theobromine primarily affects the central nervous system, cardiovascular system, and respiratory system, as well as having a diuretic effect. The first signs of poisoning in dogs include vomiting, haematemesis, and polydipsia. Other signs may include hyperexcitability, hyperirritability, tachycardia, excessive panting, ataxia, and muscle twitching. Effects may progress to cardiac arrhythmias, seizures, and death. Most symptoms will begin to appear within two hours of ingestion, but, as theobromine is metabolised slowly, it can take as long as 24 hours for them to appear and up to three days for recovery. Although there is no specific antidote, supportive management includes induction of

vomiting and administration of activated charcoal, oxygen, and intravenous fluids.

The lethal dose of theobromine is reported to be 100-500 mg/kg of body weight in dogs. However, not all types of chocolate contain the same amount of theobromine: cocoa powder and plain chocolate contain the highest concentrations (20 mg/g and 15 mg/g), milk chocolate has much less (2 mg/g), and white chocolate has the lowest concentration (0.1 mg/g). Thus, less than 100 g of plain chocolate may be fatal for a 10 kg dog.

Although it is relatively safe to give your pet a small chocolate treat occasionally, all dogs are potentially at risk from chocolate, and the safer alternative is to give special “pet chocolate” that does not contain theobromine (but can still cause obesity).

As a final thought, keen gardeners who like to spread mulch on their garden in the spring and summer months should be aware that cocoa shell mulch also contains very high levels of theobromine (25 mg/g). As it has the characteristic chocolate smell it may be attractive to dogs but is potentially lethal.

Remember, dogs love chocolate as much as we do—keep boxes of chocolates, Easter eggs, and tins of cocoa out of paws’ reach.

Fiona Finlay *consultant community paediatrician, Child Health Department, Bath NHS House, Bath (fiona.finlay@barnes-pct.nhs.uk)*,

Simon Guiton *veterinary surgeon, 12 Raby Place, Bath*