



1353-4505(94)E00025-5

## Should we Assess Appropriateness of Care in Europe?

FRED PACCAUD\* and HERVE GUILLAIN‡

\*Department of Social and Preventive Medicine, School of Medicine, University of Lausanne, 17, rue du Bugnon, 1005 Lausanne, Switzerland

‡Swiss Institute for Public Health, 21, rue du Bugnon, 1005 Lausanne, Switzerland

The appropriateness of health care interventions is becoming a central issue for health care providers and policy makers alike. Research on geographic variations in practice patterns has shown large differences in hospital admission rates, average length of stay, surgical procedure rates, and use of other health care services [1]. Since these variations cannot be fully explained by differences in the health status of the populations [2] or the supply of providers [3], they lead us to ask whether patients in some areas receive more care and in other areas less care than should be provided. In fact, the issue is not only to consider whether the volume of care is too high or too low, but, more importantly, to determine the extent to which patients receive care that is “right” for the circumstances.

A procedure can be defined as appropriate for a given indication “when the benefits of performing the procedure outweigh the risks by a sufficient margin that it is worth doing and the procedure is preferable to the alternatives” [4]. The benefits can be an increased life expectancy, relief of pain, reduction in anxiety, improved functional capacity, etc. The risks include mortality, morbidity, anxiety of anticipating the procedure, time lost from work, etc. In order to assess the appropriateness of health care interventions, Brook and colleagues developed a method to establish guidelines for the clinical practice [5]. Although the method was

initially designed for clinicians, it was payers and health services researchers who became more rapidly involved in applying this method. As a result, practice guidelines are now widely used in the US for utilization review and pre-certification purposes.

### A CONFERENCE IN LAUSANNE, SWITZERLAND

In Europe, there is a growing interest in guidelines and medical appropriateness. Research in this field is being conducted in several countries and new groups are joining those who are already familiar with the appropriateness method. In order to provide a forum for approaches and possible collaborative projects, a conference took place on 1–2 November 1993, in Lausanne, Switzerland, organized by the Department of Social and Preventive Medicine, School of Medicine, University of Lausanne. Speakers and participants came from France, Germany, Hungary, Italy, The Netherlands, Poland, Spain, Sweden, Switzerland, the UK as well as Australia, Israel and the US. Presentations focused on the use of the appropriateness method in the US, the lessons learned from multi-institutional studies, the current situation regarding the assessment of appropriateness of care in European countries and Israel, the prospects of implementing review systems similar to those currently used in the US, and the conditions that need to be met for such implementation to be successful. This article reflects several issues raised during the November meeting in Lausanne.

### TWO CRUCIAL CONDITIONS TO BE SATISFIED

Brook opened the conference by reminding the audience that any attempt to assess the quality of care may quickly provoke strong

negative reactions from the providers whose services are critically appraised. Therefore, before starting a program aimed at analysis and improvement of the appropriateness of medical care, it is important to gain the support of those health care providers concerned with the procedure for which criteria are being developed. It does not suffice to inform clinicians about the study in which they are kindly asked to participate. They must have a say in the project and perceive the appropriateness method as a tool that will help them to examine their own practice and to improve it, and even to protect themselves by offering documented evidence of appropriateness. From the experience shared by those who attended the meeting, clinicians' involvement is clearly a major factor predicting how successful an appropriateness program will be. No matter how enthusiastic administrators, managers and health service researchers are, the appropriateness of health care interventions cannot be addressed without the full participation of the clinicians who actually perform the procedures.

The introduction of appropriateness criteria in Europe also requires consideration of the differences in medical practice between the US and European countries. The scientific foundation of clinical practice is rather weak [6] and patients suffering from the same disease may be treated quite differently from country to country [7]. Despite the abundance of international conferences and a considerable exchange of information across the world, medicine remains a local business. As a result, appropriateness criteria developed in one place may not be relevant, applicable or acceptable in another, and the conduct of appropriateness studies ought to include the formulation of national or regional panels to critically appraise the guidelines established elsewhere and, if needed, to reformulate indications and rate their appropriateness anew. The members of this panel should be chosen among insightful clinicians and researchers identified as opinion leaders.

#### **POTENTIAL USES OF APPROPRIATENESS CRITERIA**

The Lausanne meeting considered the various potential uses of appropriateness criteria in

European countries: as regulatory, research, educational or quality improvement tools.

One use of appropriateness criteria is for clinical practice regulation. This is now common in the US where many payers rely on guidelines and appropriateness criteria to pre-certify reimbursement of a procedure to be done on a given patient with specific clinical characteristics. The use of guidelines and appropriateness criteria has been facilitated by the development of computer-based utilization review systems. The Medical Review System (MRS), for example, has been designed to identify those procedures that appear to be medically unwarranted [8]. At the present time, this system is used in more than 77 sites in the US and covers more than 11.5 million American lives. Another product, *QualityFIRST* software, serves as a decision-making aid in prospectively evaluating treatments and procedures prescribed by over 100,000 physicians annually in over 6000 hospital and out-patient settings in the US. These are only two examples of the many products now available to assist in identifying the most appropriate care plan in each individual case.

In Europe, to our knowledge, no health insurance company, sickness fund or regulatory body has so far implemented pre-certification or utilization review systems based on appropriateness criteria. But the pressure to contain health expenditures is growing, even if the percentage of the gross domestic product spent on health care in European countries is still lower than in the US. Once national or regional expert panels have established lists of appropriate and inappropriate indications for a variety of health care interventions, state and private insurance companies may decide to base their reimbursement on these lists. Since health plans generally do not cover drugs that are considered ineffective, the same logic could be applied to any diagnostic, therapeutic or preventive procedures. In the same line, appropriateness criteria could play a role in the planning of health resources: by considering the overuse or underuse of certain procedures, the planning process could be refined, more realistic projections obtained, and planning interventions better targeted.

However, the introduction of appropriateness criteria for regulatory purposes in Euro-

pean countries may not be propitious at this stage, and it could do a great disservice to the proponents of the appropriateness method. If clinicians' first contact with criteria is through denied payments, they are likely to perceive the appropriateness method as just one more mechanism to reduce costs and limit clinical autonomy. They could reject both the application and the underlying method, without seeing how these criteria can help them to improve their practice.

Appropriateness criteria could also be used as a research tool. For example, variations in practice pattern are an area for investigation: although one study carried out in the US showed no significant correlation between the rate of procedures and their appropriateness [9], observations involving different countries are still rare and could clarify some points [10]. Appropriateness criteria may also help to identify questions regarding clinical practice that need to be answered: a comparison of appropriateness criteria across countries may bring to light which clinical studies should be formulated and designed, allowing us to concentrate on what we most need to evaluate [11]. In this respect, the appropriateness method might be a valuable tool to establish a well-founded research agenda, in particular to prioritize the design and implementation of randomized clinical trials (RCTs). Some seem to believe that appropriateness criteria are substitutes for RCTs. This is obviously not the case; in fact, appreciations made by experts could be validated by comparing them to the results of the RCTs. The use of appropriateness criteria as a part of the research strategy in clinical practice is promising. The appropriateness method requires that health professionals make explicit judgment about what should be done or not done under specific clinical circumstances. In itself, this is already a tremendous contribution. The "black box" of medical practice needs to be opened and its content become more transparent to clinicians themselves, payers, managers, decision makers, and consumers. Only on this condition will it be possible to allocate resources without arbitrarily impeding access to care or diminishing the quality of care.

A third possible use of appropriateness criteria is clinical education: they could be integrated into the curriculum of medical students,

interns, and residents. Since the criteria developed by Brook and colleagues are based on very detailed clinical situations, they can be used in training programs more easily than other very general guidelines. Like any guidelines, appropriateness criteria will have to be revised whenever required by new knowledge produced by clinical research or feedback from those who apply them [12]. Therefore, continuing medical education programs may also include sessions to keep physicians abreast of updated criteria.

In the current and future practice of medicine, even the best expert in his own field will never be able to recall which indications are appropriate, equivocal or inappropriate. The development of computer software packages should greatly facilitate the learning, use and updating of appropriateness criteria. For example, an adaptation of the MRS or a similar product could be developed for training purposes rather than for utilization review. If early in their training period students are exposed to such products, they may well continue using them throughout their career.

Although guidelines are sometimes viewed as "cookbook medicine", the uncertainty of clinical decisions is responsible for many stressful situations in clinical practice. This problem should not be underestimated, and appropriateness criteria are tools that could reduce this uncertainty [13].

Finally, appropriateness criteria could also be used in quality assurance (QA) and quality improvement (QI) programs, where they could be seen as standards to be met. They could be used as direct measurement tools to assess the quality of the process of care actually performed in the health system: that is, they provide a measurement tool allowing the comparison of quality of care across institutions. For example, data on the rate of appropriate and inappropriate interventions could be required from hospitals as part of an accreditation process. Within institutions, in addition to volume and cost, appropriateness of procedures ordered or performed could also be taken into account when establishing physician practice profiles.

## CRITICAL APPRAISAL

Admittedly, the appropriateness method is neither a panacea nor a perfect instrument [14].

Several important dimensions of the quality of care do not enter into the method developed by Brook and colleagues [15]. Cost-effectiveness, patient satisfaction, proficiency of the performing physician or coordination of care are not considered, although they may be as important as appropriateness. For example, the indication for a surgical intervention may be entirely appropriate, but the surgeon may lack technical skills or a wrong dose of prophylactic antibiotics may be administered before the surgical procedure is initiated.

The central issue is the validity of appropriateness criteria. According to the Committee to Advise the Public Health Service on Clinical Practice Guidelines [12], "a valid practice guideline is one that, if followed, will lead to the health and cost outcomes projected for it, other things being equal". The use of appropriateness criteria will be wholly justified only when it can be reasonably argued that patients undergoing appropriate procedures do have better health outcomes than patients undergoing inappropriate procedures or patients undergoing no procedure at all when the procedure would have been appropriate for them. Since such evidence is still lacking, the effect of appropriateness criteria on the outcome of care needs to be studied, in spite of the many difficulties that will be encountered.

The reliability of the criteria established by expert panels using the method developed by Brook and colleagues needs to be further examined. Since it has been shown that the composition of the panel has an effect on appropriate ratings [16,17], it is critical to identify the best way to select panel members so that indications and criteria will be developed in a reproducible manner, independent from the groups of experts. Research conducted so far suggests that panels should include both performers and non-performers of the procedure to avoid ratings that favour either overly conservative or overly aggressive approaches. Another practical question needing a more precise answer is the extent of the expensive literature review we need to perform before panel members start rating indications according to their appropriateness.

The time lag between the availability of a set of criteria and their use in the real world is an important, though perhaps overestimated, problem which should also be addressed.

Another question is to know how "local" criteria can or should be: although scientific knowledge is universal, there is considerable agreement that medical care must be adapted to local constraints or local beliefs. The optimal limit of this "localization" process remains to be determined.

An additional question is the selection the procedures for which appropriateness criteria should be developed [18]: according to need perceived by clinicians, costs, potential health gains, or need for clearer indications?

## A TIME FOR ACTION

These problems and admitted limitations of the appropriateness method should not obscure its strengths. Appropriateness criteria are an important tool to integrate available knowledge into clinical practice, holding much promise for improving medical practice. The flaws of the appropriateness method and criteria are far from being fatal and appear to be quite correctable. The reliability as well as the validity of the method can be confirmed as it becomes more commonly used, and the dimensions of the quality of care that are currently missing could be incorporated into the criteria or at other stages in the process of care.

In fact, at the end of the Lausanne meeting, participants were clearly in favour of health professionals in European countries concentrating on efforts to implement the appropriateness method, instead of waiting for a perfect method before beginning to assess the appropriateness of health care interventions performed in European countries. The pressure to contain costs implies a choice: either we show that there are means acceptable to both clinicians and economists to reduce the volume of inappropriate care, or we run the risk that care will be soon rationed across the board, inappropriate and appropriate alike. It is therefore important that clinicians in Europe start familiarizing themselves with tools such as the appropriateness method: they should test how it can be applied, evaluate its strengths and limitations, decide which use of the method is acceptable to them, and suggest how it could be improved.

Some projects have already started in several European countries. In The Netherlands, the Dutch Inventory of Coronary Artery Treat-

ment (DUCAT) study assesses cardiovascular procedures in 10 hospitals, half of which are university hospitals. In Sweden, a prospective study is being conducted to test the same method, also for assessing cardiovascular procedures. In Switzerland, appropriateness criteria have been applied to laminectomy and upper gastrointestinal endoscopy, and other demonstration projects are being designed, for example to assess the appropriateness of carotid endarterectomy.

These regional efforts could now converge into a major international enterprise in which all European countries are invited to participate. At the end of the Lausanne meeting, it was decided to seek funding to enable interested parties to implement the appropriateness method in their respective countries, while looking for ways to share such activities as the very expensive literature review and the development of the list of indications. Large-scale economies can be expected from such a collaborative effort and comparison of data collected in a standardized manner in each institution will provide interesting and useful information on medical practice in Europe.

## REFERENCES

1. Wennberg J E, Dealing with medical practice variations: a proposal for action. *Health Affairs* 3: 7, 1984.
2. Wennberg J E, Population illness rates do not explain population hospitalization rates. *Med Care* 25: 354, 1987.
3. Wennberg J E, Freeman J L and Culp W J, Are hospital services rationed in New Haven or over-utilised in Boston? *Lancet* i: 1185, 1987.
4. Leape L L, Hilborne L H, Kahan J P, Stason W B, Park R E, Kamberg C J and Brook R H, *Coronary artery bypass graft; a literature review and ratings of appropriateness and necessity*. RAND, (Publication RAND/JRA 02). Santa Monica, 1991.
5. Brook R H, Chassin M R, Fink A, Solomon D H, Koseoff J and Park R E, A method for the detailed assessment of the appropriateness of medical technologies. *Int J Technol Assess Health Care* 2: 53, 1986.
6. Smith R, Where is the wisdom...? The poverty of medical evidence. *Br Med J* 303: 798, 1991.
7. Payer L, *Medicine and culture*. Henry Holt and Co., New York, 1988.
8. Dubois R W, Reducing unnecessary care: Different approaches to the "big ticket" and the "little ticket" items. *J Ambulatory Care Manage* 14: 30, 1991.
9. Chassin M R, Koseoff J, Park R E, Winslow C M, Kahn K L, Merrick N J, Kessev J, Fink A, Solomon D H and Brook R H, Does Inappropriate use explain geographic variations in the use of health care services? *JAMA* 258: 2533, 1987.
10. Wennberg J E, The paradox of appropriate care. *JAMA* 258: 2568, 1987.
11. Smith R, The ethics of ignorance. *J Medical Ethics* 18: 117, 1992.
12. Institute of Medicine, Field M J and Lohr K N, Eds, *Clinical practice guidelines: directions for a new program*. National Academy Press, Washington, DC, 1990.
13. Rizzo J A, Physician uncertainty and the art of persuasion. *Soc Sci Med* 37: 1451, 1993.
14. Phelps C E, The methodologic foundations of studies of the appropriateness of medical care. *N Engl J Med* 329: 1241, 1993.
15. What do we mean by appropriate health care? Report of a working group prepared for the Director of Research and Development of the NHS Management Executive. *Quality in Health Care* 2: 117, 1993.
16. Scott E A and Black N, Appropriateness of cholecystectomy in the United Kingdom—a consensus panel approach. *Gut* 32: 1066, 1991.
17. Fraser G M, Pilpel D, Hollis S, Koseoff J and Brook R H, Indications for cholecystectomy: the results of a consensus panel approach. *Quality Assurance in Health Care* 5: 75, 1993.
18. Phelps C E and Parente S T, Priority setting in medical technology and medical practice assessment. *Med Care* 28: 703, 1990.