International Journal for Quality in Health Care 1999; **Volume 11, Number 1**: pp. 13–19

An experience of utilization review in Europe: sequel to a BIOMED project

S. LORENZO[1], R. BEECH[2], T. LANG[3] AND B. SANTOS-EGGIMANN[4]

[1]Gabinete del Plan de Calidad, Fundación Hospital Alcorcón, Alcorcón, Madrid, Spain, [2]Department of Public Health Medicine, United Medical and Dental Schools of Guy's and St Thomas Hospital, London, UK, [3]Medical Informatics Unit, Paris Hospital, Paris, France and [4]Institut Universitaire de Médecine Sociale et Préventive, Lausanne, Switzerland

Abstract

Objective. To develop and test a utilization review screening tool for use in European hospitals.

Setting. In 1993 a group of researchers financed by a European Union grant reviewed the use of utilization review in Europe. They quickly noticed a lack of specifically designed instruments able to take into account the health care and cultural differences across Europe, and available for use in different health care systems. Hence, they embarked upon the task of developing and testing a utilization review screening tool for use in European hospitals.

Results. The European Union-Appropriateness Evaluation Protocol's list of reasons was developed and assessed. This is a common taxonomy that classifies days identified as unnecessary and provides a list of levels of care to identify patients' needs. This new protocol not only substitutes for the multiple previous local versions of the Appropriateness Evaluation Protocol, but will also facilitate comparisons of the varying experiences in European countries.

Main findings. Development of utilization review in Europe has been carried out mostly on a voluntary basis and the main objective was not control. The experience varies widely: from France, where utilization review is still developing and research has been implemented by local teams, to Portugal, where utilization review programmes have been initiated by government authorities. At this point different initiatives in quality improvement, and more specifically in utilization review, are being developed within the European context.

Keywords: appropriateness evaluation protocol, appropriateness of hospital use, Europe, European appropriateness evaluation protocol, utilization review

The aim of this paper is to provide insight into hospital utilization review in Europe by describing the development of a common tool, the European version of the Appropriateness Evaluation Protocol. The paper describes the authors' experiences following the process of developing of a new utilization review tool that was adapted to the European setting.

Prospective payment systems provide a basis for comparing the efficiency of hospitals. However, these systems do not indicate how well the product was delivered, and strategies to accomplish these goals may result in non-selective reduction in both appropriate and inappropriate care. The opportunity to study appropriateness of hospital utilization in a European project was thus a privileged observation point for looking at the development of utilization review in Europe.

Some characteristics of health care systems are relatively similar all over Europe. Common goals of all the health care reforms in Europe include more effective cost-containment through greater efficiency and effectiveness in service delivery whilst maintaining accessibility to health care. In addressing these goals the supply of inpatient acute care has so far received the closest scrutiny because it consumes a substantial proportion of health care spending.

One approach to controlling inpatient stay is to fund hospitals according to an agreed cost per case. An example of this approach is the utilization of case-mix measures [1] to gauge hospital product: e.g. the Diagnostic Related Groups (DRG) and Patient Management Categories (PMC) prospective payment systems as developed and used in the USA. Systems such as DRGs and PMCs seek to identify iso-resource groups in the acute care setting. Similar groupings are being developed for long-stay care [2] (RUG) and for

Address correspondence to Susana Lorenzo, Gabinete del Plan de Calidad, Fundación Hospital Alcorcón, Avda. Villaviciosa s/n, 28922 Alcorcón, Madrid, Spain. Tel: +34 91 621 9538/9400. Fax: +34 91 621 9902. E-mail: slorenzo@fhalcorcon.es

out-patient/ambulatory [3] settings: Ambulatory Visit Groups and Ambulatory Patient Groups. Such groupings provide a basis for comparing the efficiency of hospitals [4] and for defining how much health care purchasers should pay for delivering each unit of the product.

When case-mix groupings are used for reimbursement purposes, hospitals have the incentive to ensure that their costs of delivering these identified products do not exceed the agreed price [5]. The DRG system of reimbursement has resulted in a reduction of in-patient lengths of stay in the USA and in Europe [6–8]. However, these groups do not indicate whether the product was delivered to the right consumer nor do they indicate how well the product was delivered.

To address this problem, utilization review (UR) is used to determine whether specific health care services are medically necessary and whether they are provided at an appropriate level of intensity and cost [9]. The concept was developed in the late 1970s [10] with attempts to improve the management of health care resources.

The potential benefits of UR are: reducing unnecessary hospital utilization; improving the quality of care by reducing the chance of nosocomial infections or iatrogenic illness; maintaining quality of care by assuring that the hospital services provided are of sufficient duration, frequency and level of care to promote optimal health outcome; and preserving access and defining and articulating standards of care [11].

Efficiency and cost containment have become primary goals of our health care system. However, strategies to accomplish these goals may result in non-selective reduction in both appropriate and inappropriate care. It is difficult to ascertain the exact savings attributable to UR activities. Most savings reported in the literature do not appear to be based on a scientifically rigorous approach of the literature [9], whereas other authors, such as Brook [12] concluded that UR had little effect on cost control.

UR was initially developed and has been most widely used in the USA, particularly as a means to control hospital utilization under Medicare and Medicaid programmes. Here, economic pressures forced UR to focus on cost control, although the demarcation between UR and quality assurance activities is again becoming less clear [9]. In the USA the external influence of private health care purchasers has been a dominant force behind the development of UR. In most European countries, the method of funding health care means that such external pressures on health care providers are not as evident. In our setting, hospitals are currently evolving from a system based on centralized decision-making and follow-up of administrative processes that did not pay much attention to efficiency, towards models based on greater autonomy and responsibility, while paying special attention to the results obtained.

Variations in hospital admission rates across geographic areas are nearly universal, and the reasons for the differences are not well understood. Studies of the role of need, demand and supply have had inconclusive results [13–14]. Differences in physicians' practice patterns, access to and availability of timely, effective ambulatory care and inappropriate use have been suggested as reasons for those variations.

The appropriateness of hospital utilization European project: a privileged point of view to observe the development of UR in Europe

A hospital could (in theory) deliver very poor quality, unnecessary care at low cost and high efficiency [6]. In order to increase hospital efficiency and cost-effectiveness, managers in the US and Europe are devoting considerable effort to reducing unnecessary days of hospital care. There is still no definite evidence that a reduction in length of stay is associated with a parallel decrease in the numbers of unnecessary or inappropriate days of inpatient hospital care.

There are two types of inappropriate hospital utilization. Over-utilization is care which is of no benefit for the patient or which can be provided in a lower level, less costly setting. Under-utilization is care that is not sufficient in type, length, location or intensity to meet the patient's medical need [15]. Under-utilization is much more difficult to identify than over-utilization so that most of the studies that have been conducted focus on over-utilization. When used in this way UR specifically seeks to highlight inappropriate days of inpatient care and the reasons why these are occurring.

As Donabedian stated in 1982, UR methods cannot be considered anything else but screening devices, not the final, absolute indicators of quality or appropriateness [16]. The procedures used to identify inappropriate hospital utilization have been categorized according to whether they use implicit criteria, explicit criteria or a combination of both. Whereas implicit techniques are based on the reviewer's opinion, explicit methods rely on criteria auditing. Reviewers tried to develop methods that were standardized, transferable, reliable, valid and explicit [11].

Explicit review methods provide specific criteria for the reviewers. Some instruments are diagnosis-specific and some diagnosis-independent. Whereas DRG-prospective reimbursement systems lead to decreases in hospital lengths of stay, diagnosis-specific methods apply to categories of patients with specified diagnosis or signs or symptoms. The resources required for the development of these techniques may limit the analysis to certain diagnostic groups. They are limited to a few sets of specific diagnosis (i.e. coronary angiography and cholecystectomy) [17], and reviewers refer to difficulties in assigning diagnosis.

Utilization review is concerned mainly with operational efficiency and appropriateness. There are several techniques that hospital managers currently use to analyse hospital efficiency – from case-mix analysis to activity-based costing. At the same time hospital managers analyse the appropriate use of resources, comparing data on hospital utilization, focusing on drug [18,19] or diagnostic procedure use [20–24].

However, most of the research studies conducted so far have focused on the review of appropriateness of hospital

use and patient satisfaction. We will focus on appropriateness of hospital use. The studies that have been carried out were conducted using the existing generic diagnosis independent criteria lists developed in the USA in the Professional Standards Review Organizations and Peer Review Organizations movements:

- The Standarized Medreview Instrument [25]. Authors refer to low validity and reliability and it is not being used.
- The Intensity of Service, Severity of Illness and Discharge Screens review system. This instrument diffused in the USA but is not widely used in Europe. It has good reliability and moderate validity, and has been used in a few studies [26–28].
- The Appropriateness Evaluation Protocol (AEP). The validity and reliability of this instrument has been evaluated in the USA [26,29], Israel [30], Italy [31] and Spain [6,32] with satisfactory results. In the USA, it has recently been adjusted to the managed care environment (Managed Care Appropriateness Evaluation protocol).
- The Delay Tool classifies all medically unjustified hospital stays. This instrument – validated for the paediatric population in the USA [33] – assigns causes for medically unnecessary hospital days detected through other instruments. It has been used in the USA [33], Spain [34], and Switzerland [35].
- The Oxford Bed Study Instrument, based on the AEP has been used in the UK to identify factors that act as bed-blockers [36,37].

The development of a specific European UR tool

In 1993 a group of researchers financed by a European Union grant (BIOMED 1) reviewed the use of UR in Europe. Different expert groups from seven European countries participated in this process: Austria, France, Italy, Portugal, Spain, Switzerland and the UK. They quickly noticed a lack of specifically designed instruments able to take into account the health care and cultural differences across Europe, and that could be used in different health care systems [38]. Hence, the group embarked upon the task of developing and testing a UR screening tool for use in European hospitals. After a brief review of the findings and methods used in previous UR studies in Europe the remainder of this paper describes the way that the research group collaborated to develop and test a European instrument as well as future trends in UR in Europe.

While the growth of UR in Europe has so far been sporadic, it is largely recognized within Europe that at least some part of utilization of hospitals is inappropriate and that hospital patients receive services that provide no significant benefit [38]. Acute hospital beds are an expensive and scarce resource for which there is a high demand. It is therefore imperative that such resources are used efficiently and effectively. Consequently there has been considerable interest expressed by researchers, policy makers and health care professionals in the appropriate use of the highest level of inpatient care.

In the European setting, UR has avoided the assessment of the appropriateness of surgical procedures and diagnostic tests in order to concentrate on the pattern of physicians and patient's drug utilization and hospitalizations.

European activities in UR are usually limited to the retrospective review of medical records [39] based on the AEP that is the most commonly used evaluation instrument [38]. Judgement about the need for admission is based on the information available in medical records until the end of the day of admission, whereas the day of stay is assessed according to the information available up to the day of review. Given that the validity and reliability of the AEP has been evaluated worldwide [6,26,29–32], with satisfactory results, and because of its predominance in European UR studies, the research group decided to adapt the Adult Medical–Surgical version of the US review instrument (US-AEP) to the European setting. The European version of the AEP (EU-AEP) was conceived to help harmonize utilization review in Europe. The EU-AEP is based on the multiple adaptations and modifications that have been made to the US-AEP by European UR researchers [40]. Linguistic, conceptual and technical issues arose during the process of adaptation of the US instrument to the European setting. Consensus was needed between the research group members on the criteria that were going to be modified in the protocol, given the existing differences among the participating countries. The main differences were not only cultural, but in the organization and financing of the health systems of the different countries.

According to the literature, the estimated rates of inappropriate hospitalizations worldwide range from 15 to 30% [41]. Previous research shows that in Europe inappropriate hospital use figures are similar [38,42–47] although the publication of the studies that have been performed is not so frequent. UR studies in Europe were conducted by independent groups of researchers with different health care contexts [38]. The development of UR has been done mostly on a voluntary basis and the main objective was not control. The experience within European countries varies widely, from France where UR is not very developed and research has been implemented by local teams, to Portugal where UR programmes have been initiated by government authorities.

Local experiences on UR

Austria

Even though quality assurance activities have been diffused in the Austrian health service, Austrian researchers have not published any study on UR. After their involvement in the BIOMED project, UR activities are currently being introduced.

Italy

As the Italian health care system is undergoing dramatic changes, UR review tools are bound to become of great

importance in monitoring the effects of these changes. UR activities started in the 1980s using implicit judgements and evolved to others conducted using local translations of the AEP. The assessment of the effectiveness and efficiency of hospitals in order to improve their activity was the main goal. The lack of impact of the initial studies on UR in the Italian setting has been attributed to the lack of structural incentives [47]. Overall, the different studies conducted in Italy suggest, so far, that there is a large proportion of inappropriate hospital use in that country. Several investigators have described feedback programmes that have aimed at the appropriate use of drugs and transfusions, use of diagnostic tests, etc.

Italian researchers are currently conducting UR studies in this country, and the EU-AEP is being translated into Italian.

France

The development of UR in France is evolving rapidly. Implementation of hospital information systems derived from the DRG system, which are primarily financially driven, are now in place in most hospitals. The governmental health authorities initiated these systems. On the other hand, quality assurance and UR activities have not been developed systematically. They have been organized on a voluntary basis, with various methods and importance according to the attitudes and perception of evaluation by the hospital managers [48,49]. The institutions, which have been developed (National Agency and National Committee for Medical Evaluation) have so far been involved primarily with consensus conferences and clinical guidelines development. A hospital accreditation system is currently being implemented. One of the criteria for accreditation is the evaluation of activity in hospitals. Accordingly, although the AEP has been validated and used in 1990 hospital admissions, its diffusion has been modest [49]. The validation of the European version of the AEP, concerning hospital days, has received more attention within the institutions. As far as medical procedures are concerned, the experience has been limited so far. However, important disparities between regions have been reported, concerning, for example, the probability of getting appropriate microligation and treatment after an acute coronary event [50]. An increasing interest in UR in the near future may be anticipated.

Portugal

In Portugal, the National Health Service plays a major role in inpatient care. In this country, utilization review was initiated by the government health sector, with the primary goals of enhancing decision-making methods and decision support systems. Sets of information systems were designed to assess the inappropriate use of hospital inpatient resources. The analysis [51] showed that the success of UR would depend on the involvement of the physicians in the design and implementation. A pilot phase in a few hospitals was thus organized and the review performed by the physicians. The project began in the mid-1980s and a large database is currently available. After the pilot phase, the project has been expanded to other hospitals and it is being directly supported by the Ministry of Health. The DRG's database is being used to direct UR activities toward areas where problems are more likely to occur.

Spain

Studies assessing hospital utilization review in Spain started in the early 1980s, whereas some others evaluating appropriateness of ambulatory care are starting at the moment. Almost all of the studies conducted in hospitals were performed retrospectively using different local adaptations of the AEP – a large proportion of them used the medical–surgical version of the protocol and the paediatric version is currently being validated. The identified determinants of inappropriate hospitalization in Spain are related mainly to access to the different levels of care and to conservative attitudes of the physicians. Current studies are introducing the EU-AEP which has been already translated into Spanish.

UR activities are starting to become part of quality assurance programmes. These programmes are being created in almost all of the public hospitals. In the future, UR is expected to be extensively diffused, not only through research groups but also by governmental agencies. Current studies are focusing on medical procedures and monitoring through UR the efficiency of focusing on the most prevalent diagnosis [34] (hernia repair, appendectomy, etc.). Intervention programmes – associated with feedback to clinicians – are currently being introduced in order to evaluate the impact of the interventions. Some of these UR programmes might be associated with economic incentives in some cases.

Switzerland

Hospital utilization reviews were initiated in a limited number of hospitals in the very late 1980s and early 1990s [44]. They initially took place in four hospitals in Canton de Vaud receiving a global, prospective, public funding. The public health department triggered the reviews that purported to promote the hospital performance. Hospital utilization reviews were conducted on a concurrent basis, using local adaptations of the AEP together with Selker's classification for the inappropriate days detected [44]. In parallel, and in recent years, university-affiliated hospitals in Vaud and Geneva conducted sectorial and limited hospital utilization reviews. Evaluators used the original version of the AEP and explored the potential of their routine hospital information systems.

The economic crisis of the 1990s induced profound deficits in public budgets, and many cantons reacted with economic measures that were considered particularly necessary in the costly sector of acute care hospitals. This fact prompted the generalization of hospital utilization reviews in the whole of Canton de Vaud; their main objective is to contribute to the economic effort to reduce the health care public expenditure, with the assumption that hospital utilization reviews, used as an internal control tool, will improve hospital performance. As a result, a hospital utilization review covering 6 consecutive months is now under way in 16 publicly funded regional hospitals in Vaud, using a validated French translation of the EU-AEP.

The methodology adopted by the regional hospitals in Vaud is now rapidly diffusing to all other French-speaking cantons of Switzerland and, as two bilingual cantons are now involved in the process, hospital utilization reviews are likely to expand further to other regions.

UK

Since the 1950s UR studies have been reported in the scientific literature with the number of studies reported increasing in recent years. This is due to increasing pressures on acute beds because of factors such as the ageing population, the rise in emergency admissions, and the development of strategies that aim to transfer elements of care away from the hospital towards the community sector. However, in spite of this rise in published studies UR is not routinely used by health care purchasers and providers as part of their health services planning.

Discussion

Shared goals of all the attempts at reforming the health care systems in Europe include more effective cost-containment, greater efficiency and effectiveness in service delivery, and maintaining accessibility to health care. Due to the trend to shift financial risk from health insurers to provider groups we can expect that performance measures will be routinely shared by plans and providers. UR has been very diversely developed in different European countries, the level of diffusion is quite variable and it is evolving rapidly.

At the same time, standards of quality are being developed to evaluate physicians and health care organizations [52], these standards are based on existing controlled clinical studies through the evidence-based medicine approach [53, 54]. Different authors agree that although the methods of appropriateness studies have evolved and improved, they have still substantial methodological shortcomings [41,52, 55–58]. Given that methodological problems are still numerous in many aspects of UR, the diffusion of UR in Europe might be slowed down and we considered that research should be developed in this field.

At the same time, the fact that European UR researchers do not belong to intervention programmes towards the reduction of unnecessary use should be taken into account [46]. The impact of these evaluations is thus lower than in the USA. The real challenge is to see if the use of hospital beds can be improved even though there are, as yet, no published reports on the impact of these activities [55].

Cross-national organizational learning is becoming a method in the repertoire of quality improvement. It is a method that enables us to make faster and more cost-effective quality improvements [59]. The technique has already been applied in UR [40,42–44,47] and the EU-AEP is an instrument that can be used in comparisons between countries (S Lorenzo, T Lang, R Pastor, A Tampieri, B Santos-Eggimann, H Smith, A Liberati and J Restuccia, unpublished work). Traditionally Europe has followed the experiences of the USA, applying them later. At this point different initiatives in quality improvement, and more specifically in UR, are being developed within the European context, and some of them might be applied in the future in the USA. Currently, European reviewers are using the European adaptation of the AEP performed by the BIOMED project as a common tool that not only substitutes the multiple previous local versions of the AEP but that will also facilitate comparisons with the experience in other European countries. The EU-AEP list of reasons is a common taxonomy that classifies days identified as unnecessary, providing a list of levels of care to identify the patients necessities that might be useful to other countries such as the USA.

Acknowledgments

This article is based on a presentation given by Susana Lorenzo at the ISQua conference in Chicago in November 1997. It reflects the experience of the research team supported by an European Union Grant BIOMED I on Appropriateness of Hospital Use coordinated by Alessandro Liberati at Mario Negri Institute (BMH1 CT93 1053).

The authors would like to express their gratitude to all the members of the BIOMED research group on appropriateness of hospital use for their support and collaboration: Christian Koeck, Maria Schmidt, Nicole Scholtz (Austria); Bernard Huet, Thierry Lang, Hélène Logerot, Elisabeth Monnet, Patrick Six (France); Giovanni Apolone, Guido Fellin, Alessandro Liberati, Guglielmo Meregalli, Antonio Tampieri, Mauro Venegani (Italy); Margarida Bendes, Maria Da Luz Gonsalves, Elaine Pina, Margarida Santos, Joao Urbano, Valdo Costa (Portugal); Jordi Gol, Susana Lorenzo, Roberto Pastor, Rosa Suñol (Spain); Thierry Blanc, Fred Pacaud, Briggite Santos-Eggimann (Switzerland); Giles Glover, Cara O'Neill, Declan O'Neill, Maggie Pearson, Helen Smith (UK).

References

1. Sanderson H, Mountney L. The development of patient groupings for more effective management of health care. *Eur J Publ Health* 1997; **7:** 210–214.

2. Fries BE, Schneider DP, Foley WJ. Refining a case mix measure for nursing homes resource utilization groups (RUG-III). *Med Care* 1994; **32:** 668–685.

3. Parkin D, Hutchinson A, Philips P, Coates J. A comparison of diagnosis related groups and ambulatory visit groups in day-case surgery. *Health Trends* 1993; **25:** 41–44.

4. Muller A. Medicare prospective payment reforms and hospital utilization: temporary or lasting effects? *Med Care* 1993; **31:** 296–298.

5. Fetter RB. Hospital payment based on diagnosis-related groups. *J Soc Health Syst* 1992; **3:** 4–15.

6. Alonso J, Muñoz A, Antó JM. Using length of stay and inactive days in the hospital to assess appropriateness of

utilization in Barcelona, Spain. *J Epidemiol Commun Health* 1996; **50:** 196–201.

7. Ruiz-Lapuente MA, Royo Bordonada MA, Bermejo Fraile B *et al.* Appropriateness of admission and hospital stays in inguinal hernia without complications (in Spanish). *Med Clin (Barc)* 1997; **108:** 259–262.
8. Oterino D, Peiró S, Marchan C, Portella E. Inappropriate hospitalization: reasons and determinants. *Eur J Publ Health* 1996; **6:** 126–132.
9. Ermann D. Hospital Utilization Review: past experience, future directions. *J Health Politics, Policy Law* 1990; **5:** 683–704.
10. Hofdijk J. The development of patient groupings for more effective management of health care. *Eur J Publ Health* 1997; **7:** 215–221.
11. Payne SMC, Restuccia JD, Ash A *et al.* Using utilization review to improve hospital efficiency. *Hosp Health Serv Admin* 1991; **36:** 473–490.
12. Leape LL, Rolla EP, Solomon DH *et al.* Relation between surgeons' practice volumes and geographic variation in the rate of carotid endarderectomy. *N Engl J Med* 1989; **321:** 653–657.
13. Halfon N, Newacheck PW. Childhood asthma and poverty: differential impacts and utilization of health services. *Pediatrics* 1993; **91:** 56–61.
14. Diehr P, Cain K, Connell F, Volinn E. What is too much variation? The null hypothesis in small area analysis. *Health Serv Res* 1990; **24:** 741–771.
15. Payne SMC. Identifying and managing inappropriate hospital utilization. A policy synthesis. *Health Serv Rev* 1987; **22:** 709–769.
16. Donabedian A. *Explorations in Quality Assessment and Monitoring: The Criteria and Standards of Quality,* vol. II. Ann Arbor, MI: Health Administration Press, 1982.
17. Winslow CM, Solomon DH, Chassin MR *et al.* The appropriateness of carotid endarderectomy. *N Engl J Med* 1988; **318:** 721–727.
18. Buetow SA, Sibbad B, Cantrill A, Halliwell S. Prevalence of potentially inappropriate long-term prescribing in general practice in the United Kingdom; 1980–95: systematic literature review. *Br Med J* 1996; **313:** 1371–1374.
19. INSALUD. *Pharmaceutical Use Indicators in the National Health Institute* (in Spanish). Madrid: INSALUD (National Health Institute), 1996.
20. Doughty A, Nash SI, Gift DA. Deployment and utilization of MR imaging in Michigan: observations of a statewide data base. *Radiology* 1992; **185:** 53–61.
21. Marco F, Perez-Caballer AJ, Sanchez Barbero JC *et al.* Arthroscopic verification of semilum cartilage by Nuclear Magnetic Resonance Diagnosis (in Spanish). *Rev Esp Cir Osteoart* 1993; **28:** 213–217.
22. Bransby-Zachary MAP, Sutherland GR. Unnecessary X-ray examinations. *Br Med J* 1989; **298:** 1294.
23. Power KJ, Norman J. Pre-registration house surgeons. A questionnaire study of anaesthesia-related knowledge and approach to preoperative investigations. *Anaesthesia* 1992; **47:** 518–522.
24. Walker D, Williams P, Tawn J. Audit of requests for preoperative chest radiography. *Br Med J* 1994; **309:** 772–773.
25. Systemetrics. *Standarized Medreview Instrument: SMI: Reviewer's Training Manual.* Santa Barbara, CA: McGraw Hill, 1984.
26. Strumwasser I, Paranjpe NV, Ronis DL *et al.* Reliability and validity of utilization review criteria. Appropriateness evaluation protocol, standardized review instrument and intensity–severity–discharge criteria. *Med Care* 1990; **28:** 95–111
27. Inglis AL, Coast J, Gray SF *et al.* Appropriateness of hospital utilization. The validity and reliability of the Intensity–Severity–Discharge review System in a United Kingdom acute hospital setting. *Med Care* 1995; **33:** 952–957.
28. Coast J, Peters TJ, Inglis A. Factors associated with inappropriate emergency hospital admission in the UK. *Int J Qual Health Care* 1996; 8: 31–39.
29. Gertman PM, Restuccia JD. The Appropriateness Evaluation Protocol: a technique for assessing unnecessary days of hospital care. *Med Care* 1981; **19:** 855–871.
30. Rishpon S, Lubacsh S, Epstein LM. Reliability of a method of determining the necessity for hospitalization days in Israel. *Med Care* 1986; **24:** 279–282.
31. Apolone G, Fellin G, Tampieri A *et al.* Appropriateness of hospital use. Report from an Italian study. *Eur J Publ Health* (in press).
32. Peiró S, Portella E. Reliability of the appropriateness evaluation protocol (in Spanish). *Med Clin (Barc)* 1996; **107:** 124–129.
33. Selker HP, Beshansky JR, Pauker SG, Kassirer JP. The epidemiology of delays in a teaching hospital. The development and use of a tool that detects unnecessary hospital days. *Med Care* 1989; **27:** 112–129.
34. Lorenzo S. Utilization Review: A Spanish overview (in Spanish). *Rev Calidad Asistencial* 1997; **12:** 140–146.
35. Santos-Eggimann B, Sidler M, Schopfer D, Blanc T. Comparing results of concurrent and retrospective designs in a hospital utilization review. *Int J Qual Health Care* 1997; **9:** 115–120.
36. Anderson P, Meara J, Brodhurst S. Use of hospital beds: a cohort study of admissions to a provincial teaching hospital. *Br Med J* 1988; **297:** 910–912.
37. Victor CR, Khakoo AA. Is hospital the right place? A survey of inappropriate admissions to an inner London NHS trust. *J Publ Health Med* 1995; **16:** 286–290.
38. Liberati A, Apolone G, Lang T, Lorenzo S. A European project assessing the appropriateness of hospital utilization: background, objectives and preliminary results. *Int J Qual Health Care* 1995; **7:** 187–199.
39. Restuccia JD. The evolution of hospital utilization review methods in the United States. *Int J Qual Health Care* 1995; **7:** 253–260.
40. Lang T, Liberati A, Tampieri A *et al.* An European version of the Appropriateness Evaluation Protocol: goals and presentation. *Int J Tech Assess* (in press).
41. Phelps CE. The methodologic foundations of studies of the appropriateness of medical care. *N Engl J Med* 1993; **329:** 1241–1245.

42. Lang T, Davido A, Logerot H, Meyer L. Appropriateness of admissions: the French experience. *Int J Qual Health Care* 1995; **7:** 233–237.

43. Lorenzo S, Suñol R. An overview of Spanish studies on appropriateness of hospital use. *Int J Qual Health Care* 1995; **7:** 213–218.

44. Santos-Eggimann B, Paccaud F, Blanc T. Medical appropriateness of hospital utilization: an overview of the Swiss experience. *Int J Qual Health Care* 1995; **7:** 227–232.

45. Bañeres J, Alonso J, Broquetas J, Antó JM. Inappropriate hospital admissions and days of stay in neoplasic and chronic pulmonary obstructive disease patients (in Spanish). *Med Clin (Barc)* 1993; **100:** 407–411.

46. Oterino D, Peiró S, Marchan C, Portella E. Inappropriate hospitalization: reasons and determinants. *Eur J Publ Health* 1996; **6:** 126–132.

47. Fellin G, Apolone G, Tampieri A, Liberati A. Appropriateness of hospital use: an overview of Italian studies. *Int J Qual Health Care* 1995; **7:** 219–226.

48. Amouretti M, Beraud C, Saint-Martin E. Medical audit in France: from ideal to reality. *Br Med J* 1992; **304:** 428–430.

49. Giraud A. Medical audit in France: historical perspectives. *Br Med J* 1992; **304:** 426–428.

50. French Network of the Ischemic Heart Disease Registries, MONICA. *Myocardial Infarction in France. Results from the Three French Registries, Lille, Strasbourg and Toulouse 1985–92* (in French). Paris: Fédération Française de Cardiologie, 1996.

51. Bentes M, Da Luz G, Santos M, Pina E. Design and development of a utilization review program in Portugal. *Int J Qual Health Care* 1995; **7:** 235–241.

52. Kassirer JP. The quality of care and the quality of measuring it. *N Engl J Med* 1993; **229:** 1263–1265.

53. Epstein AM. The outcomes movement – will it get us where we want to go? *N Engl J Med* 1990; **323:** 266–270.

54. Evidence Based Medicine Working Group. Evidence based medicine. *J Am Med Assoc* 1992; **268:** 2420–2425.

55. Black N. European collaboration on hospital bed use: a commentary. *Int J Qual Health Care* 1995; **7:** 185–186.

56. Lorenzo S. Utilization review methods: limitations (in Spanish). *Med Clin (Barc)* 1996; **107:** 22–25

57. Wennberg JE. On the appropriateness of small-area analysis for cost-containment. *Health Affairs* 1996; **15:** 164–167.

58. Dresser M, Feingoldn L, Rosenkranz SL, Coltin KL. Clinical quality measurement. Comparing chart review and automated methodologies. *Med Care* 1997; **35:** 539–552.

59. Øvretveit J. Would it work for us? Learning from quality improvement in Europe and beyond. *J Qual Improve* 1997; **23:** 7–22.

Accepted for publication 16 September 1998