Chapter 1: Introduction

The annotated bibliography consists of articles that are directly or indirectly related to the RAND Appropriateness Method (RAM). The RAM is a method to determine appropriate care. A procedure is used appropriately if its application - to a particular patient - provides the expectation of more benefit than harm. Benefits and harms are defined in terms of the patient’s physical, functional, and subjective well-being. The following presents a short overview of the different steps involved in studies using the RAM.

When a topic has been selected, the first step of a study using the RAM is to perform a detailed literature review to synthesise the latest available scientific evidence on the procedure to be rated. At the same time, a list of indications is produced in the form of a matrix which categorizes patients who might present for the procedure in question in terms of their symptoms, past medical history and the results of relevant diagnostic tests.

The literature review and the list of indications are sent to the members of an expert panel. These panel members individually rate the appropriateness of using the procedures for each indication on a nine-point scale, ranging from extremely inappropriate (=1) to extremely appropriate (=9) for the patient described in the indication. The panel members assess the benefit-risk ratio for the "typical patient with specific characteristics receiving care delivered by the typical surgeon in the typical hospital."

After this first round of ratings, the panel members meet for 1 or 2 days under the leadership of a moderator. During this meeting, the panellists discuss their previous ratings, focusing on areas of disagreement, and are given the opportunity to modify the original list of indications and/or definitions. After discussing each chapter of the list of indications, they re-rate each indication individually. The two-round process is focused on detecting consensus among the panel members. No attempt is made to force the panel to consensus. In examining the potential use of the procedure for these indications, the method can determine situations when the procedure is inappropriate (that is, the risks outweigh the benefits) or necessary (that is, the procedure is the only possible means of providing substantial benefit for the patient).

The articles on the RAM have been grouped in the following categories:

A) Descriptions of RAND Appropriateness Methodology (RAM)
   A.1) RAM: general descriptions and basic definitions
   A.2) More general papers in which RAM is discussed
   A.3) Necessity: general descriptions and basic definitions
   A.4) Methodological issues
   A.5) Supporting software

B) RAM outcomes (literature reviews, panel ratings or chart audits by topic)
   B.1) Assessment of appropriateness of care (possibly with B.2 or B.3)
   B.2) Panel results (possibly with B.3)
   B.3) Literature review

C) Within panel comparisons
   C.1) Across panelist characteristics
   C.2) Across methods of assessing appropriateness
   C.3) Across countries/regions
D) Multiple panel comparisons
   D.1) Across panelist characteristics
   D.2) Across methods of assessing appropriateness
   D.3) Across countries/regions

E) Non-RAM studies related to appropriateness

Where an author abstract is available, we have copied it in this bibliography. The categories are not mutually exclusive which means that some articles fit in more than one category. In such cases, the article is placed in what we believe is the most appropriate category. Within each category, the articles are ordered chronologically. The RAND Europe Appropriateness library contains most of the articles, reports and books listed in the annotated bibliography. We are very interested in hearing from you if you:

- know of articles, reports or books that should be added to the annotated bibliography;
- can send us copies of the articles, reports or books listed in the annotated bibliography that we do not have copies of;
- have suggestions to improve the annotated bibliography;
- want to know more about RAND, RAND Europe or the RAND Appropriateness Method.

Please contact:

Mirjam van het Loo
c/o RAND Europe
Newtonweg 1
2333 CP LEIDEN
The Netherlands
tel: +31 524 51 70
fax: +31 524 51 91
e-mail: m.vanhetloo@randeurope.org
Chapter 2: Content-coded listing of articles related to the RAND Appropriateness Method (RAM)

A) Descriptions of RAND Appropriateness Methodology (RAM)

A.1) RAM: general descriptions and basic definitions


The standard way to assess medical technologies is to conduct a randomized clinical trial. Patients are randomly assigned to groups receiving alternative treatments, and outcomes are monitored over a long period of time. Randomized clinical trials provide invaluable knowledge about the outcome of medical technologies, but they suffer from several disadvantages. They are time-consuming and expensive, sample sizes are limited, and results from multiple trials may be contradictory. In addition, randomized trials are often performed under ideal rather than usual circumstances, and thus their results must be generalized cautiously. In order to judge appropriateness, a method that combines data from trials and other studies with the opinions of experts is needed. This note describes such a method, and discusses its application to rate the appropriate-ness of six medical and surgical procedures: coronary angiography, coronary artery bypass surgery, cholecystectomy, upper gastrointestinal endoscopy, colonoscopy, and carotid endarterectomy.


Inappropriate health care is certainly costly and possibly injurious. Research assessments of the rate of delivery of inappropriate health care have shown substantial variations in the delivery of inappropriate care among providers and geographic regions. Clinical practice guidelines, presenting the collective judgment of clinical experts, have been promected as a means to decrease undesirable variation in physician practice and thereby improve the quality of health care. This report presents the results of a pilot study of the use of the appropriateness method as part of the Agency for Health Care Policy and Research Clinical Practice Guideline development process.


This paper describes a collaborative effort among several organizations in the United States (the RAND/American Medical Association (AMA/Academic Medical Center
Consortium (AMCC) Clinical Appropriateness Initiative) and presents, as an illustrative example, one of the four projects of this initiative, the abdominal aortic aneurysm surgery project that was directed by the author. The Clinical Appropriateness Initiative was conceived as a long-term program to develop and maintain a current set of clinical practice guidelines that would play an integral role in enhancing the quality of care delivered in the USA. However, several problems were encountered in its inception that ultimately resulted in early termination of the project. The primary reason for the demise of the abdominal aortic aneurysm surgery project and the other components of the Clinical Appropriateness Initiative was failure to enlist the support of AMCC clinical leaders prior to initiating the project. Numerous opportunities to optimize the generation and implementation of the practice guidelines for the intention of improving health care quality were identified during the conduct of the Clinical Appropriateness Initiative. Health care quality improvement initiatives throughout Europe may benefit from the RAND/AMA/AMCC experience and have the ability to address many of the scientific and policy issues raised by the Clinical Appropriateness Initiative.


The paper describes the RAND/UCLA appropriateness method and discusses reasons for the methodological choices that are made. The RAND/UCLA appropriateness method produces within a 6-month to 1-year period, (1) a detailed literature review; (2) outcome evidence tables that cover the areas of efficacy, cost, complications, and use; (3) a structure of indications for all possible uses of the procedure; and (4) ratings for each of the indications, both in terms of appropriateness and necessity. This information could be used to form clinical guidelines for use in physician offices or hospitals, to provide information to patients, or to help third-party carriers or regulators improve both the quality and appropriateness of care. Reliability and validity of the RAND/UCLA method have been tested, but much more methods work needs to be done. This additional investment in research should result in a better method to assess appropriateness.


A.2) More general papers in which RAM is discussed


This paper reports on the need for more research on outcomes and appropriateness of medical care to develop a more rational system for preventing, punishing, and compensating for medical malpractice. Malpractice claims are clearly linked to outcomes of care, and a large proportion of court decisions settling these claims confuse bad outcomes with inappropriate care because of the lack of definitive research on how specific medical interventions affect patient outcomes. The author suggests that the medical malpractice system might be improved by conducting more research on outcomes and appropriateness of care to inform medical decision-making and to set standards relevant only to a certain class of malpractice problems, ones in which it was not appropriate to take a specific action. Developing more explicit standards of care treats both doctor and patient more equitably by offering a statement about range of expected risks and benefits of the intervention under consideration as well as providing a structure to ease problems of malpractice.


Concerns regarding significant levels of inappropriate medical services--as high as 20% or more--continue to influence discussions regarding medical care quality, utilization, and costs. The basis of these concerns are findings from a series of studies of the appropriateness of use of several medical and surgical services provided in the late 1970s and early 1980s--cardiac pacemaker implants, carotid endarterectomy, coronary artery bypass surgery, coronary angiography, and upper gastrointestinal endoscopy. More recent data from Medicare peer review organizations, however, indicate lower levels of unnecessary hospital admissions and medical services. Despite uncertainties regarding the extent of inappropriate care, additional efforts are required to better define appropriate medical care. A promising effort to meet this need is the development of practice parameters, which include practice guidelines and standards.


Resources available to provide adequate health care in western countries must compete with other priorities such as education and the environment. At the same time the allocation of health resources often does not correspond to the real needs of patients. We have developed a method that combines what is known in the literature with expert
physician judgment to generate clinically valid appropriateness guidelines. The method involves a modified Delphi approach, including a detailed literature review, consultations with experts, and three rounds of panel ratings. Clinical scenarios, or indications, are rated appropriate, uncertain, or inappropriate. Appropriate means that the procedure is worth doing for the given indication if the health benefit exceeds the health risk. We have conducted panels relating to coronary procedures in four countries. Application of the indications to individual cases has demonstrated that the amount of inappropriate care is too large to be ignored. Dissemination of appropriateness results might take the form of public disclosure or as part of the physician/patient exchange to improve performance. Indications for which a procedure is frequently performed and which are rated uncertain should be considered to be the focus of controlled clinical trials.

A.2.05 Working Group for the Director of Research and Development of the NHS Management Executive, “What do we Mean by Appropriate Health Care?”, Quality in Health Care, Volume 2, 1993, pp. 117 - 123.

Conclusion: The working group believes that there should be a wider public and professional debate about the meaning of appropriate care, and a wide debate also about linking public and professional perspectives of appropriateness. Allocation of resources will depend in part upon this, as well as on evidence of efficacy and effectiveness. The working group hopes that funding will be available to explore many of the research issues described in the article.

A.2.06 Ballard, D.J., “Lo sviluppo di linee guida; l’esperienza del programma americano sull’ approprietezza clinica e le sue implicazioni per i sistemi sanitari europei”, l’Ospedale, Volume 47, No. 9, 1994, pp. 171-184.


This paper explores assumptions inherent in results generated by a leading measure of appropriateness and concludes that there are considerable uncertainties about the measure’s meaning, the magnitude of bias that it contains, and the degree to which its application can be generalized. Some of these uncertainties could be resolved if the tacit assumptions inherent in the generation of the criteria could be made explicit. Existing measures of appropriateness are not yet sufficiently robust to be used with confidence to influence or control the delivery of health care. They may have a use as an aid rather than a constraint in clinical decision making. A randomized controlled trial could resolve whether patients achieve better outcomes if their care is influenced by appropriateness criteria.


A major focus of the current health care debate is the notion that a substantial proportion of the health care delivered in Canada is inappropriate. There are two types of appropriateness: appropriateness of a service and appropriateness of the setting in which care is provided (i.e., inpatient v. outpatient or home care). Measuring both types objectively requires the comparison of observed patterns of care with explicit criteria for appropriate care. The few studies of appropriateness conducted in Canada have shown that inappropriate services are provided and inappropriate settings are used. Reducing inappropriate health care delivery could involve active strategies for the implementation of guidelines and better cooperation and coordination within the health care system. However, lower rates of health care delivery or even inappropriate health care will not necessarily translate into higher quality care or lower costs overall.


As health expenditures continue to consume increasingly large proportions of national budgets, ways must be found to assume that money spent on health is used for effective services. Only a small proportion of medical decisions are based on scientific evidence about their outcomes, thus it is not surprising that wide variations exist in clinical practice. In recent years, questions have been raised about how medical decisions are made and the proportion of medical procedures and services that are performed for appropriate reasons. One method that has been developed to quantify the amount of inappropriate use is the so-called ‘RAND appropriateness method,’ which is based on a structured review of the scientific literature and the collective judgment of an expert panel. Measured by this method, a number of procedures have been shown to have high rates of inappropriate or uncertain use. The challenge is to find ways to eliminate ineffective services and procedures to free resources for those that have been proven effective. Further research is needed to improve the method and to find acceptable ways its findings can be used to promote effective care.


The first steps in quality assessment of results of performance in gynecology and obstetrics were taken by the Swiss Association of the OB/GYN Clinics of the Swiss Society of Obstetrics and Gynecology in the 1980's. Currently, medical data of more than 800,000 patients are available, allowing a good comparison of the quality of results. Concerned that purchasing coverage for health services was made exclusively on the basis of cost, resulting in lower public health quality, the Swiss Society of OB/GYN set up a Board of Quality Assessment in 1995. Not only the quality of results, but also the quality of indications, structures and processes are considered by using a modified
Donabedian model. Moreover, standardized patient information forms have been worked out for 15 gynecological and five obstetrical operations. Since hysterectomy is the most common major gynecological operation in Switzerland, the evaluation of the quality of the indications is of substantial interest. Guided by the RAND Appropriateness Method, criteria for quality assessment with respect to appropriateness and necessity of hysterectomy have been drawn up. Swiss guidelines and 125 different clinical scenarios are defined by a panel of 22 OB/GYN experts. The aim of this project is to offer an interactive second opinion system, called "Swiss Guidelines for Hysterectomy", and anonymous self-assessment of quality. Appropriate-ness and necessity of medical procedures have to be defined by medical professionals providing a proper balance between the patient's desires and financial demands.


A.3) **Necessity: general descriptions and basic definitions**


The extent of unnecessary surgery has been the object of considerable speculation and occasional wild accusation in recent years. Most evidence of the existence of unnecessary surgery, such as information from studies of geographic variations and the results of second surgical opinion programs, is circumstantial. However, results from the few studies that have measured unnecessary surgery directly indicate that for some highly controversial operations the fraction that are unwarranted could be as high as 30 percent. Most unnecessary surgery results from physician uncertainty about the effectiveness of an operation. Elimination of this uncertainty requires more efficient production and dissemination of scientific information about clinical effectiveness. In the absence of adequate data from scientific studies, the use of a consensus or expert opinion, disseminated by means of comprehensive practice guidelines, offers the best opportunity to identify and eliminate unnecessary surgery.


This is a report on the extension of the concept of the appropriateness of a procedure to the necessity, or crucial importance, of that procedure. To state that a procedure is crucial means that withholding the procedure would be deleterious to the patient's health. Appropriateness and necessity ratings for six procedures were obtained using a modified Delphi panel process developed in earlier work. Panels were composed of practicing clinicians who were recognized leaders in their fields. The panels included both performers and non-performers of the procedure under discussion. For most procedures and panelists, necessity was related to appropriateness, but was distinct from it. The proportion of indications for which the procedure was crucial varied in clinically consistent ways both among and within procedures. However, panelists did not achieve a consensus on necessity. Further research is suggested to refine the method to promote consensus and to validate further the ratings of necessity. In conclusion, necessity ratings can be used together with appropriateness ratings to address not only the overuse of procedures, but also to indicate limited access to care through underuse of procedures.

A.3.03 **Bergthold, L.A., “Medical Necessity: Do We Need It?”, *Health Affairs*, 1995.**

The article summarizes the history of the term medical necessity as it has been used in public and private plans and the way it has evolved from an insurance concept to a rationing tool. It will describe how and why the term was defined during the recent national health care reform debate, to whom the definitions are important, and propose some clarification of the meaning of medical necessity within a changing delivery system. Do we need the term medical necessity at all?

The term "medical necessity" is used ubiquitously in health care, but its meaning and implementation vary substantially among providers, payers, and patients. This ambiguity has led some to suggest that cost-effectiveness be used as a basis for decision rules. This paper presents an analytical framework that is familiar to clinicians and shows that medical necessity and cost-effectiveness do not provide deterministic rules for clinical decision making. First, 2 x 2 tables are used to show the tradeoff between the sensitivity and specificity of decision rules. Then, the example of asymptomatic abdominal aortic aneurysm is used to show that these tradeoffs can be seen as a continuum of decision rules on a receiver-operating characteristic curve. Society can therefore choose a decision threshold on the basis of medical necessity that optimizes the number of lives saved or any other desired outcome, but the tradeoff between sensitivity and specificity cannot be avoided. Applying cost-effectiveness criteria may change the decision threshold because cost-effectiveness itself involves inherent tradeoffs that create additional ambiguity for clinical decisions. The conclusion is that decision rules based on medical necessity or cost-effectiveness should not be considered deterministic. Rather, decision rules are useful when they make assumptions explicit and specify tradeoffs so that clinicians, patients, and payers can make better decisions.


Background: Attempting to explain the marked variation in utilization of medical procedures has vexed health policy analysts for nearly three decades. Most health services research to date has been directed at identifying and reducing excessive utilization. Little attention has been given to underuse of care. The Los Angeles Cardiac Underuse Project Overview: A research group at the University of California, Los Angeles (UCLA), performed two separate, published studies seeking to measure underuse of coronary angiography and coronary artery revascularization (bypass surgery and angioplasty), two critical links in the chain of care leading from initial diagnosis of coronary artery disease to definitive treatment. In each study, the necessity criteria developed by the panel were used to identify patients needing an invasive procedure. Results: Within this population of patients (sampled predominantly from public hospitals), substantial underuse of clinically necessary coronary angiography (41% without refusers) and revascularization (23% without refusers) was detected. In this select population of patients, receiving a necessary revascularization procedure appeared to both reduce the risk of death and improve quality of life. Discussion: Despite limitations of the method, detection of underuse is feasible, valid, and affordable in the context of overall health care expenditures. Moreover, the case for implementing "underuse prevention" systems is increasingly compelling. Measuring and disseminating data on underuse of expensive but highly beneficial procedures would provide health care consumers (patients and employers) with useful information and enable health care providers to develop quality improvement strategies aimed at rational use of health care resources.

A.4) Methodological issues


Consensus methods are being used increasingly to solve problems in medicine and health. Their main purpose is to define levels of agreement on controversial subjects. Advocates suggest that, when properly employed, consensus strategies can create structured environments in which experts are given the best available information, allowing their solutions to problems to be more justifiable and credible than otherwise. This paper surveys the characteristics of several methods (Delphi, Nominal group, and models developed by the National Institutes of Health and Glaser) and provides guidelines for those who want to use the techniques. Among the concerns these guidelines address are selecting problems, choosing members for consensus panels, specifying acceptable levels of agreement, properly using empirical data, obtaining professional and political support, and disseminating results.


We sought the voluntary cooperation of a randomly selected sample of community physicians and hospitals in five states for a study of how appropriately they performed coronary angiography, carotid endarterectomy, and upper gastrointestinal tract endoscopy. Ninety percent of 913 sampled physicians (n = 819) consented to a review of up to 20 of their 1981 Medicare patients' records. These physicians represented seven different specialties and subspecialties and performed 4988 procedures, 92% of the desired sample. Only three of 230 hospitals did not participate. We attribute our method's success primarily to the formation of a network to connect the branches of the profession, respect for office and hospital practice routine, confidentiality, and the development of carefully designed medical record abstraction systems. We conclude that, with effort, cooperative research among disparate segments of the medical community can become a reality even if the topic studied is relatively sensitive.


Understanding the clinical appropriateness of a procedure's use may be critical in explaining geographic variations in its use. Little is known, however, about whether data on appropriateness can be obtained from a medical record. A national panel of physicians formulated a list of 300 mutually exclusive, detailed clinical indications for performing coronary angiography. Using this list, we compared the reasons physicians perform coronary angiography as revealed in medical records with those given in interviews with the physicians who actually did the procedure. Thirty-five of 47 eligible billing entities (74%) from two Los Angeles Professional Standards Review Organization areas participated. These physicians practiced in 14 hospitals and accounted for 81% of all
angiographies performed on Medicare patients in the two areas. Sixty-six records (approximately two per physician) were reviewed; physician interviews were conducted by two trained data collectors who were blinded to each other's results. Ninety-one percent agreement was reached on the specific indication for performing coronary angiography when information from the record review and interview was compared. We conclude that medical records yield valid information on why coronary angiography is performed and that they are a suitable source to use in judging the appropriateness of that use.


Increasing attention is being paid to data on geographic differences in population-based rates of use of medical and surgical procedures. To understand these differences and to determine what level of use is appropriate, a method is needed to judge the clinical appropriateness of services. We recently developed and tested such a method in two large, urban geographic areas. Eighty-one medical records from a randomly selected sample of 30 billing entities (46 physicians) who per-formed upper gastrointestinal endoscopy (UGIE) on Medicare patients were abstracted. Ninety-four percent of physicians who were asked agreed to participate. Reliability testing showed 99% agreement on items abstracted from a subset of records independently reviewed by two abstrac-tors. Based on the abstractions, each patient could be assigned at least one (mean 2.2) specific clinical indication for which UGIE was performed. Using ratings derived from a previously held panel meeting, it was possible to evaluate the appropriateness of the indications for each UGIE.


A study was undertaken to test the hypothesis that consultants who are willing to participate in expert panels are similar, in terms of routinely available characteristics, to those who are not participating. All consultants in acute specialties in North-east Thames Region were asked to participate in a series of expert panels. Routinely available data was used to compare those who agreed to participate with those who declined or did not reply. Consultants who are willing to participate in expert panels are similar to those who are not in terms of years since qualification, specialty, sex, country of graduation, and possession of higher degrees. Consultants working in district general hospitals seem to be more likely to be willing to participate than those employed in teaching hospitals (37% versus 26%, p less than 0.02), although this difference may be accounted for by errors in the list of teaching hospital consultants.


The RAND-UCLA Health Services Utilization Study previously analyzed the appropriateness of use of carotid endarterectomy based on a literature review and global expert judgments. In this study, for 45 of the same clinical indications used in the RAND-UCLA Study, the authors compare the appropriateness judgments based on the global
judgment method to appropriateness ratings based on probability estimates of specific outcomes that were provided by the same panel of experts. The authors asked these experts to estimate, for each clinical indication, the likelihood of important medical outcomes (i.e. stroke within a year) in the presence and absence of endarterectomy. Using decision analysis, the appropriateness of endarterectomy for these 45 indications was then calculated. For only two of the eight physicians were the Spearman rank-order correlations between these two methods of judging appropriateness significant and positive. (Correlations for the eight physicians ranged from 0.45 to -0.38). This result was produced by: 1) the tendency of the experts to estimate relatively poor outcomes for seriously ill patients whether or not endarterectomy was performed; and 2) a far less consistent effect of clinical factors on outcome estimates than on global judgments. Better methods for incorporating probability estimates into a global rating process must be developed. The authors found excellent agreement between the panelists’ relative outcome estimates for common endarterectomy indications and the observed stroke rate for these same indications, suggesting that one promising method is to use specific outcome data to "anchor" expert judgments.


**Conclusion:** Based on the RAND data, the answer to the above question is decidedly “no”, most of the variation is explained by additional appropriate procedures in high-use areas. High-use and low-use areas have much the same mix of appropriate and inappropriate procedures. The RAND result admittedly raises more questions than it answers. If inappropriate use does not account for variations, what does? Is it disease incidence, patients’ access to care, patients’ propensity to seek care, physicians’ decision making, physician supply, or what? Could it be that ‘appropriate’ is too inclusive a category? Can variations be explained by procedure use that is appropriate but not strictly necessary? To what extent is there ‘underuse’ of ‘necessary’ procedures in low-use areas (or even in high-use areas)? More research is needed.


This report documents the medical record abstraction form and guidelines for appropriateness of hysterectomy used in the HMO Quality of Care Consortium study of this procedure. The abstraction form was designed to follow the format of a medical record so that abstraction would be both accurate and efficient. For each item or group of items, the section of the medical record that was to serve as the source of data was specified. In addition, items derived from a particular portion of the medical record were grouped. To standardize the abstraction process, a detailed set of guidelines was prepared to accompany the abstraction form. The guidelines define medical terms, specify data sources from the medical record, and provide important medical synonyms. A separate form was developed for use by the physician overreader who was responsible for reviewing the data collected on the abstraction form by the HMO abstractor and reviewed by the nurse supervisor at RAND in order to make necessary clinical judgments. As with the medical records abstraction guidelines, the physician overreader guidelines provide item by item instructions for making the required clinical judgments.

Objective: To assess criteria used for detecting underuse of coronary artery revascularization procedures in terms of patient outcomes. Design: Retrospective cohort study using medical records supplemented by a telephone survey and review of county death records. Setting: Four public hospitals and two academically affiliated private hospitals in Los Angeles County, California. Participants: A total of 671 patients who had coronary angiography between June 1, 1990 and September 30, 1991, and who met explicit clinical criteria for the necessity of coronary artery bypass graft (CABG) surgery or percutaneous transluminal coronary angioplasty (PTCA). Main outcome measures: For all patients (n = 671), we estimated the association between receipt of necessary revascularization and mortality (median follow-up after angiography, 797 days) after adjusting for potential confounders. For the patients completing the telephone interview (n = 374), we examined the relationship between receipt of necessary revascularization and frequency of chest pain. Results: Patients who received necessary revascularization within 1 year of angiography had lower mortality than those who did not (8.7% vs 15.8%, P = .01), and this association persisted after adjustment for extent of coronary artery disease, clinical symptom complex, ejection fraction, and cardiac surgical risk index (adjusted odds ratio = 0.49; 95% confidence interval, 0.28 to 0.86). The same general results were obtained whether revascularization was received within 1 year or within 30 days of the catheterization, whether panelists' ratings or individual clinical variables were entered as covariates, and whether the statistical procedure used was logistic regression or Cox proportional hazards analysis. In addition, among patients responding to the telephone survey, those receiving necessary revascularization had less chest pain at
follow-up (P = .03). **Conclusions:** Among patients meeting criteria for the necessity of revascularization, those receiving a revascularization procedure within 1 year had lower mortality than those treated medically. These results support the validity of the RAND/UCLA criteria for detecting underuse of these procedures, but more research is needed to confirm the findings and to determine the validity of guidelines for other procedures.


**Objective:** Spinal manipulation is an efficacious therapy for some patients with low back pain (LBP). In this pilot study, we tested the feasibility of assessing the appropriateness of chiropractic spinal manipulation for patients with LBP. **Methods:** Criteria for the appropriate and inappropriate use of spinal manipulation for low back pain were developed using the RAND/UCLA appropriateness method. Two separate expert panels, one multidisciplinary and one all chiropractic, each rated a comprehensive array of clinical scenarios for appropriateness. A random sample of practicing chiropractors was selected, and data were collected from ten randomly selected office records from each participating clinician. Assessment of the appropriateness for the use of spinal manipulation was made by comparing the care delivered with the appropriateness criteria determined by each expert panel. **Results:** Eight of thirteen (62%) eligible chiropractors agreed to participate. For the remainder, by the multidisciplinary panel's criteria, 38% of care was appropriate and 26% of care was inappropriate. By the all-chiropractic panel's criteria, the same cases were judged 74% appropriate and 7% inappropriate. The two panel's appropriateness ratings were in agreement on 48% of all cases. **Conclusions:** In this geographic area, the rate of appropriate care is between 38% and 74% and the rate of inappropriate care is between 7% and 19%, depending on the criteria used to assess appropriateness. Data from other geographic areas of the U.S. will be needed before inferences to a larger population may be drawn, and we have demonstrated that such a study is feasible.


The analysis of variation in the use of health care services, and particularly of practice variation, has been the subject of renewed interest because of the view that the inappropriate use of procedures could be a major cause of these differences. In this article, recent literature is reviewed and the results of personal research are described on both the variation in care provision and on appropriate assessment. In the few studies that have focussed on both subjects no evidence has yet been found to suggest that practice variation is to be explained by differences in appropriateness rates. However, there are still many methodological pitfalls in both variation analyses (statistical problems) and appropriateness assessment (reliability of the judgement), implying that this conclusion is far from definitive. More research should therefore be conducted on methodological questions of variation analysis and appropriateness assessment. Furthermore in variation analysis the relative contribution of all potential determinants has to be studied on the various levels of care provision. Finally, to study the relationship between practice
variation and appropriateness of care, the clinical problem and not the procedure should be the starting point.


**Objective:** To assess the feasibility of the appropriateness method in the Agency for Health Care Policy and Research (AHCPR) Clinical Practice Guideline Development process, and to compare the results of the appropriateness method with those obtained using evidence tables and an informal consensus method. **Setting:** AHCPR Low Back Problems Clinical Practice Guideline. **Design:** Two different group process methods with the same panel of experts were used in observational comparison of results of and satisfaction with guideline development. **Data collection:** Practice guideline statements were created for topics using the conventional AHCPR method; then six months later new guideline statements for four topics were created using the appropriateness method. Panelist satisfaction with each process and resulting set of guideline statements was assessed. **Principal findings:** Results of the appropriateness method for TENS, discography, and traction showed no disagreement among panel members and no appropriate indications for their use in the patient scenarios considered. These results are qualitatively similar to the guideline statements produced using evidence tables and informal consensus. Clinical practice guideline statements about electro-diagnostics created from appropriateness ratings were much more clinically specific than those created using evidence tables and informal consensus. Neither informal consensus building nor the appropriateness method was clearly preferred by a majority of panelists. **Conclusions:** It is feasible to use the appropriateness method in the AHCPR Clinical Practice Guideline Development process, and in some instances it produces more clinically specific guideline statements than does informal consensus.


**Objective:** To assess the clinical consistency of expert panelists’ ratings of appropriateness for coronary artery bypass surgery. **Design:** Quantitative analysis of panelists’ ratings. **Participants:** Nine participants (three cardiothoracic surgeons, four cardiologists, and two internists) convened by RAND to establish criteria for the appropriateness of coronary artery bypass surgery. **Main outcome measures:** Percentage of indication-pairs given clinically inconsistent ratings (i.e. higher rating assigned to one member of an indication-pair when rating should have been equal or lower). **Results:** In the final round of appropriateness ratings, among 1785 pairs of indications differing only on a single clinical factor (e.g., three-vessel vs. two-vessel stenosis), 6.6% were assigned clinically inconsistent ratings by individual panelists, but only 2.7% received inconsistent ratings from the panel as a whole (using the median panel rating as the criterion). Internists provided fewer inconsistent ratings (4.6%) than either cardiologists (7.8%) or cardiothoracic surgeons (6.3%) (p < 0.001). More inconsistencies were noted when the factor distinguishing otherwise identical indications was symptom severity (inconsistency rate, 13.2%) or intensity of medical therapy (13.2%) than when it was number of stenosed vessels (3.8%) or proximal left anterior descending (PLAD) involvement (1.9%). Contrary to expectations, panelists’ inconsistency rates increased between the
initial and final rounds of appropriateness ratings (from 3.9% to 6.6%, p < 0.001). Panelists’ mean ratings across indications were only weakly correlated with individual inconsistency rated (r = 0.18, p = NS). **Conclusions:** The RAND/UCLA method for assessing the appropriateness of coronary revascularization generally produces criteria that are clinically consistent. However, research is needed to understand the sources of panelists’ inconsistencies and to reduce inconsistency rates further.


**Background:** Expert panels are frequently used to make recommendations for the use of health services, but there exists concern about the validity of such recommendations. We took advantage of a natural experiment to assess the predictive validity of an expert panel’s ratings of the appropriateness of carotid endarterectomy. **Methods:** We compared the results of randomized controlled trials of carotid endarterectomy with ratings of appropriateness of potential indications for carotid endarterectomy were created by a multidisciplinary group of physicians in 1984. Seven major randomized controlled trials of carotid endarterectomy have since been published. We assessed the concordance between the clinical trials’ results and the classification of the appropriateness of performing carotid endarterectomy on a sample of 1981 Medicare patients. **Results:** The seven clinical trials have provided direct confirmatory evidence for 12 appropriate indications, six uncertain indications, and nine inappropriate indications. These 27 indications accounted for 23 percent of all carotid endarterectomy procedures delivered to the Medicare population in 1981. The seven trials provided inferred confirmatory evidence for an additional two appropriate and 15 inappropriate indications, which accounted for an additional 5 percent of 1981 cases. In no cases have subsequent clinical trials provided evidence refuting the appropriateness ratings. **Conclusions:** The predictive validity of this expert panel’s ratings of the appropriateness of performing carotid endarterectomy has been very good in the instances where there are enough data to reach a conclusion.


**Background:** To assess the overuse and underuse of medical procedures, various methods have been developed, but their reproducibility has not been evaluated. This study estimates the reproducibility of one commonly used method. **Methods:** We performed a parallel, three-way replication of the RAND-University of California at Los Angeles appropriateness method as applied to two medical procedures, coronary revascularization and hysterectomy. Three nine-member multidisciplinary panels of experts were composed for each procedure by stratified random sampling from a list of experts nominated by the relevant specialty societies. Each panel independently rated the same set of clinical scenarios in terms of the appropriateness of the relevant procedure on a risk-benefit scale ranging from 1 to 9. Final ratings were used to classify the procedure in each scenario as necessary or not necessary (to evaluate underuse) and inappropriate or not inappropriate (to evaluate overuse). Reproducibility was measured by overall agreement and by the kappa statistic. The criteria for underuse and overuse derived from
these ratings were then applied to real populations of patients who had undergone coronary revascularization or hysterectomy. **Results:** The rates of agreement among the three coronary-revascularization panels were 95, 94, and 96 percent for inappropriate-use scenarios and 93, 92, and 92 percent for necessary-use scenarios. Agreement among the three hysterectomy panels was 88, 70, and 74 percent for inappropriate-use scenarios. Scenarios involving necessary use of hysterectomy were not assessed. The three-way kappa statistic to detect overuse was 0.52 for coronary revascularization and 0.51 for hysterectomy. The three-way kappa statistic to detect underuse of coronary revascularization was 0.83. Application of individual panels' criteria to real populations of patients resulted in a 100 percent variation in the proportion of cases classified as inappropriate and a 20 percent variation in the proportion of cases classified as necessary. **Conclusions:** The appropriateness method is far from perfect. Appropriateness criteria may be useful in comparing levels of appropriate procedures among populations but should not by themselves be used to direct care for individual patients.


**Objective:** Examine the precision and completeness of information in medical records (MR) for evaluating the appropriateness of operative indications for lumbar disc surgery. **Design:** Retrospective review of MR. **Setting:** University Hospital Department of neurosurgery. **Subjects:** MR of 100 patients. **Interventions:** None. **Main Outcome Measure:** Percent of pre-defined, detailed appropriateness criteria present in the MR. Percent of cases capable of being unequivocally classified as to the appropriateness of the operative indication. **Results:** The criteria were present and precise for 52% of the items (range 9 - 90%); present but imprecise for 38% of the items and absent for 10%. Due to this imprecision, the appropriateness of only 7% of the operative indications could be unequivocally determined retrospectively. **Conclusion:** Medical records are of limited use in assessing the appropriate management of care. The evaluation of the process of care should therefore be done prospectively.


**Background:** Variation in expert opinion and lack of a systematic methodology hinder the development of reliable clinical practice guidelines. However standardized protocols have been defined to quantify, combine and summarize expert judgments. In addition, statistical methods may help to outline guidelines based on simplified models of these judgments. **Methods:** To test this hypothesis, stepwise logistic regression (SLR) and classification tree pruning (CTP) were used to predict the results of two expert panels(USA, 1992/ Switzerland, 1995) on laminectomy in sciatica conditions. Both panels, using the RAND-UCLA explicit method, assessed whether the procedure would be inappropriate or of potential use in 720 case scenarios combining 7 relevant factors. **Results:** Laminectomy was rated as inappropriate in 60%/70% of the scenarios by the
US/Swiss panels respectively. Either statistical method, in both panels, based its simplest model on the same 4 factors: imaging test results, disability, neurological findings and conservative treatment trials (in decreasing order), the influence of 2 other factors - duration of pain and nerve root irritation - being only marginal. The correct classification rates of the models were 89%/93% for SLR and 93%/85% for CTP. Adopting the CTP US algorithm as a guideline would lead to consider performing laminectomy only in patients with imaging evidence of hernia, relatively severe disability, reflex abnormalities and previous nonsurgical treatment. Adherence to the corresponding CTP Swiss algorithm would result in less restrictive conditions. **Conclusion:** The statistical techniques proved useful instruments to structure and simplify appropriateness criteria developed by expert panels and to outline parsimonious decision models for clinical practice.
A.5) Supporting software


Because of the complexities of medical care, results from randomised control trials, even when available cannot be easily interpreted to produce clear-cut recommendations. Studies looking at quality of care have in fact documented that some patients are not receiving procedures that they need, and others are receiving procedures that they do not need. Researchers at the RAND Corporation and UCLA (USA) have developed a structured methodology to produce explicit indications using multidisciplinary panel of experts. They hypothesized that producing explicit guidelines for care would help to improve patients’ outcome by addressing both the “overuse” (i.e. people that do not need a procedures get it) of procedures/interventions. To perform the analysis of the indications produced through RAND/UCLA method we have developed two SAS® macro routines %TREE and %AGREE, using macro language that makes the program available for any kind of data, already presented at SEUGI’95. Clinical practice guidelines can be identified, verified and simplified by the use of logical analysis and application of decision table techniques. To summarize the clinical practice guidelines produced by the two macros mentioned above we have developed a third new macro, %PATTERN. The results of the method will be illustrated using the example of indications for breast cancer surgery developed by an Italian panel of doctors from different specialties indications.

A.5.02 Clivio, L., Scedan (Scenarios Developer and Analyzer), November 1997.

This program lets an easy implementation of definable consensus methods for medical and health services research. Two consensus methods commonly adopted are based on ‘Delphi’ and ‘Expert panel’ techniques. With this program it is possible to define a number of scenarios, identified upon the definition of three entities, called variables, levels and uncompatibilities (uncompatible levels group). The scenarios can be rated by a number of definable ‘experts’ (panelists) about a specific procedure. The procedures have to be defined with a set of parameters that specify the selected analysis method. The selected experts can give more than one rate for each scenario, in different customizable rounds. For each round, it is possible to see some statistics about the last rating before making up one’s mind to rate.
B) RAM outcomes (literature reviews, panel ratings or chart audits by topic)

B.1) Assessment of appropriateness of care (possibly with B.2 or B.3)


**Topic(s): coronary angiography**

Using ratings of appropriateness derived from an expert physician panel, we measured how appropriately physicians in 1981 performed coronary angiography in a randomly selected, community-based sample of cases in the Medicare population. We studied large geographic areas (three sites) in three states, representing regions of high and low use. The high-use site had fewer procedures classified as appropriate (72%) than either low-use site (77% and 81%, respectively). Over all sites, 17% of procedures were classified as inappropriate. Patients in the high-use site were older, had less severe angina, and were less intensively medically treated than patients in either of the low-use sites. Patients without angina who had not undergone exercise testing constituted the most common subgroup of inappropriate cases. Although overall differences in appropriateness were not large, practice differences do exist. This analysis of practice differences among study sites provides the clinical basis for understanding the small, but significant, differences in the appropriateness of use of coronary angiography. The finding of 17% inappropriate use may be cause for concern.


We studied the appropriateness of use of coronary angiography, carotid endarterectomy, and upper gastrointestinal tract endoscopy and its relationship to geographic variations in the rates of use of these procedures. We selected geographic areas of high, average, and low use of these procedures and randomly sampled Medicare beneficiaries who had received one of the procedures in 1981. We determined the indications for the procedures using a detailed review of medical records and used previously developed ratings of appropriateness to assign an appropriateness score to each case. Differences among sites in levels of appropriateness were small. For example, in the high-use site for coronary angiography, 72% of the procedures were appropriate, compared with 81% in the low-use site. Coronary angiography was performed 2.3 times as frequently in the high-use site compared with the low-use site. Under the conditions of this study, we did find significantly levels of inappropriate use: 17% of cases for coronary angiography, 32% for carotid endarterectomy, and 17% for upper gastrointestinal tract endoscopy. We conclude that differences in appropriateness cannot explain geographic variations in the use of these procedures.

Topic(s): upper gastrointestinal endoscopy

Study objective: To determine how appropriately physicians in 1981 did upper gastrointestinal endoscopy in a randomly selected, community-based sample of Medicare patients. Design: We developed a comprehensive and clinically detailed list of 1069 indications for upper gastrointestinal endoscopy. A national panel of nine clinicians rated the appropriateness of the indications. We categorized the indications as appropriate, inappropriate, or equivocal. We did a clinically detailed medical record review of a random sample of 1585 patients having upper gastrointestinal endoscopy to assess the appropriateness of using upper gastrointestinal endoscopy. Setting: Patients were sampled from large geographic areas in three states. Two areas represented high use, and one area, low use. Patients: Random sample of patients 65 years of age or older receiving diagnostic upper gastrointestinal endoscopy. Interventions: None; the study was retrospective. Measurement and results: Patient characteristics, histories, and clinical indications for upper gastrointestinal endoscopy were similar across low- and high-use areas. Overall, 72% of the endoscopies were done for appropriate indications, 11% for equivocal indications, and 17% for inappropriate indications. Upper gastrointestinal bleeding (26%), follow-up to an abnormal upper gastrointestinal series (21%), dysphagia (18%), and dyspepsia (15%) were the most frequent clinical reasons for doing endoscopy. Inpatient endoscopies were more often appropriate and less often inappropriate than outpatient endoscopies. Conclusions: This analysis of practice patterns among study sites provides the clinical basis for understanding the use of upper gastrointestinal endoscopy. The finding of 17% inappropriate use may be cause for concern.


Topic(s): coronary artery bypass surgery

Information about how appropriately procedures are performed is vital to the understanding of the impact of technology and to the success of efforts to channel its use appropriately. While the efficacy of coronary artery bypass surgery has been addressed in several large-scale, randomized trials, there is little information about how appropriately the procedure is actually being used in the community. We determined the appropriateness of coronary artery bypass surgeries performed in three randomly chosen hospitals in a western state. We determined appropriateness by comparing data obtained from a detailed medical record review with a list of 488 indications. This list, developed by a national panel of physicians, covered all possible reasons for performing the procedure. Three hundred eighty-six cases from the years 1979, 1980, and 1982 were examined. Fifty-six percent of the surgeries were performed for appropriate reasons, 30% for equivocal reasons, and 14% for inappropriate reasons. The percentage of appropriate surgeries varied by hospital, from 37% to 78%, but did not vary by patient age. Eliminating the performance of inappropriate procedures may lead to reductions in health care expenditures or to improved patient outcomes.
Carotid endarterectomy is a commonly performed but controversial procedure. We developed from the literature a list of 864 possible reasons for performing carotid endarterectomy, and asked a panel of nationally known experts to rate the appropriateness of each indication using a modified Delphi technique. On the basis of the panel's ratings, we determined the appropriateness of carotid endarterectomy in a random sample of 1302 Medicare patients in three geographic areas who had had the procedure in 1981. Thirty-five percent of the patients in our sample had carotid endarterectomy for appropriate reasons, 32 percent for equivocal reasons, and 32 percent for inappropriate reasons. Of the patients having inappropriate surgery, 48 percent had less than 50 percent stenosis of the carotid artery that was operated on. Fifty-four percent of all the procedures were performed in patients without transient ischemic attacks in the carotid distribution. Of these procedures, 18 percent were judged appropriate, as compared with 55 percent judged appropriate in patients with transient ischemic attacks in the carotid distribution. After carotid endarterectomy, 9.8 percent of patients had a major complication (stroke with residual deficit at the time of hospital discharge or death within 30 days of surgery). We conclude that carotid endarterectomy was substantially overused in the three geographic areas we studied. Furthermore, in situations in which the complication rate is equal to or above the study's aggregate rate, carotid endarterectomy would not be warranted, even in cases with an appropriate indication, because the risks would almost certainly outweigh the benefits.

We previously reported substantial disagreement among expert physician panelists about the appropriateness of performing six medical and surgical procedures for a large number of theoretical indications. A recently completed community-based medical records study of about 4,500 patients who had one of three procedures—coronary angiography, upper gastrointestinal endoscopy, and carotid endarterectomy—shows that many of the theoretical indications are seldom or never used in practice. However, we find that there is also substantial disagreement (5, 25, or 32 per cent for angiography, endoscopy, or endarterectomy, respectively) about the appropriateness of indications used in actual cases if disagreement is defined by first discarding the two extreme of nine ratings, then looking for at least one rating near the bottom (1 to 3) and one near the top (7 to 9) of the 9-point scale. Patients should know that a substantial percentage of procedures are performed for indications about which expert physicians disagree.

**Background and methods:** In a nationally representative population 65 years of age or older, we have demonstrated that about one quarter of coronary angiographies and upper gastrointestinal endoscopies and two thirds of carotid endarterectomies were performed for reasons that were less than medically appropriate. In this paper we examine whether specific characteristics of patients (age, sex, and race), physicians (age, board-certification status, and experience with the procedure), or hospitals (teaching status, profit-making status, and size) predict whether a procedure will be performed appropriately. **Results:** In general, we found that little of the variability in the appropriateness of care (4 percent or less) could be explained on the basis of standard, easily obtainable data about the patient, the physician, or the hospital. For all three procedures, however, performance in a teaching hospital increased the likelihood that the reasons would be medically appropriate (P = 0.09 for angiography, p = 0.30 for endoscopy, and P < 0.01 for endarterectomy). In addition, angiographies were more often performed for appropriate reasons in older or more affluent patients (P < 0.01 for both). Being treated by a surgeon who performed a high rather than a low number of procedures decreased the likelihood of an appropriate endarterectomy by one third, from 40 to 28 percent (P < 0.01). **Conclusions:** Appropriateness of care cannot be closely predicted from many easily determined characteristics of patients, physicians, or hospitals. Thus, for the present, if appropriateness is to be improved it will have to be assessed directly at the level of each patient, hospital and physician.


**Topic(s):** carotid endarterectomy, upper gastrointestinal endoscopy, coronary angiography

**Background and methods:** In a nationally representative population 65 years of age or older, we have demonstrated that about one quarter of coronary angiographies and upper gastrointestinal endoscopies and two thirds of carotid endarterectomies were performed for reasons that were less than medically appropriate. In this paper we examine whether specific characteristics of patients (age, sex, and race), physicians (age, board-certification status, and experience with the procedure), or hospitals (teaching status, profit-making status, and size) predict whether a procedure will be performed appropriately. **Results:** In general, we found that little of the variability in the appropriateness of care (4 percent or less) could be explained on the basis of standard, easily obtainable data about the patient, the physician, or the hospital. For all three procedures, however, performance in a teaching hospital increased the likelihood that the reasons would be medically appropriate (P = 0.09 for angiography, P = 0.30 for endoscopy, and P < 0.01 for endarterectomy). In addition, angiographies were more often performed for appropriate reasons in older or more affluent patients (P less than 0.01 for both). Being treated by a surgeon who performed a high rather than a low number of procedures decreased the likelihood of an appropriate endarterectomy by one third, from 40 to 28 percent (P less than 0.01). **Conclusions:** Appropriateness of care
cannot be closely predicted from many easily determined characteristics of patients, physicians, or hospitals. Thus, for the present, if appropriateness is to be improved it will have to be assessed directly at the level of each patient, hospital, and physician.


Topic(s): cholecystectomy

Background: Previous studies have reported variation in the population-based use rate of diagnostic and therapeutic procedures. Cholecystectomy is one of the most common surgical procedures, and we conducted this study to assess whether in Israel the use of this procedure varied by region and whether differences in use can be related to differences in appropriateness of use. In Israel, there is a pre-paid health insurance system and all surgeons are salaried. Methods: Age-adjusted rates of cholecystectomy in four hospitals, each serving a defined population in Israel, were calculated. Two hundred and sixty-six potential clinical indications for performing cholecystectomy were rated as to their appropriateness by a panel of 9 expert physicians. A trained team abstracted the medical records of all patients who underwent the operation in the four Israeli hospitals in 1986 (n = 702) and recorded the clinical indication for the surgery. Results: The population-based age-adjusted rates of cholecystectomy varied over threefold among the four hospitals. 29% of the cholecystectomies were performed for less than appropriate reasons, and this figure varied by hospital from 36% to 17% (p = 0.002). However, appropriateness did not vary systematically with the population-based use rate. Conclusion: Cholecystectomy was performed frequently for inappropriate or equivocal reasons, even in a country in which resources are limited, and physicians are salaried. Efforts to improve surgical decision making should be undertaken.


Topic(s): hysterectomy

Objective: To correlate the listing of multiple preoperative indications for hysterectomy with the risk of non-confirmation of the preoperative diagnosis. Methods: Records of 171 women undergoing consecutive hysterectomies for all indications at a large teaching hospital were reviewed for preoperative indication(s), compliance with published preoperative validation criteria for cases in which tissue pathology was not expected, and histologic verification of the preoperative diagnosis for cases in which tissue pathology was expected. Rates of confirmation (histologic verification plus successful compliance with validation criteria) of the preoperative diagnosis were compared between subgroups of cases in which single indications were listed (N = 124) or multiple indications were listed (N = 47) preoperatively. Results: The rate of confirmation of single indications (115 of 124 cases, 93%) was significantly higher than the rate of confirmation of even one indication in cases in which multiple indications were listed (28 of 47 cases, 60%, P < .0001; relative risk for non-confirmation of multiple indications = 1.55). Multiple indications were more likely to be listed when tissue pathology was not expected, representing 49% of validatable indications as compared with only 18% of histologically verifiable indications (P < .0001). Overall, the rate of compliance with validation criteria
(70%) was significantly lower than the rate of histologic verification (90%) (P < .01).

**Conclusion:** These data suggest that listing of multiple preoperative indications for hysterectomy is associated with both decreased appropriateness, as reflected in decreased compliance with generally accepted preoperative validation criteria, and decreased diagnostic accuracy, as reflected in lower rates of histologic verification.


**Topic(s):** hysterectomy

**Objective:** To develop and test a method for comparing the appropriateness of hysterectomy use in different health plans. **Design:** Retrospective cohort study. **Setting:** Seven managed care organizations. **Patients:** Random sample of all non-emergency, non-oncological hysterectomies performed in the seven managed care organizations over a 1-year period. Patients who were not continuously enrolled in a plan for 2 years prior to their hysterectomy were excluded. **Main outcome measures:** Proportion of women undergoing hysterectomy in each plan for inappropriate clinical reasons according to ratings derived from a panel of managed care physicians. **Results:** Overall, about 16% of women underwent hysterectomy for reasons judged to be clinically inappropriate. Only one plan had significantly more hysterectomies rated inappropriate compared with the group mean (27%, unadjusted). Adjusting for age and race did not affect the rankings of the plans and had little effect on the numeric results. **Conclusion:** The rates of inappropriate use of hysterectomies are similar to those for other procedures and vary to a small degree among health plans. This information may be useful to purchasers when they consider which health plans to offer their employees.


**Topic(s):** coronary angiography, coronary artery bypass surgery

To determine whether patients are less likely to receive an inappropriate procedure in countries that devote fewer resources to health care than does the United States, we studied how appropriately coronary angiography and coronary artery bypass surgery were performed in the Trent region of the United Kingdom. The medical records of 320 patients who underwent coronary angiography and 319 who underwent coronary artery bypass surgery in 1987 and 1988 were randomly selected for review. Despite the United Kingdom's more limited use of coronary angiography and coronary artery bypass surgery, a substantial proportion were still performed for less than appropriate reasons, by both U.S. and U.K. criteria. Merely reducing the rate of use of these procedures will not be sufficient to eliminate such inappropriate use.

Topic(s): coronary angiography

Objective: To determine the appropriateness of use of coronary angiography in New York State. Design: Retrospective randomized medical record review. Setting: Fifteen randomly selected hospitals in New York State that provide coronary angiography. Patients: Random sample of 1335 patients undergoing coronary angiography in New York State in 1990. Main outcome measures: Percentage of patients who underwent coronary angiography for appropriate, uncertain, or inappropriate indications. Results: Approximately 76% of coronary angiographies were rated appropriate; 20%, uncertain; and 4%, inappropriate. Inappropriate use did not vary significantly between the elderly (i.e., patients aged 65 years and older) and non-elderly, 4.7% and 3.9%, respectively. Although the rate of inappropriate use varied from 0% to 9% among hospitals, the difference was not significant. Rates of appropriateness did not vary by hospital location (upstate vs. downstate), volume (fewer than 750 procedures annually or at least 750 procedures annually), teaching status, or whether revascularization was available at the hospital where angiography was performed. Conclusions: Although coronary angiography was used for few inappropriate indications in New York State, many procedures were performed for uncertain indications in which the benefit and risk were approximately equal or unknown.


Topic(s): coronary angiography

In recent years there has been a steady increase in number of coronary angiography procedures and coronary angioplasty. Since these interventions are relatively expensive, we tried to evaluate the appropriateness of coronary angiography indications. A retrospective pilot study was undertaken on all residents in Padua who had had a coronary angiography performed in 1988. In order to take into account patients who, although resident in Padua, had a coronary angiography performed outside Padua or even abroad, all Italian Hemodynamic Centers and the Office for Foreign Cures Authorization of ULSS 21 were contacted. Follow-up was stopped on April, 1991. Patients were first checked if living by consulting the Register's Office of the Community of Padua; they were then interviewed by letter on state of health. One hundred twenty-four patients underwent 138 coronary angiography. Utilization rate is 5.6 per 10,000 people (CI 4.6-6.6). This figure is lower than the present USA utilization rate, and is similar to the rate of late 70's. From patient records, 3 out 8 groups emerged as predominant indication for coronary angiography: unstable angina (37%), valvular heart disease (20%) and recent myocardial infarction (20%). According the ACC/AHA guidelines, indication was considered "appropriate" in 69%, "inappropriate" in 7%, "doubtful" in 20% and impossible to evaluate in 5% of cases. Although this classification may have been built up with incomplete information, it is note-worthy that the percentage of inappropriate indication is comparable to that of other reports. A further observation is that not in every
case the treatment assigned at the time of diagnostic angiography was really carried out on the patients.


**Topic(s):** percutaneous transluminal coronary angioplasty

**Objective:** To determine the appropriateness of use of percutaneous transluminal coronary angioplasty (PTCA) in New York State. **Design:** Retrospective randomized medical record. **Setting:** Fifteen randomly selected hospitals in New York State that provide PTCA. **Patients:** Random sample of 1306 patients undergoing PTCA in New York State in 1990. **Main outcome measures:** Percentage of patients who underwent PTCA for indications rated appropriate, uncertain, and inappropriate. **Results:** The majority of patients received PTCA for chronic stable angina, unstable angina, and in the post-myocardial infarction period (up to 3 weeks). Fifty-eight percent of PTCAs were rated appropriate; 38%, uncertain; and 4%, inappropriate. The inappropriate rate varied by hospital from 1% to 9% (P = .12); the uncertain rate, from 26% to 50% (P = .02); and the combined inappropriate and uncertain rate, from 29% to 57% (P < .001). There was no difference in appropriateness when the institutions were grouped by volume (fewer than 300 procedures annually or at least 300 procedures annually), location (upstate vs. downstate), or by teaching status. **Conclusions:** Few PTCAs were performed for inappropriate indications in New York State. However, the large number of procedures performed for indications that were rated uncertain as to their net benefit requires further study and justification at both clinical and policy levels.


**Topic(s):** coronary artery bypass graft surgery

**Objective:** To determine the appropriateness of use of coronary artery bypass graft surgery in New York State. **Design:** Retrospective randomized medical record review. **Setting:** Fifteen randomly selected hospitals in New York State that provide coronary artery bypass graft surgery. **Patients:** Random sample of 1338 patients undergoing isolated coronary artery bypass graft surgery in New York State in 1990. **Main outcome measures:** Percentage of patients who had bypass surgery for appropriate, inappropriate, or uncertain indications; operative (30-day) mortality; and complications. **Results:** Nearly 91% of the bypass operations were rated appropriate; 7%, uncertain; and 2.4%, inappropriate. This low inappropriate rate differs substantially from the 14% rate found in a previous study of patients operated on in 1979, 1980, and 1982. The difference in rates was not due to more lenient criteria but to changes in practice, the most important being that the fraction of patients receiving coronary artery bypass grafts for one- and two-vessel disease fell from 51% to 24%. Individual hospital rates of inappropriateness (0% to 5%) did not vary significantly. Rates of appropriateness also did not vary by hospital location, volume, or teaching status. Operative mortality was 2.0%; 17% of patients suffered a complication. Complication rates varied significantly among hospitals (P <
.01) and were higher in downstate hospitals. **Conclusions:** The rates of inappropriate and uncertain use of coronary artery bypass graft surgery in New York State were very low. Rates of inappropriate use did not vary significantly among hospitals, or according to region, volume of bypass operations performed, or teaching status.


**Topic(s):** coronary angiography, coronary revascularization

**Objective:** To evaluate the appropriateness of performing coronary angiography and revascularization in a Swedish population. **Design:** Prospective population study of questionnaires and medical records. **Setting:** All the hospitals in southwestern Sweden that perform coronary angiography and revascularization. **Patients:** Random sample of 831 patients (with chronic stable angina) on the waiting list for coronary angiography or revascularization in southwestern Sweden in September 1990. **Main outcome measure:** Percentage of patients referred for coronary angiography or revascularization for appropriate, uncertain, or inappropriate indications. **Results:** Of the patients referred for angiography, 89% were classified as appropriate, 9% as uncertain, and 2% as inappropriate. The percentages are similar for patients referred for coronary artery bypass graft surgery and for angioplasty (91% and 86%, respectively, classified as appropriate). The majority of patients had chest pain rated as Canadian Cardiovascular Society classes II through IV (93%), despite maximum anti-ischemic therapy in 90% of these patients. **Conclusions:** Few patients were referred for coronary angiography or revascularization for inappropriate or uncertain indications. The percentage of these patients who are from southwestern Sweden is similar to the percentage recently reported from New York State.


**Topic(s):** coronary angiography, percutaneous transluminal coronary angioplasty, coronary artery bypass graft surgery

**Objective:** To determine whether there are differences between women and men in the appropriateness of use of cardiovascular procedures. **Design:** Retrospective chart review. **Setting:** Thirty hospitals located in New York State. **Patients:** Random sample of 3979 patients undergoing coronary angiography, percutaneous transluminal coronary angioplasty, or coronary artery bypass graft surgery in 1990. **Measures:** We evaluated two measures: (1) the percent of women and men who underwent cardiovascular procedures for appropriate, uncertain, and inappropriate indications and (2) for coronary angiography patients, the prognostic exercise stress treadmill score that predicts before the coronary angiogram the 5-year probability of death from a cardiovascular event. **Results:** The inappropriate rate of use of cardiovascular procedures was low and not significantly different for men and women (4% vs. 5% for coronary angiography; 4% vs. 3% for percutaneous transluminal coronary angioplasty; and 2% vs. 3% for coronary artery bypass graft surgery, respectively), and the use of these procedures for uncertain reasons also did not vary significantly by gender. There was also no significant gender difference in the predicted risk of death from a cardiovascular event for coronary
angiography patients: 24% of men and 22% of women were at high risk (i.e., < 75% 5-year survival rate) and 20% and 16%, respectively, were at low risk (i.e., > or = 95% 5-year survival rate). **Conclusion:** Based on two indicators, the RAND appropriateness score and the Duke prognostic exercise treadmill score, we were unable to find any evidence of a difference in the clinical appropriateness of use of these three cardiovascular procedures between women and men.


**Topic(s):** coronary artery bypass graft surgery, percutaneous transluminal coronary angioplasty


**Topic(s):** tympanostomy tubes

**Objective:** To describe the clinical reasons tympanostomy tubes are proposed for children and to assess their appropriateness. **Design:** Analysis of data previously collected prospectively by a national utilization review (UR) firm during a two-step UR process to assess the medical appropriateness of tympanostomy tube placement. Nurses interviewed otolaryngologists’ and primary care physicians’ office staff to collect clinical data. For a randomly selected subsample of cases found inappropriate, we reviewed subsequent interviews of the otolaryngologists by physician reviewers, who looked for possible extenuating clinical circumstances or additional clinical data that might have changed the appropriateness category. **Setting:** Otolaryngologists’ practices from 49 states and the District of Columbia. **Patients:** All 6611 children younger than 16 years who were insured by clients of the UR firm and whose proposal to receive tympanostomy tubes was reviewed by this system from January 1, 1990, through July 31, 1991. The insurance companies in the study insured 5.6 million Americans at the time of the study. **Main outcome measure:** The medical appropriateness of tympanostomy tube surgery according to explicit criteria developed by an expert panel using the RAND/University of California-Los Angeles modified Delphi method. **Results:** A total of 6429 (97%) of the cases were proposed for recurrent acute otitis media, otitis media with effusion, or both. Making generous clinical assumptions, 41% of the proposals for these reasons had appropriate indications, 32% had equivocal indications, and 27% had inappropriate ones. Considering the additional information available from the subsample review, the proportion appropriate was 42%, equivocal 35%, and inappropriate 23%. **Conclusion:** About one quarter of tympanostomy tube insertions for children in this study were proposed for inappropriate indications and another third for equivocal ones.

**Topic(s):** coronary angiography

**Objective:** To assess the appropriateness of using coronary angiography in two major medical centers in Israel. **Methods:** (a) Development of dichotomous appropriateness rating by consensus of an expert panel for possible indications to use coronary angiography; (b) applying the panel's rating to clinical data gathered from interviews and records of 499 patients who underwent coronary angiography. **Results:** Angiographies were performed inappropriately in 58% of patients: in 56% neither prognostic stratification was performed, nor was comprehensive medical therapy prior to angiography undertaken. Referral to catheterization from an acute care hospital, compared to ambulatory referral, correlated independently with an inappropriate angiogram. **Conclusion:** In the Israeli public health system there is still a high rate of inappropriate coronary angiography, most of the cases being due to error in management before performing angiographies.


**Topic(s):** upper gastrointestinal endoscopy

Work by this group has shown that there is a wide range of opinion as to patients' suitability for endoscopy. In a recent study, 1297 questionnaires were sent to a random selection of doctors, including 350 general physicians, 400 surgeons, 477 gastroenterologists, and 70 general practitioners. The respondent was asked to indicate whether or not he would refer the patient described by each case vignette for endoscopy. Depending on the indication, the positive referral rate varied from 4.5% to 99% overall, and from 4.5% to 63.8% for all those clinical situations that the working party felt to be inappropriate. A second study examined the appropriateness of 400 consecutive cases referred from four units within one health region; these cases were judged independently, and without conferring, by a panel of seven gastroenterologists. The same cases were rated by software that incorporated American opinion (the RAND criteria). Although only 45 (11%) of the cases were classed as inappropriate by the British panel, 120 cases (31%) assessed by the American software were rated inappropriate. These differences occurred largely because in the USA it is recommended that one month's antiulcer treatment be tried before considering endoscopy for dyspepsia and thus many referrals were seen as inappropriate by the American database. Of the 45 cases found to be inappropriate by the British doctors no important abnormality was found at endoscopy; whereas of 120 cases judged inappropriate by the RAND criteria, three duodenal and two gastric ulcers, and one gastric cancer were diagnosed at gastroscopy. This study attempts a quantitative assessment of inappropriate use and serves to encourage further work to define appropriateness.

Topic(s): hysterectomy

The degree to which national expert panel survey ratings of the appropriateness of hysterectomy differed from those of a random sample of practicing community gynecologists was determined. Community gynecologists rated hysterectomy as more appropriate on six of eight cervical dysplasia scenarios. Experts agreed among themselves on 19 of 32 indications (intraclass correlation coefficient = 0.66); community gynecologists agreed on 12 of 32 indications (intraclass correlation coefficient = 0.50). Although few differences of opinion existed between experts and community gynecologists, for common clinical scenarios there was a large variation of opinion about the appropriateness of hysterectomy within each group. For areas of clinical uncertainty in which experts’ opinions are used in guideline development, additional measures such as process of care, quality of life, and patient preference should be included in discussions about guidelines.


Topic(s): coronary angiography

Background: At Harvard Community Health Plan (HCHP), Brookline, Mass, a mixed-model health maintenance organization (HMO), coronary angiography is performed at numerous community and tertiary-level teaching hospitals. Objective: To determine the appropriateness of coronary angiography within HCHP according to RAND (1992) criteria and to examine the relationship between the appropriateness rating and (1) the clinical indication for catheterization and (2) the extent of anatomic disease. Method: A retrospective, randomized hospital medical record review of 292 patients enrolled in HCHP who underwent coronary angiography in 1992, stratified by four distinct HCHP subgroups. Results: Of the coronary angiographies reviewed, 78% were rated appropriate, 16% uncertain, and only 6% inappropriate across the entire sample. Ratings were comparable in all subdivisions of HCHP despite an incidence rate of catheterization in one of the three HMO divisions that was 60% and 40% higher than in the other two divisions. The lowest appropriateness ratings were for Asymptomatic patients (43%) and those with Chest Pain of Uncertain Origin (35%) (capital letters refer to the RAND clinical indication criteria mentioned above). A rating of necessity was not a better discriminator of anatomic disease than a rating of appropriateness alone: 82% and 84%, respectively, were found to have disease by angiography. Conclusion: The low HCHP rate of inappropriateness for coronary angiography is comparable with the RAND 1992 New York State data. This finding, coupled with marked differences in the incidence rate of this procedure among the HCHP divisions, is consistent with either major differences in the sickness of the HMO's sub-populations or, more likely, a lack of specificity of the RAND criteria for coronary angiography.

Topic(s): coronary artery bypass graft surgery, percutaneous transluminal coronary angioplasty


Topic(s): upper gastrointestinal endoscopy

**Background and study aims:** This prospective study tested the appropriateness of referrals for upper gastrointestinal endoscopy in an open-access endoscopy unit, using the criteria of the American Society for Gastrointestinal Endoscopy. It also examined whether there was any relationship between appropriateness of use and the presence of significant lesions detected by endoscopy. **Methods:** Four hundred fifty consecutive upper gastrointestinal endoscopies were studied prospectively. The referral indication was recorded by the endoscopist before the procedure was performed, and was compared with the current criteria of the American Society for Gastrointestinal Endoscopy and with endoscopic findings. **Results:** The appropriateness of referral was assessed in 442 consecutive endoscopies. Of these, 252 (57%) were judged to be appropriate. In 168 (88%) of the 190 endoscopies rated as inappropriate, the reason was that the patient had not undergone empirical anti-ulcer therapy before endoscopy. The probability of finding a significant lesion did not differ between the endoscopies judged to be appropriate (50%) and those judged to be inappropriate (46%). **Conclusions:** Upper gastrointestinal endoscopy was frequently used for inappropriate indications. The main reason for inappropriate use was insufficient treatment, or no treatment, of dyspeptic symptoms prior to endoscopy. In this study, the criteria for appropriateness did not predict the probability of finding a significant endoscopic lesion.


Topic(s): coronary angioplasty

Rating the indications for invasive cardiac procedures often is regarded as one key method for quality assessment in cardiology. The aim of this study was to evaluate the validity of the method proposed by the RAND Corporation and the University of California, Los Angeles, appropriateness and necessity rating. Two hypotheses were tested: 1. The acute and long-term benefit of PTCA is less clearly in cases rated "uncertain" than in cases rated "appropriate" or "necessary", and 2. in cases rated "necessary", successful PTCA improves the patients long-term prognosis, since RAND's definition of "necessary" implies that a successful PTCA will avert major harm from this patient. **Patients:** Five hundred and one consecutive patients dilated between 1981 and 1984 were included. **Baseline data:** Age 52.5 ± 8.2 years, male gender 87.0%, Canadian Heart Classification (CHC) class I or II 40.5%, positive and extremely positive stress test 54.4% and 33.3%, single vessel disease 72.5%. The long-term follow-up was determined by questionnaire 91 ± 21 months after PTCA, and a complete 5-year follow-up was
available in 95.4% of all patients. Results: According to RAND's 1991 published criteria, 1.0% of all indications were rated "inappropriate", 21.4% "uncertain", 27.1% "appropriate", and 50.5% "necessary". Before PTCA, patients rated "uncertain" were less symptomatic and had a higher exercise capacity than patients rated "necessary" (CHC I or II: 85.0% vs. 8.7%, p < 0.001; 693 ± 214 vs. 520 ± 251 watts*min, p < 0.001). Following PTCA, "uncertain" patients reported less often a lasting symptomatic improvement or freedom of symptoms than "necessary" patients (63.9% vs. 76.3%, p = n.s.), and they experienced no relevant improvement in exercise capacity (10 ± 263 vs. 139 ± 308 watts*min, p < 0.05). The acute success rate was significantly lower for patients who underwent PTCA for indications rated "revascularization necessary" or "CABG necessary", p < 0.03. This resulted in a high rate of CABG within the first year for patients with these ratings (29.0% and 36.1%, compared to 15.0% to 20.6% in cases rated "uncertain", "appropriate", or "PTCA necessary", p < 0.02). Only in patients with the rating "PTCA necessary" (n = 184), a significant difference in long-term survival was observed between successful and unsuccessful uncomplicated cases (5-year-survival-probability 97.8 ± 1.3% vs. 85.5 ± 6.0%, p < 0.001). In contrast, "uncertain" cases (n = 107) had a 5-year-survival-probability of 94.8 ± 2.5% following successful and 94.4 ± 5.4% following unsuccessful PTCA. Conclusion: These results are in agreement with both hypotheses and therefore support the validity of RAND's 1991 appropriateness and necessity criteria for PTCA in a historic series of patients. Thus, even if these particular ratings may be outdated today, the technique of appropriateness and necessity rating is expected to be a useful and valid tool for quality assessment in PTCA.


Topic(s): coronary angiography, percutaneous transluminal coronary angioplasty

Objective: To assess the indications for coronary angiography and percutaneous transluminal coronary angioplasty (PTCA) according to the criteria of the RAND Corporation's expert panel ratings; to compare the results with those already published in the literature; and to examine the method with respect to its appropriateness as a measure of quality control. Patients and methods: The parameters necessary for rating according to the RAND Corporation's published criteria were prospectively obtained in 116 consecutive patients (89 men, 27 women; mean age 59.4 +/- 10.7 years) undergoing coronary angiography and 138 patients (112 men, 26 women; mean age 61.5 +/- 9.4 years) undergoing PTCA. Results: For coronary angiography the >>inappropriate rate << was 22.4%, for >> uncertain indications << 15.5%, >> appropriate indications << 24.1%, and for >> necessary indications << 37.9%. 38.6% of coronary angiographies, performed in the course of angiographic control after PTCA, were >> inappropriate <<, but only 12.5% of other coronary angiographies (P < 0.01). In 48.6% of >> appropriate << or >> necessary << procedures invasive treatment followed, compared with 19.2% of those rated >> inappropriate << (P < 0.05). In 10.1% of patients the indications for PTCA were judged >> inappropriate <<, in 32.6% as >> uncertain <<, in 13.0% as >> appropriate << and 44.2% as >> necessary <<. There was no correlation between ratings and the acute results of PTCA. These findings pertaining to indications for coronary angiography and PTCA correspond to those reported in the literature. Conclusions: As the RAND criteria take inadequate account of individual peculiarities, they are not suitable for individual clinical decisions. But they are useful as screening method in a quality control project,
because procedures for which indications have not been adequately proven are singled out by an unfavorable rating and can thus be thoroughly analyzed in the individual case.


**Topic(s):** coronary artery bypass graft surgery

**Objective:** To compare the appropriateness of use of coronary artery bypass graft (CABG) surgery in Academic Medical Center Consortium hospitals as judged 1) according to criteria developed by an expert panel, 2) according to revisions of those criteria made by cardiac surgeons from the Academic Medical Center Consortium, and 3) by review of cases by the surgeons responsible for those cases. **Design:** Retrospective, randomized medical record review. **Setting:** 12 Academic Medical Center Consortium hospitals. **Patients:** Random sample of 1156 patients who had had isolated CABG surgery in 1990. **Main outcome measures:** 1) Percentage of patients with indications for which CABG surgery was classified as appropriate, Inappropriate, or of uncertain appropriateness and 2) percentage of cases in which CABG surgery was judged inappropriate or uncertain for which ratings changed after local case review. **Results:** Data were retrieved from medical records by trained abstractors using an explicit data collection instrument. Cases in which CABG surgery was judged to be inappropriate or uncertain were individually reviewed by the responsible surgeons. According to the expert panel ratings, 83% of the CABG operations (95% CI, 81% to 85%) were necessary, 9% (CI, 8% to 10%) were appropriate, 7% (CI, 5% to 8%) were uncertain, and 1.6% (CI, 0.6% to 2.5%) were inappropriate. These rates are almost identical to those found in a previous study that was done in New York State and that used the same criteria (in that study, 91% of operations were classified as necessary or appropriate, 7% were classified as uncertain, and 2.4% were classified as inappropriate). Rates of inappropriate procedures varied from 0% to 5% among the 12 member hospitals (P = 0.02). The Academic Medical Center Consortium cardiac surgeons revised 568 (24%) of the indications used by the expert panel. However, because those revisions altered the appropriateness ratings in both directions and affected only 50 cases (4%), the net effect of the revisions was slight: The rate of inappropriate CABG surgery increased from 1.6% to 1.9%. Local review found that data collection errors had caused erroneous ratings in 12.5% of 64 cases in which surgery had been classified as inappropriate or uncertain. **Conclusions:** The Academic Medical Center Consortium hospitals had low rates of inappropriate and uncertain use of CABG surgery, regardless of the criteria used for assessment. Even though surgeons from the Consortium revised the appropriateness ratings extensively, their revisions had a negligible effect on the overall assessment of appropriateness. However, because of potential data collection errors, appropriateness criteria should be used for individual case audits only if supplemented by subsequent physician review.

**Topic(s):** open-access endoscopy

**Background:** Open-access endoscopy allows non-gastroenterologist physicians the opportunity to directly schedule elective common endoscopic procedures for their patients without having them first examined in the gastrointestinal clinic. There are few data as to whether non-gastroenterologist physicians in the United States schedule patients for appropriate indications. **Objectives:** To examine our practice to see whether patients undergoing open-access endoscopy were scheduled for appropriate indications and to see whether there were differences among physicians in various medical specialties. **Methods:** We prospectively tracked 310 consecutive patients scheduled for open-access esophagogastroduodenoscopy (EGD) and colonoscopy by non-gastroenterologist physicians over a 9-month period in our academic practice setting to determine whether the indications for performing the procedures were appropriate. The American Society for Gastrointestinal Endoscopy criteria (revised in 1992) were used as the standard for comparison. **Results:** Primary care physicians (family practitioners and general internists) did a superior job of scheduling patients for appropriate indications for EGD and colonoscopy than did non-primary care physicians (internal medicine subspecialists and surgeons): 97.0% vs. 81.3% for EGD (P = .04) and 84.9% vs. 66.7% for colonoscopy (P = .02), respectively. **Conclusions:** Primary care physicians were significantly more likely to schedule patients for open-access EGD and colonoscopy for appropriate indications than were non-primary care physicians. The frequency of inappropriate indications for colonoscopy referrals was greater than for EGD. The reasons for the differences among primary care physicians, surgeons, and internal medicine subspecialists require further exploration.


**Topic(s):** cataract surgery

**Purpose:** To develop criteria for the appropriateness of cataract surgery (extracapsular cataract extraction or phacoemulsification with planned implantation of a posterior chamber intraocular lens) and to apply these criteria to patients from ten academic medical centers. **Methods:** The study is a retrospective case series from ten academic medical centers. One thousand one hundred thirty-nine patients who had had cataract surgery in 1990 at the medical centers were selected randomly. Patients, identified by specific ICD-9-CM or CPT-4 codes, had no other ocular surgery performed at the same time as cataract surgery. Rates of inappropriate, uncertain, appropriate, and appropriate and crucial surgeries were determined by application of the criteria established by a multidisciplinary expert panel. **Results:** Approximately 2% of the procedures were classified as inappropriate, after adjusting for missing or nonspecific visual function by use of discriminant analysis. Ninety-one percent of the procedures were classified as appropriate (52%) or appropriate and crucial (39%). Seven percent were designated as uncertain, either due to a median rating in the uncertain range or to disagreement in ratings among the panelists. Significant variation occurred in the results among the different institutions: inappropriate surgeries ranged from 0% to 4%, uncertain from 1%
to 14%, appropriate from 35% to 66%, and appropriate and crucial from 21% to 62% (P=0.02). **Conclusion:** A small percentage of cataract surgeries was performed at these ten academic medical centers for inappropriate indications using the study criteria. Given the large number of cataract surgeries performed annually, the small percentage of uncertain and inappropriate surgeries may translate into a large number of surgeries performed for less than appropriate or appropriate and crucial indications. Significant variation existed among the institutions in the distribution of appropriate and crucial and appropriate compared with uncertain and inappropriate surgeries.


**Topic(s):** upper gastrointestinal endoscopy

**Background and aims:** Efforts to reduce costs in health care may raise concerns about underuse of medical procedures. This study prospectively assessed underuse of upper gastrointestinal endoscopy in a cohort of patients in whom we have recently published data on overuse of endoscopy. **Methods:** Underuse was identified by formal necessity criteria for endoscopy, obtained by an explicit panel process. Outpatients were consecutively included in two clinical settings. Setting A consisted of 20 primary care physicians and 7215 patient visits that occurred within 1 month. Setting B consisted of 920 visits that occurred during 3 weeks at an outpatient clinic. **Results:** During 8135 visits, 611 patients complained of upper digestive symptoms; 63 of them underwent endoscopy. Underuse was identified in 72 patients (11.8%). The two clinical situations mainly responsible for underuse of endoscopy were underinvestigated peptic symptoms resistant to treatment and dysphagia. At first follow-up, 29 of the patients with initial underuse still fulfilled criteria of necessity (underuse rate, 4.7%). One-year follow-up showed underuse of endoscopy in 5 patients. **Conclusions:** This prospective evidence shows that underuse of a medical procedure exists. The estimated overuse and underuse of endoscopy in this cohort were approximately equal (5%). Improving quality of care will require reductions of both overuse and underuse of medical procedures.


**Topic(s):** upper gastrointestinal endoscopy

**Background:** This prospective observational study was aimed at evaluating the appropriateness of use of upper gastrointestinal endoscopy (UGE) in primary care in a country with open access to and high availability of the procedure. **Methods:** Outpatients were consecutively included in two clinical settings: Setting A (20 primary care physicians during 4 weeks) and B (university-based outpatient clinic during 3 weeks). In patients undergoing UGE, appropriateness of referral was judged by explicit Swiss criteria developed by the RAND/UCLA panel method. **Results:** Patient visits (8135) were assessed. Six hundred eleven patients complained of upper gastrointestinal
symptoms. Physicians decided to perform UGE in 63 of these patients. Twenty-five (40%) of the endoscopies were rated appropriate, 7 (11%) equivocal, and 31 (49%) inappropriate. Overuse of UGE occurred in 5.1% (setting A: 4.7%; setting B: 6.5%; p = 0.39) of the patients who presented with upper gastrointestinal symptoms. The decision to perform UGE in previously untreated dyspeptic patients was the most common clinical situation resulting in overuse. **Conclusions:** Inappropriate use of UGE is high in Switzerland. However, to better reflect primary care decision making, overuse should be related not only to patients referred for a medical test, but also to the number of patients who complain of the symptoms that would be investigated by the procedure.


**Topic(s):** coronary angiography

**Objective:** To estimate the extent of under use of coronary angiography and to determine whether women, ethnic minorities and poor and uninsured patients are less likely than their counterparts to receive necessary coronary angiography. **Design:** Retrospective cohort study employing chart review and patient interviews. **Setting:** Four teaching hospitals: three government owned (public) and one private university medical center in Los Angeles, California. **Patients:** Three hundred and fifty two patients who had a positive exercise stress test between 1 January 1990 and 30 June 1991 and met explicitly defined criteria for the necessity of coronary angiography established by a multidisciplinary expert panel. **Main outcome measures:** Percentage of patients who received necessary coronary angiography within 3 and 12 months following exercise stress testing, adjusted for demographic and clinical characteristics using logistic regression. **Results:** Overall 43% received necessary coronary angiography within 3 months and 56% within 12 months of the stress test. Women were less likely than men to receive necessary coronary angiography. Adjusted odds ratio (AOR) 0.54, 95% confidence interval (CI) 0.34-0.90 for angiography within 3 months of the stress test; AOR 0.47, 95% CI 0.29-0.77 for angiography within 12 months of the stress test. Public hospital patients underwent necessary coronary angiography less often than private hospital patients. AOR 0.40, 95% CI 0.23-0.79 for within 3 months; AOR 0.52, 95% CI 0.30-0.91 for within 12 months. **Conclusions:** Under use of coronary angiography can be measured and occurs to a significant degree. It is important to develop standards of quality to address and safeguard against under use of necessary medical care.


**Topic(s):** coronary revascularization

**Objectives:** Our main objective was to apply a new method to determine whether coronary revascularization procedures are underused, especially among African-Americans and uninsured patients. **Background:** Although overuse of revascularization procedures has been studied, underuse as defined clinically has not been examined before. **Methods:** The study was conducted at four public and two academically affiliated
private hospitals in Los Angeles; 671 patients who underwent coronary angiography between June 1, 1990 and September 30, 1991 and met explicit clinical criteria for coronary revascularization were included. The main outcome measure was the proportion of patients undergoing an indicated procedure within 12 months (ascertained by medical record review and confirmed with a telephone survey). Adjusted relative odds of undergoing an indicated procedure for African-Americans and patients in public hospitals compared with whites and patients in private hospitals were calculated. **Results:** Overall, 75% of patients underwent a revascularization procedure. Of 424 patients requiring bypass surgery, 107 angioplasty and 140 either bypass surgery or angioplasty, 59%, 66% and 75% underwent the procedure, respectively. African-Americans were less likely than whites to undergo operation (adjusted odds ratio [OR] 0.49, p < 0.05) and angioplasty (adjusted OR 0.20, p < 0.05). Patients in public hospitals were less likely than those in private hospitals to undergo angioplasty (adjusted OR 0.10, p < 0.005). **Conclusions:** Underuse of coronary revascularization procedures is measurable and occurs to a significant degree even among insured patients attending private hospitals. Underuse is especially pronounced among African-Americans and patients attending public hospitals. Future cost-containment efforts must incorporate safeguards against underuse of necessary care.


**Topic(s):** surgery of lumbar disc hernia and spinal stenosis

**Study design:** This prospective study examines the appropriateness of indications for surgery of herniated intervertebral disc and spinal stenosis in patients undergoing surgery in a university hospital setting. **Objective:** To evaluate the appropriateness of surgery using explicit criteria developed by an expert panel in the United States. **Summary and background data:** The use of surgery for herniated intervertebral disc and spinal stenosis varies widely within and among countries. It has been postulated that the main reason for treatment failure is poor selection of candidates for the procedure. **Methods:** The authors prospectively evaluated appropriateness of surgical indications for herniated lumbar intervertebral disc or spinal stenosis in 328 consecutive patients undergoing the operation in two university neurosurgery departments. Outcome was measured 1 year after surgery by a standardized interview. **Results:** Indications for surgery were considered to be appropriate or equivocal in 202 (62%) patients and inappropriate in 126 (38%). Among the 126 inappropriate procedures, 66 were so rated because of insufficient activity restriction before the procedure. One year after surgery, 74% of the patients perceived the results of the operation as good or very good. **Conclusions:** Appropriateness as measured by the criteria established by the American panel identified a large percentage of day-to-day practice in the two surgical units as inappropriate. However, use of criteria that include new findings about lack of efficacy of bed rest probably would lower this percentage. Criteria of appropriateness of medical and surgical procedures, developed through the panel process, need to be updated regularly.

**Topic(s):** hysterectomy

The report presents recommendations based on appropriateness ratings of indications for hysterectomy. The recommendations were developed to provide an authoritative guide to help physicians decide whether to recommend hysterectomy for non-emergency, non-malignant disease. Readers are assumed to be knowledgeable about the full spectrum of gynecological disease and its management, including indications for and interpretation of diagnostic tests, and are qualified to provide pre-, intra-, and post-operative care. The recommendations, which represent expert judgments about the appropriateness of hysterectomy for a large number of clinical scenarios, were developed by a working group of obstetricians/gynecologists. The appropriateness ratings and the literature review that accompany these recommendations are provided in two companion documents.


**Topic(s):** carotid endarterectomy

**Objective:** To examine specifically the influence of estimated perioperative mortality and stroke rate on the assessment of appropriateness of carotid endarterectomy. **Data sources/study setting:** An expert panel convened to rate the appropriateness of a variety of potential indications for carotid endarterectomy based on various rates of perioperative complications. We then applied these ratings to the charts of 1,160 randomly selected patients who had carotid endarterectomy in one of the 12 participating academic medical centers. **Study design:** An expert panel evaluated indications for carotid endarterectomy using the modified Delphi approach. Charts of patients who received surgery were abstracted, and clinical indications for the procedure as well as perioperative complications were recorded. To examine the impact of surgical risk assessment on the rates of appropriateness, three different definitions of risk strata for combined perioperative death or stroke were used: Definition A, low risk < 3 percent; Definition B, low risk < 5 percent; and Definition C, low risk < 7 percent. **Principal findings:** Overall hospital-specific mortality ranged from 0 percent to 4.0 percent and major complications, defined as death, stroke, intracranial hemorrhage, or myocardial infarction, varied from 2.0 percent to 11.1 percent. Most patients (72 percent) had surgery for transient ischemic attack or stroke; 24 percent of patients were asymptomatic. Most patients (82 percent) had surgery on the side of a high-grade stenosis (70-99 percent). When the thresholds for operative risk were placed at the values defined by the expert panel (Definition A), only 33.5 percent of 1,160 procedures were classified as "appropriate." When the definition of low risk was shifted upward, the proportion of cases categorized as appropriate increased to 58 percent and 81.5 percent for Definitions B and C, respectively. **Conclusions:** Despite the high proportion of procedures performed for symptomatic patients with a high degree of ipsilateral extracranial carotid artery stenosis and generally low rates of
surgical complications at the participating institutions, the overall rate of "appropriateness" using a perioperative complication rate of ≤ 3 percent was low. However, the rate of "appropriateness" was extremely sensitive to judgments about a single clinical feature, surgical risk. These data show that before applying such "appropriateness" ratings, it is crucial to perform sensitivity analyses in order to assess the stability of the results. Results that are robust to moderate in variation in surgical risk provide a much sounder basis for policy making than those that are not.


**Topic(s):** coronary artery bypass grafting, percutaneous transluminal coronary angioplasty

**Objective:** To determine the appropriateness of intention to treat decisions concerning coronary artery bypass grafting (CABG) and percutaneous transluminal coronary angioplasty (PTCA) for patients with coronary artery disease in The Netherlands.

**Design:** Prospective study of intention to treat decisions using a computerized expert system.

**Setting:** “Presentation” sessions in 10 tertiary referral heart centers in 1992.

**Patients:** 3207 consecutive patients: 1618 CABG and 1589 PTCA candidates.

**Main outcome measure:** Percentage of invasive treatment decisions rated appropriate, uncertain, or inappropriate by the expert system.

**Results:** PTCA decisions were common for patients with one-vessel disease and CABG decisions for patients with three-vessel and left main disease. PTCA decisions outnumbered CABG decisions in acute myocardial infarction. Of CABG decisions, 84% were rated appropriate, 12% uncertain, and 4% inappropriate. The proportions for PTCA decisions were 39% appropriate, 31% uncertain, and 29% inappropriate. Type C lesion was the main determinant of inappropriateness of PTCA decisions. If type C lesions were downgraded to type A/B lesions the rate of inappropriate PTCA decisions dropped to 6%.

**Conclusions:** Clinicians in tertiary referral centers in The Netherlands favored CABG if vessel disease was extensive or involved the left main artery, and PTCA for patients with less extensive disease and with acute myocardial infarction. Few CABG decisions were inappropriate. The main determinant of inappropriateness of PTCA decisions was its intended use in patients with type C lesions.


**Topic(s):** colonoscopy

**Background:** Prospective data describing the appropriateness of use of colonoscopy based on detailed panel-based clinical criteria are not available. **Methods:** In a cohort of 553 consecutive patients referred for colonoscopy to two university-based Swiss outpatient clinics, the percentage of patients who underwent colonoscopy for appropriate, equivocal, and inappropriate indications and the relationship between appropriateness of use and the presence of relevant endoscopic lesions was prospectively assessed. This
assessment was based on criteria of the American Society for Gastrointestinal Endoscopy and explicit American and Swiss criteria developed in 1994 by a formal panel process using the RAND/UCLA appropriateness method. **Results:** The procedures were rated appropriate or equivocal in 72.2% by criteria of the American Society for Gastrointestinal Endoscopy, in 68.5% by explicit American criteria, and in 74.4% by explicit Swiss criteria (not statistically significant, NS). Inappropriate use (overuse) of colonoscopy was found in 27.8%, 31.5%, and 25.6%, respectively (NS). The proportion of appropriate procedures was higher with increasing age. Almost all reasons for using colonoscopy could be assessed by the two explicit criteria sets, whereas 28.4% of reasons for using colonoscopy could not be evaluated by the criteria of the American Society for Gastrointestinal Endoscopy (p < 0.0001). The probability of finding a relevant endoscopic lesion was distinctly higher in the procedures rated appropriate or equivocal than in procedures judged inappropriate. **Conclusions:** The rate of inappropriate use of colonoscopy is substantial in Switzerland. Explicit criteria allow assessment of almost all indications encountered in clinical practice. In this study, all sets of appropriateness criteria significantly enhanced the probability of finding a relevant endoscopic lesion during colonoscopy.


**Topic(s):** unstable angina


**Topic(s):** coronary angiography

**Objective:** To evaluate the appropriateness of referral following coronary angiography in Sweden. **Design:** Prospective population survey and review of medical records. **Setting:** Seven of the 8 hospitals in Sweden that perform coronary revascularization. **Patients:** Consecutive series of 2786 patients who underwent coronary angiography in Sweden between May 1, 1994 and January 15, 1995. **Main Outcome Measures:** Percentage of patients referred for CABG or PTCA for necessary, appropriate, uncertain and inappropriate indications and the percentage of patients referred for continued medical therapy who met necessity criteria for revascularization. **Results:** Almost half the patients were referred for CABG, 25% for PTCA and 26% for continued medical therapy. CABG was judged necessary for 68% of patients, appropriate for 1%, uncertain for 12% and inappropriate for 20%. For PTCA the figures were 16%, 11%, 28% and 46%, respectively. Of the 597 patients referred for inappropriate coronary revascularization, 46 percent were referred after undergoing coronary angiography during an admission for unstable angina. Among patients referred for continued medical therapy, 4.3% met necessity criteria for undergoing coronary revascularization but were considered too healthy by their physician for revascularization. **Conclusions:** Many patients were referred for PTCA and CABG for inappropriate or uncertain indications. Three factors accounted for the high inappropriate rate including referring patients who: (1) did not have significant coronary artery disease; (2) were treated with suboptimal doses of medications; and (3) were referred during an admission for unstable angina.
These data raise questions as to the degree of overuse of coronary revascularization in Sweden and the clinical reasoning behind the Swedish appropriateness criteria.


Topic(s): coronary artery bypass grafting, percutaneous transluminal coronary angioplasty


Topic(s): perioperative use of recombinant erythropoietin


Topic(s): upper gastrointestinal endoscopy


Topic(s): coronary angiography


Topic(s): laminectomy

We have developed criteria to determine the appropriate indications for lumbar laminectomy, using the standard procedure developed at the RAND corporation and the University of California at Los Angeles (RAND-UCLA). A panel of five surgeons and four physicians individually assessed 1000 hypothetical cases of sciatica, back pain only, symptoms of spinal stenosis, spondylolisthesis, miscellaneous indications or the need for repeat laminectomy. For the first round each member of the panel used a scale ranging from 1 (extremely inappropriate) to 9 (extremely appropriate). After discussion and condensation of the results into three categories laminectomy was considered appropriate in 11% of the 1000 theoretical scenarios, equivocal in 26% and inappropriate in 63%. There was some variation between the six categories of malalignment, but full agreement in 64% of the hypothetical cases. We applied these criteria retrospectively to the record of 196 patients who had had surgical treatment for herniated discs in one Swiss University hospital. We found that 48% of the operations were for appropriate indications, 29% for equivocal reasons and that 23% were inappropriate. The RAND-
UCLA method is a feasible, useful and coherent approach to the study of the indications for laminectomy and related procedures, providing a number of important insights. Our conclusions now require validation by carefully designed prospective clinical trials, such as those which are used for new medical techniques.


Topic(s): upper gastrointestinal endoscopy

Objectives: To describe and compare both overuse and underuse of diagnostic upper gastrointestinal endoscopy in different settings. 

Design: Merging of data from three prospective observational studies. The appropriateness and necessity of indications for gastroscopy were evaluated using explicit criteria developed by a standardized expert panel method (RAND-UCLA). Inappropriate endoscopies represent overuse. Necessary indications not referred for the procedure constitute underuse. Setting: Three primary care outpatient clinics, 20 general practices, 3 gastroenterology practices, 2 district and 1 university hospitals. Subjects: A third of the collective were consecutive ambulatory patients with upper abdominal complaints, while the other two-thirds were ambulatory and hospitalized patients referred for the procedure. 

Main outcome measures: Proportions of overuse and underuse in the different settings. Results: 2885 patients were included (mean age: 49 years, 52% males, 2442 outpatients), 1858 patients underwent ≥ 1 endoscopy. Among 2086 endoscopies, 805 (39%) were inappropriate, most of which were performed for dyspepsia (83%). Overuse was higher in young, foreign, female patients and lower in inpatient settings, the latter reflecting a different distribution of presenting symptoms. Among 1646 patient visits in primary care, overuse represented 148 endoscopies (9%). Underuse was identified in 104 of the same patient visits (6%). It was higher as patient age increased; there were no significant differences between men and women. Conclusions: Rates of overuse and underuse depend mainly on case presentation and patient characteristics. Both over- and underuse should be addressed to maintain and improve quality of care.


Topic(s): coronary artery bypass grafting, percutaneous transluminal coronary angioplasty
B.2) Panel results (possibly with B.3)


Topic(s): coronary artery bypass graft surgery

This report is part of a larger study to determine the extent to which the appropriate use of medical procedures is related to overall rates of use among geographic areas. It is one of six volumes that report the ratings of panels of physicians on the appropriateness of performing each of six procedures for a wide variety of clinically specific indications. The six procedures were coronary angiography, coronary artery bypass graft surgery, cholecystectomy, diagnostic gastrointestinal endoscopy, colonoscopy, and carotid endarterectomy. This volume contains a literature review of coronary artery bypass graft surgery, a comprehensive set of indications for performing this procedure, and ratings of the appropriateness for each indication.


Topic(s): coronary angiography


Topic(s): upper gastrointestinal endoscopy

This report contains a review of the published literature on the diagnostic use of upper gastrointestinal endoscopy from the procedure's introduction in the 1960s through 1981. In addition, it includes data on its findings, complications, costs, efficacy, utilization, and indications. It rates the appropriateness of each indication, and provides a medical record abstraction form and guidelines for its use.


Topic(s): colonoscopy

B.2.05 Merrick, N.J., A. Fink, R.H. Brook, R.E. Park, J. Kosecoff, C.P. Roth, L. McCloskey, J.W. Keesey, and D.H. Solomon, *Indications for Selected Medical and Surgical Procedures--A Literature Review and Ratings of Appropriateness: Carotid

**Topic(s):** carotid endarterectomy

This report is part of a larger study to determine the extent to which the appropriate use of medical procedures is related to overall rates of use among geographic areas. It is one of six volumes that report the ratings of panels of physicians on the appropriateness of performing each of six procedures for a wide variety of clinically specific indications. The six procedures were coronary angiography, coronary artery bypass graft surgery, cholecystectomy, diagnostic gastrointestinal endoscopy, colonoscopy, and carotid endarterectomy. This volume contains a literature review of carotid endarterectomy, a comprehensive set of indications for performing this procedure, and ratings of the appropriateness for each indication.


**Topic(s):** coronary angiography, coronary artery bypass graft surgery, percutaneous transluminal coronary angioplasty, carotid endarterectomy, colonoscopy, upper gastrointestinal endoscopy

We convened three panels of physicians to rate the appropriateness of a large number of indications for performing a total of six medical and surgical procedures. The panels followed a modified Delphi process. Panelists separately assigned initial ratings, then met in Santa Monica, California where they received reports showing their initial ratings and the distribution of the other panelists ratings. They discussed the indications and revised the indications lists, then individually assigned final ratings. There was generally better agreement on the final ratings than on the initial ratings. Based on reasonable criteria for agreement and disagreement, and excluding one outlying procedure, the panelists agreed on ratings for 42 to 56 per cent of the indications, and disagreed on 11 to 29 per cent.


**Topic(s):** cholecystectomy

This report is part of a larger study to determine the extent to which the appropriate use of medical procedures is related to overall rates of use among geographic areas. It is one of six volumes that report the ratings of panels of physicians on the appropriateness of performing each of six procedures for a wide variety of clinically specific indications. The six procedures were coronary angiography, coronary artery bypass graft surgery, cholecystectomy, diagnostic gastrointestinal endoscopy, colonoscopy, and carotid endarterectomy. This volume contains a literature review of cholecystectomy, a comprehensive set of indications for performing this procedure, and ratings of the appropriateness for each indication.

Topic(s): endoscopy, cholecystectomy

We evaluated the effect of patients' comorbidity on the appropriateness of performing esophagogastroduodenoscopy or cholecystectomy. A nine-member national physician panel rated 1118 brief clinical scenarios for patients without comorbidity. Ratings were then repeated for patients with increasing degrees of comorbidity. As comorbidity changed from none to medium, 60% of those scenarios that were originally rated as appropriate for endoscopy and cholecystectomy remained appropriate. As high comorbidity was introduced, only 13% of such scenarios remained appropriate for endoscopy, while 33% remained appropriate for cholecystectomy. These findings suggest that, although clinical reasons for performing procedures are a powerful determinant of when they should be used, comorbidity is also important and needs to be included in any assessment of the appropriateness of procedure use.


Topic(s): percutaneous transluminal coronary angioplasty

This report summarizes information on percutaneous transluminal coronary angioplasty (PTCA) contained in research published between 1982 and October 1990. In particular, it deals with research on the efficacy, complications, costs, utilization, and indications for PTCA. The report includes a comprehensive set of indications for performing this procedure, and ratings of the appropriateness and necessity for each indication.


Topic(s): coronary artery bypass graft

This report summarizes the evidence for efficacy and risk according to generally accepted indications for coronary artery bypass graft surgery. It summarizes information contained in 322 papers published between 1971 and 1990. It also provides a description of the expert panel process and the results of their ratings of appropriateness and necessity.


Topic(s): spinal manipulation for low-back pain
This report contains the indications and ratings for appropriateness for spinal manipulation that reflect the findings of a nine-member panel of back-pain experts. The panel members rated the appropriateness of indications twice, using a nine-point scale. The initial ratings of appropriateness were done individually and without group discussion. The second-round ratings used a structured method based on consensus procedures. The results showed that these physicians were able to formulate detailed lists of indications for spinal manipulation for low-back pain and rate their appropriateness. There was an increase in agreement and a decrease in disagreement between the initial round and the final round. The final round rated 1,550 indications, and there was agreement on 36 percent and disagreement on 12 percent of the indications. The level of disagreement was somewhat less than in previous ratings of medical procedures by all-medical panels.


**Topic(s):** abdominal aortic aneurysm surgery

This report reviews the literature on the benefits and risks of abdominal aortic aneurysm (AAA) surgery, provides a comprehensive list of potential indications for AAA surgery, and presents the ratings of appropriateness and necessity of those indications by a panel of experts. The ratings of appropriateness and necessity are the final results of a two-stage modified Delphi process and were assigned by a multispecialty panel of physicians with expertise in the diagnosis and management of AAA.


**Topic(s):** hysterectomy

This report presents a literature review of the state of the art in treating conditions for which hysterectomy represents one alternative, a comprehensive set of indications for performing the procedure, and ratings of appropriateness for each indication. The authors reviewed 348 articles and 19 textbooks that addressed the utilization, effectiveness, complications, costs, and indications associated with hysterectomy. The review of effectiveness was organized around the non-oncologic clinical conditions for which hysterectomy is used: dysmenorrhea, endometriosis, adenomyosis, recurrent uterine bleeding, leiomyoma, chronic pelvic pain, endometrial hyperplasia, pelvic relaxation, pelvic inflammatory disease, and cervical dysplasia. The indications and ratings for hysterectomy reflect the findings of a nine-member panel of physicians associated with managed health care organizations. The ratings of appropriateness are the final result of a two-stage process. The indications were initially rated independently by each physician without discussion or contact with the other panelists. The panel then assembled and after discussion, each panelist independently and confidentially rated each indication again. The median rating is used as the appropriateness score. The results show that physicians were able to formulate comprehensive lists of indications for hysterectomy and come closer during the second round of ratings to agreement on their appropriateness ratings.
Panelists' agreement about the ratings increased from 50.1 to 65.7 percent after the second round. Disagreement also decreased significantly from 13.9 to 4.5 percent.


Topic(s): coronary angiography

This report contains a literature review of coronary angiography, a comprehensive set of indications for performing this procedure, and ratings of the appropriateness and necessity for each indication. It primarily concentrates on coronary angiography studies; however, studies examining percutaneous transluminal coronary angioplasty (PTCA) and coronary artery bypass graft (CABG) surgery are also included to establish the efficacy of coronary angiography in coronary artery disease patients. Findings show that where significant coronary artery disease is defined as a 50 percent reduction in luminal diameter, coronary artery disease occurred in almost 80 percent of patients sampled; with a 70 percent reduction in luminal diameter, there was 72 percent occurrence. Using either criterion, significant left main or three-vessel coronary artery disease was present in approximately 35 percent of procedures sampled. Data on complications associated with coronary angiography show that patients with left main coronary artery disease have a threefold greater risk of myocardial infarction and sevenfold greater risk of death. Other complications include cerebrovascular accidents (0.07%), ventricular arrhythmias (0.47%), local vascular injury (0.59%), and contrast dye reactions (0.23%). The average myocardial infarction rate was 0.07% and the average mortality rate was 0.10% for 1981-1989. The authors draw conclusions concerning the efficacy of coronary angiography for nine specific clinical conditions, based on a comparison of CABG surgery, medical therapy, and observational studies of PTCA: Coronary angiography is efficacious for patients with both chronic stable angina and unstable angina, whose angina cannot be satisfactorily controlled by medical therapy; evidence supports the efficacy of coronary angiography in acute myocardial infarction patients who experience cardiogenic shock or an evolving myocardial infarction, and in myocardial infarction survivors who experience angina or have evidence of myocardial ischemia; following revascularization, the only patients shown to clearly benefit from coronary angiography are those who develop severe angina; there are no studies on the efficacy of coronary angiography in survivors of sudden cardiac death without angina; there are conflicting data on efficacy with patients with chest pain of uncertain origin; and there are no data to support efficacy of coronary angiography in patients with ventricular arrhythmias, or in asymptomatic patients.


Topic(s): carotid endarterectomy

**Topic(s):** spinal manipulation for low-back pain

This report presents results from the third stage of the RAND Appropriateness of Spinal Manipulation for Low-Back Pain Study. The study is designed to ascertain the clinical criteria for the appropriate use of spinal manipulation for low-back pain from chiropractors and medical specialists, and then to investigate the use of chiropractic services in a random sample of practicing chiropractors. The third stage describes the methods used for assessing the appropriateness of spinal manipulation by a panel of expert chiropractors and presents the panel's actual ratings of 1570 indications for the use of spinal manipulation for low-back pain. The panel members rated the appropriateness of indications twice, using a nine-point scale. Results showed that these chiropractic physicians were able to formulate detailed lists of indications for spinal manipulation for low-back pain and rate their appropriateness for spinal manipulation.


**Topic(s):** cholecystectomy

A consensus panel approach was used in Israel to develop a list of clinical indications for which there was agreement that cholecystectomy should be performed. Nine physicians from different disciplines were asked to score a list of 266 clinical indications for cholecystectomy. Each indication was scored on a scale of 1 (inappropriate, i.e. health risks exceed health benefits) to 9 (appropriate, i.e. benefits exceed risks). Each indication also included one of four comorbidity level (none to high). Agreement and disagreement were defined and panelists met to discuss, modify and rescore the list. The composition of the panel and definitions of agreement had a considerable impact on the preparation of a list of agreed, appropriate indications for cholecystectomy. Gastroenterologists in the panel were less likely to recommend surgery than either surgeons or general internists both before and after the panel discussion. Following the discussion the level of agreement (defined as after discarding the highest and lowest score all of the remaining seven panelists were in a 3-point range) increased from 39% to 46% (p < 0.08) and disagreement decreased from 27% to 18% (p < 0.01). Fifty-nine of the 266 indications were considered appropriate with agreement.


**Topic(s):** cholecystectomy

Indications as to which patients should undergo cholecystectomy remain, at least in part, a matter of controversy. In 1987, a panel of nine Israeli physicians from different specialties established a list of indications for the performance of cholecystectomy based on the literature available at the time. The panel agreed that cholecystectomy was
appropriate for 59 indications and that it was inappropriate for 58. The major indications for surgery were biliary colic and acute cholecystitis. Patients who were asymptomatic or had vague symptoms were not recommended to undergo surgery unless they had stones in the common bile duct and were less than 71 years of age. Patients with pancreatitis were recommended for surgery in the common bile duct and did not have a history of alcohol abuse. Performing a cholecystectomy at the same time as abdominal surgery was being performed for other reasons was indicated only if the patient was symptomatic from his gall-stones.


Topic(s): coronary angiography

This publication provides the results of a two-stage Delphi process by which a panel of Canadian physicians rated the appropriateness and necessity of indications for the use of coronary angiography. Chapter One, "Methods and Appropriateness Results," summarizes the methods by which the panelists rated the appropriateness of indications for coronary angiography and discusses how those ratings were analyzed. The actual ratings of appropriateness are provided in Appendix A. Chapter Two, "Methods and Necessity Results," summarizes the methods by which the panelists rated the necessity of appropriate indications for coronary angiography and discusses how those ratings were analyzed. The actual ratings of necessity are shown in Appendix B. Chapter Three, "Definitions Used by Panel in Final Ratings of Indications," provides the definitions that were used by the panelists in making their final round of ratings. This publication is part of a three-volume series. Two volumes report ratings for the appropriateness and necessity of performing coronary artery bypass graft surgery, percutaneous transluminal coronary angioplasty and coronary angiography. The other volume provides the medical records abstraction form and guidelines for its use.


Topic(s): cataract surgery


Topic(s): coronary artery bypass graft surgery, percutaneous transluminal coronary angioplasty

This publication provides the definitions used by a nine-member panel of Canadian physicians who rated the appropriateness and necessity of indications for the use of coronary artery bypass graft surgery and percutaneous transluminal coronary angioplasty, the methods by which the ratings of the panel were analyzed, and the final panel ratings
by indications. The results of the panel ratings of appropriateness are provided in Appendix A. The results of the panel ratings of necessity are in Appendix B. This publication is part of a three-volume series. Two volumes report ratings for the appropriateness and necessity of performing coronary artery bypass graft surgery, percutaneous transluminal coronary angioplasty and coronary angiography. The other volume provides the medical records abstraction form and guidelines for its use.


**Topic(s): prostatectomy**

**Study objective:** The use of formal consensus development to determine appropriate indications for prostatectomy and to identify factors underlying clinical decisions about appropriateness is described. **Design:** A nominal group technique was used. **Settings:** The study took place in an academic research institution. **Participants:** The panel consisted of six urologists and three general practitioners. **Measurement and main results:** The panel identified agreed indications for prostatectomy, expressed in terms of different combinations of type of retention, type and severity of symptoms, and level of comorbidity. Agreement was reached for 67% of the indications considered. For acute on chronic retention, surgery is indicated, regardless of symptom severity, if life expectancy is greater than one year. For acute or chronic retention, surgery is generally indicated if symptoms are severe, or if symptoms are moderate and life expectancy is greater than five years. For patients with neither acute nor chronic retention, surgery is indicated if symptoms are severe, or if these are moderate and life expectancy is greater than five years. For chronic or acute retention surgery is inappropriate if symptoms are mild and life expectancy is less than one year, or if there is no retention and only mild symptoms. An "appropriateness score" was developed. This confirmed that in general the ratings were internally consistent, that the panel attached little weight to mild symptoms, that a combination of irritative and obstructive symptoms was no more indicative of surgery than obstructive symptoms alone, and that the type of symptom was less important than the other factors considered. **Conclusions:** The results provide a basis for population based surveys of the need for prostatectomy.


**Topic(s): coronary artery bypass graft, percutaneous transluminal coronary angioplasty**


**Topic(s): manipulation and mobilization of the cervical spine**

An overview of the RAND appropriateness method is provided along with a summary of key results of the recent multidisciplinary expert consensus panel assessing the appropriateness of manipulation and mobilization in the cervical spine. Appropriate use
of such results is discussed. The RAND appropriateness method was used which involved a synthesis of literature, creation of a clinical indications list based on the literature, and input from informed sources. An expert multidisciplinary panel of clinicians was selected and a Delphi rating round for appropriateness of cervical spine manipulation and mobilization was performed. Results were tabulated and the panel was convened for second round of appropriateness ratings. Over 1400 clinical scenarios (indications) were rated. The panel demonstrated clear agreement on 40% of the indications and clear disagreement on 2% of them. Regarding the appropriateness of cervical manipulation or mobilization for the indications 43% of the indications were rated inappropriate for the intervention with 41% ranking as uncertain and 16% considered appropriate. The level of panel disagreement was higher with manipulation compared to mobilization. The utility of the results of this appropriateness panel is discussed relative to practicing chiropractors, non-physicians, and the chiropractic profession as a whole.


**Topic(s):** hospitalization of adolescents

**Objective:** Rates of psychiatric hospitalization and lengths of stay for adolescents have been a focus of recent controversy. With the advent of managed care, hospital systems and third-party payers are looking for ways to decide when hospitalization is indicated. The authors sought to determine whether experts could agree on the appropriateness of putative indicators for hospitalization of adolescents for conduct disorder or substance abuse. **Method:** Using a methodology developed at the RAND Corporation and previously applied to procedures in medicine and surgery, the authors developed a list of possible indications for hospitalization of adolescents with conduct disorder and/or substance abuse. A nine-member panel of experts in these areas, balanced by geography, academics/clinical practice, and whether the expert was in charge of a hospital unit, then rated the appropriateness of each indication twice under a modified Delphi procedure. **Results:** Using prespecified definitions for agreement, after the initial rating the panel had low levels of disagreement (11%) and moderate levels of agreement (28%) on more than 1,900 possible indications for hospitalization. Despite an expanded number of indications, the panel reduced disagreement to less than 5% and increased agreement to more than 55% after the second round of ratings. **Conclusions:** The consensus achieved compared favorably with the results of similar panels judging the appropriateness of procedures in medicine and surgery. The methodology is applicable to studies of the appropriateness of pharmacological or psychotherapeutic interventions in both child and adult psychiatry. The results of such studies can form the basis for rational utilization review and treatment authorization decisions.


**Topic(s):** benign hypertrophy of the prostate

This document describes the standards of appropriate use of surgery in Benign Hypertrophy of the Prostate (HBP). These standards are a result of the “method of
appropriate use”. This method combines the scientific evidence with the opinion of a panel of experts. The purpose of the method is to establish in which specific clinical circumstances the performance of surgical procedure may be appropriate, inappropriate, or uncertain. This is the first time this methodology has been used in Spain. The research team made a review of literature and prepared the list of clinical situations (1344). The panel was formed by 11 nationally prestigious urologists. The method consists of two rounds, in the first one each panelist had to score individually the appropriate indications of transurethral surgery of HBP on a scale of 1 to 9. In the second one, they meet for a day to discuss the scores and the indications, and they re-score the indications. **Results:** After the meeting the number of indications decreased to 588. The results of scores reached a level of agreement of 65%. The rating of indications of surgery of HBP based on the scores of the second round were: 184 appropriate, 245 uncertain and 159 inappropriate. **Recommendations/conclusions:** These standards may be used for the preparation of guidelines for clinical practice, for knowing the percentage with which in practice this procedure has been made for appropriate, doubtful, or inappropriate reasons, and may also be used to determine the needs of the procedure.


**B.2.29** Hagenfeldt, K., and B. Brorsson, “När är det lämpligt att göra hysterektomi?: RAM-metoden en hjälp att hitta rätt behandlingsalternativ” (“When is it Justified

**Topic(s):** hysterectomy


**Topic(s):** hysterectomy


**Topic(s):** hysterectomy

This document reports the methods and results of an expert panel process convened to develop ratings of appropriateness of hysterectomy for non-emergency, non-oncology indications. The ratings were developed as part of a project to evaluate alternative methods for improving clinical decision-making regarding hysterectomy. The indications and ratings reflect the findings of a nine-member panel of physicians who rated the indications twice. The initial ratings were made individually and without group discussion. The second round of ratings were made during a two-day meeting, within the framework of a modified Delphi methods that is often used to bring groups closer to agreement.


**Topic(s):** preoperative use of recombinant erythropoietin

**Background:** Blood transfusions have risks. Preoperative use of recombinant erythropoietin (rHuEPO) can potentially decrease the need for transfusion but has costs and risks associated with it. **Objective:** To develop indications for the preoperative use of rHuEPO alone and in conjunction with preoperative autologous donation (PAD). **Outcomes:** Each indication was rated under three separate contexts; clinical risk vs. benefit, cost constraint, cost and blood supply constraint. **Panelists:** Nine physicians representing multiple clinical specialties, practice environments and geographic locations. **Evidence:** Evidence tables and literature summary (using a computerized bibliographic search of the Medline database from January 1985 - August 1996) provided to the panelists. **Consensus process:** Two round modified Delphi-consensus method (RAND Appropriateness Method). We developed 264 indications by permuting six clinical
factors (age, history of transfusion incompatibility, hemoglobin level, expected blood loss, anemia of chronic disease, presence of cardiovascular or cardiopulmonary disease, patient anxiety). Panelists rated each indication on a 9-point appropriateness scale. Median scores and measures of agreement were determined. **Statistical analysis:** Paired comparison t-tests for equality of ratings were used to judge the significance of cost constraints or cost and blood supply constraints. Analysis of variance was performed to determine the impact of each clinical factor on the ratings. **Results:** 54% of indications were rated appropriate, 18% uncertain, and 28% inappropriate. Expected blood loss had the greatest impact on the ratings (high expected blood loss had a 5.9 point more appropriate rating (on a 9-point scale) than low expected blood loss, p<0.0001). Preoperative hemoglobin also significantly influenced the ratings (ratings for patients with low Hgb had scores 4.4 points more appropriate than ratings for patients with high Hgb, p<0.0001). Compared to the clinical context, the ratings under the cost constraints were 1.0 less appropriate for rHuEPO alone (1.2 less appropriate for rHuEPO and PAD, p<0.0001). The ratings for patients with moderate expected blood loss were significantly influenced by the cost constraint (less appropriate). **Conclusions:** Expected blood loss and preoperative hemoglobin were the best indicators of rHuEPO appropriateness. Different contexts modify the appropriateness ratings of an expensive drug like rHuEPO.

**B.2.33 Members of the Swiss Society of Gynecology and Obstetrics, Swiss Consensus Guidelines for Hysterectomy, draft, 1997.**

**Topic(s): hysterectomy**

The quality of the indication of hysterectomy is widely discussed at present. In early 1996 the committee for quality assurance of the Swiss Society of Gynecology and Obstetrics decided to set up nationally accepted guidelines for the indication of hysterectomy. A modified Delphi approach was used. In a first step general guidelines and actions prior to hysterectomy were defined. An expert panel rated 74 frequent indications twice for appropriateness (more benefits than risks for the patient), once for necessity (n = 34; procedure has to be offered or discussed with the patient) and discussed the guidelines prior to the procedure. In the homerating before the first panel there was an agreement of 48%. In 45% we observed neither agreement nor disagreement and in 7% we found disagreement among the experts. After the panel discussion an agreement percentage of 89% was noted. In 11% we observed indetermination agreement, but no disagreement among the panelists. The necessity ratings showed agreement in 68% and indetermination in 32%. The average median rating for appropriateness on a 1-9 point scale (1 = extremely inappropriate, 9 = extremely appropriate or necessary) was 5.4 over all single indications and 7.8 for necessity. After a second panel for consensus there was no more indetermination for appropriateness and necessity and no more disagreement. The results of appropriateness and necessity consensus presented in this paper reflect the findings of a 17 member Swiss panel. This joint effort by a medical society may be a step into the direction of a peer controlled health care system.

**B.2.34 Members of the Swiss Society of Cardiology, Swiss Society of Internal Medicine, Swiss Society of Thoracic and Cardiovascular Surgery, “Results of a Swiss Consensus Conference on Coronary Angiography”, Schweizerische Medizinische Wochenschrift, Volume 127, 1997, pp. 1179-1190.**

**Topic(s): coronary angiography**
Coronary angiography is used as a diagnostic procedure to identify patients with coronary artery disease who need to undergo revascularization. The quality of its indication is widely discussed at present. The Swiss Societies of Cardiology, of Internal Medicine and of Thoracic and Cardiovascular Surgery decided to set up a consensus conference to evaluate the indications for angiography. For this a modified RAND approach (Delphi method) was used. An expert panel rated 374 indications for coronary angiography twice for appropriateness (more benefits than risks for the patient), and once for necessity (procedure has to be offered or discussed with the patient). In the panel an agreement percentage of 54% was noted. In 40% we observed neither agreement nor disagreement and in 6% we found disagreement among the panelists. The necessity ratings showed similar results: 48% agreement, 46% indetermination and 6% disagreement. The average median rating for appropriateness on a 1-9 point scale (1 = extremely inappropriate, 9 = extremely appropriate or necessary) was 6.3 over all given single indications and 7.3 for necessity. The results of appropriateness and necessity ratings presented in this paper reflect the findings of a 15-member Swiss panel.


Topic(s): coronary revascularization

Percutaneous transluminal coronary angioplasty (PTCA) and coronary artery bypass grafting (CABG) are major coronary revascularization procedures to relieve angina, prevent myocardial infarction and improve long-term survival. The quality of their indication is widely discussed at present. The Swiss Societies of Cardiology, of Internal Medicine and of Thoracic and Cardiovascular Surgery decided to set up a consensus conference to evaluate the indications for coronary revascularization. For this a modified RAND approach (Delphi method) was used. An expert panel rated 992 indications for coronary revascularization twice for appropriateness (more benefits than risks for the patient), and once for necessity (procedure has to be offered or discussed with the patient). In the panel an agreement percentage of 65% was noted. In 35% we observed neither agreement nor disagreement, and there was almost no disagreement among the panelists. The results of the necessity ratings were 48% agreement, 49% indetermination and 3% disagreement. The average median rating for appropriateness on a 1-9 point scale (1 = extremely inappropriate, 9 = extremely appropriate or necessary) was 7.7 over all given single indications and 7.2 for necessity. The results of appropriateness and necessity ratings presented in this paper reflect the findings of a 15-member Swiss panel.


Topic(s): coronary revascularization

Objective: To assess the appropriateness of indications for coronary artery bypass graft (CABG) surgery and percutaneous transluminal coronary angioplasty (PTCA). Methods: A modified Delphi group judgement process with input from a panel of six interventional cardiologists and six cardiopulmonary surgeons. There was one clinician from each of the 12 tertiary referral heart centers in The Netherlands. Main outcome measure: Ratings by
panel members, on a 1 to 9 scale, of indications presented as a choice between two treatments (CABG v medical treatment, PTCA v medical treatment, and CABG v PTCA) for 1182 model cases. Each case represented a unique combination of clinical features in terms of symptoms, medical history, and results of tests. Ratings were analyzed with respect to degree of agreement among panelists, degree of appropriateness of indications, and panel's preference for invasive or medical treatment. **Results:** The panel agreed on 58.6% and disagreed on 3.2% of the indications. The panel opted for invasive treatment in 48.2% and medical treatment in 22.8%, and had no clear preference for either method in 29.0% of the cases. When compared with medical treatment, CABG was more often rated appropriate than PTCA: 35.4% v 21.6% (P < 0.001). Panel scores depended on severity of anatomical disease. For instance, for 51.5% of the model cases with one-vessel disease not including the proximal left anterior descending artery, the panel preferred medical treatment to invasive treatment, while the latter was preferred in 18% of the cases. In cases with type C lesions, the panel frequently rated PTCA as inappropriate. Panel scores were also affected by non-clinical factors. Cardiologists and surgeons rated the procedure of their own specialty higher than the alternative invasive intervention. **Conclusions:** The panel method yields logically consistent scores of the appropriateness of indications for carrying out medical procedures. It may be an aid in formulating clinical practice guidelines.


**Introduction and objectives:** The large differences in rates of use of clinical procedures among regions, hospitals and physicians raise questions as to whether some population groups are receiving inappropriate procedures or others are not receiving necessary ones. The objective of this study is to develop criteria for the appropriate use of coronary revascularization in Spain. **Methods:** Following the RAND appropriateness method, criteria were developed for the appropriate use of coronary revascularization (percutaneous transluminal coronary angioplasty and coronary artery bypass graft) in Spain. A literature review was produced as well as a comprehensive and mutually exclusive list of 1,826 indications for coronary revascularization. A panel of 12 experts (4 interventional cardiologists, 4 non-interventional cardiologists and 4 cardiovascular surgeons) rated the appropriateness of each indication on a scale from 1 (highly inappropriate) to 9 (highly appropriate). The ratings were made twice; anonymously in the first round, and during a 2-day meeting in the second round. In accordance with the panelists’ mean rating and level of agreement, each indication was classified as appropriate, uncertain or inappropriate for revascularization, angioplasty and bypass surgery. **Results:** Criteria have been developed for the appropriate use of angioplasty and bypass surgery which can be applied to patients with coronary artery disease. The combination of clinical characteristics makes it possible to classify patients with a high degree of specificity. **Conclusions:** These criteria can be used retrospectively, to measure the proportions of appropriate use, or prospectively, as an aid to decision making in order to promote the appropriate use of coronary revascularization.

Topic(s): laminectomy

A multi-disciplinary panel of 9 specialists developed criteria of appropriateness for low-back surgery, using a standardized procedure. The appropriateness of 1000 potential surgical indications for lumbar laminectomy covering sciatica, back pain only, symptoms of spinal stenosis, spondylolisthesis, miscellaneous indications and repeat laminectomy was evaluated by panelists applying a 9-point scale (1 = extremely inappropriate, 9 = extremely appropriate). Overall, laminectomy was considered appropriate in 11% of all 1000 theoretical scenarios. Twenty-six percent of the theoretical indications were considered equivocal and inappropriate in 63%. The intrapanel agreement rate for all indications was 64%. Applying these criteria retrospectively to 196 patients undergoing surgical treatment for herniated discs in a University Hospital showed that 48% of the interventions were appropriate, 29% equivocal and 23% inappropriate. The RAND/UCLA method proved to be a feasible, useful and coherent approach to develop criteria for determining the appropriateness of laminectomy and related procedures.


Topic(s): abdominal aortic aneurysm surgery

Objective: To compare an expert panel’s global assessment of appropriateness of elective surgery for abdominal aortic aneurysms (AAA) with their assessment of the effect of surgery on the probability of 5-year mortality. Methods: Nine expert panel members rated the appropriateness of 120 scenarios for elective AAA repair on a nine-point scale, and also estimated the 5-year probability of AAA-related death and of non-AAA related death among 30-day survivors of AAA surgery and among patients with unoperated AAA. These probabilities were used to determine differences in 5-year probability of mortality of surgery vs. no surgery for each scenario. Three categories of appropriateness were defined based on these differences: inappropriate (<0%), equivocal (0-5%), and appropriate (>5%). Results: The distribution of scenarios was inappropriate (39%), equivocal (12%), and appropriate (49%) based on probability estimates and inappropriate (43%), equivocal (22%), and appropriate (36%) based on global assessment. There was poor agreement between the two methods, with a Kappa coefficient =0.28 (95% CI: 0.23 to 0.32). Although a higher proportion of scenarios were rated as appropriate using probability estimation rather than global judgment, the level of agreement among members of the panel was similar, Kappa coefficient = 0.07 (95% CI: - 0.07 to 0.72). Conclusions: Experts disagree about the appropriate indications for elective surgery for AAA. Explicit estimates used in a decision analysis may provide a better assessment of appropriate indications than the global judgment of experts. Global assessment of the appropriateness of AAA surgery based on panel members’ review of research evidence for increased survival appears to include implicitly their valuation of outcomes.

Topic(s): benign prostatic hyperplasia


Topic(s): laminectomy


Topic(s): upper and lower gastrointestinal endoscopy


Topic(s): upper and lower gastrointestinal endoscopy


Topic(s): upper and lower gastrointestinal endoscopy


Topic(s): upper and lower gastrointestinal endoscopy


Topic(s): upper and lower gastrointestinal endoscopy


Topic(s): upper and lower gastrointestinal endoscopy

Topic(s): upper and lower gastrointestinal endoscopy


Topic(s): upper and lower gastrointestinal endoscopy


Topic(s): coronary artery bypass graft, percutaneous transluminal coronary angioplasty

Background: Studies have shown that clinical speciality has a strong influence on appropriateness ratings. We examined the effect of clinical speciality on physician recommendations for the performance of coronary artery bypass graft surgery (CABG) and percutaneous transluminal coronary angioplasty (PTCA) in Spain. Methods: Following the RAND appropriateness method, a 10-member panel composed of 4 cardiovascular surgeons (CVs), 2 interventional cardiologists (ICs), and 4 non-interventional cardiologists (NICs) rated the appropriateness of 1826 hypothetical indications: 936 for coronary revascularization and 890 for preference between PTCA and CABG. For all revascularization indications and for all PTCA-CABG indications for which revascularization was rated appropriate by the panel, we calculated the mean appropriateness rating, by panellist and by speciality group, and the proportion of indications rated with preference for PTCA, with preference for CABG, and with no preference, by speciality group. Results: The ICs had a higher mean rating across all revascularization indications (7.8) than either the NICs (5.7) or the CVs (5.9). For the 509 indications rated appropriate for revascularization by the panel, the ICs preferred PTCA in 54% of indications, versus 39% for the NICs and 25% for the CVs. The CVs preferred CABG in 58% of indications, versus 32% for the NICs and 20% for the ICs. Conclusions: Appropriateness ratings varied by clinical speciality. Physicians who perform a procedure may be more aggressive in recommending its use than non-performers. Appropriateness panels should be multidisciplinary to accurately reflect the judgements of the different types of physicians involved in patient care.


Topic(s): upper and lower gastrointestinal endoscopy


Topic(s): upper and lower gastrointestinal endoscopy

Topic(s): upper and lower gastrointestinal endoscopy


Topic(s): upper and lower gastrointestinal endoscopy


Topic(s): upper and lower gastrointestinal endoscopy


Topic(s): upper and lower gastrointestinal endoscopy


Topic(s): upper and lower gastrointestinal endoscopy


Topic(s): upper and lower gastrointestinal endoscopy
B.3) Literature review


Topic(s): acute medical care for the elderly

This Note analyzes the literature regarding appropriateness of acute care provided to the elderly. The authors identified 17 articles that explicitly cited appropriate or inappropriate care (including undercare, overcare, and misuse) provided in hospital and ambulatory settings and for procedures, and 19 articles that presented data on the appropriateness of medication use in the elderly. Virtually every study found at least double-digit levels of inappropriate care. Perhaps as much as one-fifth to one-quarter of acute hospital services or procedures were felt to be used for equivocal or inappropriate reasons, and two-fifths to one-half of the medications studied were overused in outpatients. The few studies that examined underuse or misuse of services also documented the existence of these phenomena. This was especially true for the ambulatory care of chronic physical and mental conditions and concerned the use of low-cost technologies (visits, preventive services, some medications). Thus, the authors conclude that there appears to be a substantial problem in matching acute services to the needs of elderly patients. This mismatch occurs both in terms of overuse and underuse, at least for areas where research has been conducted.


Topic(s): spinal manipulation for low-back pain

This report contains a literature review on spinal manipulation treatment for low-back pain, covering the literature from 1955 to 1991, and gathering data from 76 sources, including 22 controlled trials of the use of spinal manipulation for low-back pain. Based on limited data, it is estimated that about 5 percent of the U.S. population uses chiropractors annually. Without systematic reports on the frequency of complications, anecdotal evidence suggests that the serious complications of spinal manipulation include death, paraplegia, and advancement of unrecognized coexisting medical disorders because of misdiagnosis; the rate of these occurrences is probably low. The literature on the efficacy of spinal manipulation is of uneven quality; given that caveat, support is consistent for the use of spinal manipulation as a treatment for patients with acute low-back pain and an absence of other signs of lower-limb nerve-root involvement.


Topic(s): spinal manipulation for low-back pain
Spinal manipulation as a treatment for musculoskeletal complaints has been practiced for centuries. In the last 50 years, the use of spinal manipulation has been equated with the practice of chiropractic, and, in part because of this, the use of spinal manipulation has been labeled an unorthodox treatment by the medical profession. Recent research favorable to the chiropractic treatment of patients with low-back pain along with the current emphasis on patient outcomes, has helped stimulate a re-appraisal of the role of spinal manipulation. In this article, reprinted from Annals of Internal Medicine, the authors review the scientific literature on the use, complications, and efficacy of spinal manipulation for low back pain. Techniques for spinal manipulation can be broadly categorized as one of two types: nonspecific long-lever manipulations and specific, short-lever, high-velocity spinal adjustments. Long-lever manipulations use the femur, shoulder, head, or pelvis to manipulate the spine in a nonspecific manner, whereas short-lever spinal adjustments use a specific contact point on a process of a vertebra to affect a specific vertebral joint. It is this second method that is most closely identified with chiropractic practice. This article examines the use of lumbar spine manipulation of all types to treat low-back pain, and concludes that spinal manipulation is of short-term benefit in some patients, particularly those with uncomplicated, acute low-back pain, but that data are insufficient concerning the efficacy of spinal manipulation for chronic low-back pain.


Topic(s): percutaneous transluminal coronary angioplasty


Topic(s): coronary artery bypass graft surgery


Topic(s): manipulation and mobilization of the cervical spine

*Study design:* Cervical spine manipulation and mobilization were reviewed in an analysis of the literature from 1966 to the present. **Objectives:** To assess the evidence for the efficacy and complications of cervical spine manipulation and mobilization for the treatment of neck pain and headache. **Summary of background data:** Although recent research has demonstrated the efficacy of spinal manipulation for some patients with low back pain, little is known about its efficacy for neck pain and headache. **Methods:** A structured search of four computerized bibliographic data bases was performed to identify articles on the efficacy and complications of cervical spine manual therapy. Data were summarized, and randomized controlled trials were critically appraised for study quality. The confidence profile method of meta-analysis was used to estimate the effect of spinal manipulation on patients' pain status. **Results:** Two of three randomized controlled trials showed a short-term benefit for cervical mobilization for acute neck
pain. The combination of three of the randomized controlled trials comparing spinal manipulation with other therapies for patients with subacute or chronic neck pain showed an improvement on a 100-mm visual analogue scale of pain at 3 weeks of 12.6 mm (95% confidence interval, -0.15, 25.5) for manipulation compared with muscle relaxants or usual medical care. The highest quality randomized controlled trial demonstrated that spinal manipulation provided short-term relief for patients with tension-type headache. The complication rate for cervical spine manipulation is estimated to be between 5 and 10 per 10 million manipulations. **Conclusions:** Cervical spine manipulation and mobilization probably provide at least short-term benefits for some patients with neck pain and headaches. Although the complication rate of manipulation is small, the potential for adverse outcomes must be considered because of the possibility of permanent impairment or death.

**B.3.07 Wyss, P., J. Schilling, and U. Haller, Review of Recent Literature for First Consensus Panel: Swiss Indication Catalogue Appropriateness and Necessity of Hysterectomy, University Hospital of Zurich, Dept. OB/GYN and Institute for Public Health University of Zurich, August 1996.**

**Topic(s): hysterectomy**


**Topic(s): hysterectomy**

This literature review summarizes information from 326 articles and books containing information on (1) the natural history, epidemiology, and medical treatment of non-oncology and non-emergency conditions for which hysterectomy is one treatment option; (2) the appropriateness and rate of use of hysterectomy; and (3) the effectiveness, complications, and costs associated with hysterectomy.


**Topic(s): gastrointestinal endoscopy**


**Topic(s): percutaneous transluminal coronary angioplasty, medical therapy**


**Topic(s): benign prostatic hyperplasia**

Topic(s): upper and lower gastrointestinal endoscopy
C) Within panel comparisons

C.1) Across panelist characteristics


Where information about the appropriateness of a surgical procedure is lacking, expert panels have been used to establish guidelines for medical practitioners. Such a panel was convened to assess the appropriateness of percutaneous transluminal coronary angioplasty and coronary artery bypass graft surgery in the Netherlands. The panel, consisting of interventional cardiologists and cardiothoracic surgeons, used a modified Delphi process to rate 1126 clinical indications over two rounds. This article describes the degree of change in both agreement amongst members and in the appropriateness ratings over the two rounds, and examines the internal consistency of the ratings of individual panelists. Over the two rounds, agreement increased. Although most appropriateness ratings remained unchanged, there was significant movement from equivocal ratings to determinate ratings. While individual members showed some degree of inconsistency in their scoring, the panel as a whole scored very consistently. The observed changes in appropriateness were consistent with expectations, showing that the appropriateness method is used logically and consistently by panelists.


Topic(s): abdominal aortic aneurysm surgery

**Objective:** To investigate how the composition of multispecialty physician panels is associated with both the summary ratings assigned by such panels and the agreement of such panels with the recommendations of specialty societies. **Data sources/study setting:** We examined the final ratings assigned by a nine-member multispecialty RAND Corporation physician panel regarding indications for abdominal aortic aneurysm surgery and the recommendations of a specialty society representing vascular surgeons who perform the same surgery. **Study design:** The panel was retrospectively divided into two sub-panels, one composed of the three vascular surgeons on the panel and the other composed of the six remaining physicians. We analyzed the two sub-panels’ rating patterns with respect to each other and with respect to concurrent guidelines generated by the Joint Council of the Society of Vascular Surgery and the North American Chapter of the International Society for Cardiovascular Surgery. **Principal findings:** Of the 782 indications considered by the panel appropriateness, the vascular surgeons had an average of mean ratings for appropriateness of 5.1, significantly higher than the 4.5 average of the other physicians. Across the 221 indications considered by the panel for necessity, the vascular surgeons had an average of mean necessity ratings of 6.8, significantly higher than the 5.8 average of the other physicians. The vascular surgeons’ rankings of agreement with the guidelines of the Joint Council were significantly higher than those of the physician panelists from other specialties. **Conclusions:** Statements of clinical
appropriateness and necessity produced by summarizing ratings assigned to indications by expert panel members may disguise marked underlying disagreements among well-defined groups of practitioners within these panels. In the case of abdominal aortic aneurysm management, these disagreements within the RAND panel explain the marked discrepancy between the RAND multidisciplinary panel ratings and the recommendations issued by vascular surgeon professional societies.


**Topic(s):** coronary artery bypass graft surgery, percutaneous transluminal coronary angioplasty

**Objective:** To assess the role of physician specialty and the medium by which patient data are presented on the choice of treatment for patients with cardiovascular disease.

**Design:** Prospective study.

**Setting:** 10 heart centers in the Netherlands. Within each center, the recommended treatment is determined by a team consisting of cardiologists and/or cardiovascular surgeons.

**Patients:** 3628 patients found to have significant vessel disease who were referred to heart centers as candidates for percutaneous transluminal angioplasty (PTCA) or coronary artery bypass graft surgery (CABG).

**Main outcome measures:** The proportions of patients nominated for PTCA, CABG and noninvasive (medical) therapy. **Results:** Teams consisting only of cardiologists were more likely to choose PTCA than teams consisting only of surgeons, while the reverse was true for CABG. Teams with both specialties were between the two extremes. Team determination was based at least partly on the severity of vessel disease and other patient characteristics. When these patient variables were controlled by regression, there remained a substantial effect of team composition on treatment choice. For post myocardial infarction patients, direct presentation of the case to the team by the referring cardiologist reduced somewhat the likelihood that a patient would be recommended for CABG. **Conclusions:** The patient’s characteristics notwithstanding, the likelihood that a patient receives a given treatment is partly dependent on the specialties represented in the team making the decision to treat. The medium of the case presentation may also be of influence in some groups of patients.
C.2) Across methods of assessing appropriateness


*Objective:* To compare criteria for coronary revascularization developed by the expert panel process and by decision analysis. *Method:* We reviewed the medical records of 3080 chronic stable angina patients who either underwent coronary artery bypass graft surgery (CABG) or percutaneous transluminal coronary angioplasty (PTCA) and determined the agreement between appropriateness ratings made by two expert physician panels, one from the United States and the second from The Netherlands. We also evaluated the agreement between these panels’ appropriateness ratings and a decision analytic model's effectiveness categories. *Results:* There was poor agreement between U.S. and Dutch panel appropriateness ratings for PTCA (kappa = 0.03) and slight agreement for bypass surgery (kappa = 0.18). Dutch ratings had substantial agreement with the decision analytic models effectiveness categories for both PTCA and CABG (kappa = 0.83 and 0.79, respectively) whereas there was no systematic agreement between U.S. ratings and the decision analytic model for PTCA and poor agreement for CABG (kappa = 0.00 and 0.18, respectively). *Conclusions:* Although the level of agreement between expert panels and decision analysis on when a procedure is appropriate or effective may vary by procedure and the strength of the scientific evidence, we found that Dutch physicians agree much more strongly with decision analysis than U.S. physicians.
C.3) Across countries/regions


Topic(s): coronary angiography, coronary artery bypass graft surgery

Objective: To compare the appropriateness of use of coronary angiography (CA) and bypass surgery (CABS) in the early 1980s in one Canadian hospital and several American hospitals using explicit case-based criteria. Procedure rates were much lower in Canada during this period. Subjects: Canadian subjects, 502 Manitoba residents undergoing CA at one hospital in 1981-82 were tracked forward to determine whether and when they underwent CABS. CA comparisons were made with 351 Canadian patients aged 55 years or older; 1677 American CA patients aged 65 years or older and were drawn in 1981 from three hospitals. Two hundred and forty-five patients undergoing CABS within 12 months of CA made up the Canadian sample. American CABS patients were sampled from three hospitals during 1979-82. Manitoba patients assigned to medical treatment after CA were also appraised. Methods: Criteria were derived by an American panel of clinicians following a Delphi process; indications were rated appropriate, equivocal and inappropriate. A trained abstractor reviewed charts and assigned ratings. Results of exercise electrocardiograms were not available in the Manitoba data set, hence sensitivity analyses were performed to determine how differing proportions of positive exercise electrocardiograms might affect the Manitoba results. Results: Even assuming only 50% of treadmill tests were positive in the Manitoba sample, the proportion of inappropriate CA was lower in Manitoba than for the American hospitals: 9% versus 15 to 18%. For CABS, only one of the three comparison American hospitals approached the Manitoba hospital level of appropriateness. Conclusions: In the early 1980s, there was more appropriate use of CA and CABS in a Manitoba hospital compared with several American hospitals. Many Canadian patients undergoing CA and treated medically met American indications for appropriate use of CABS. Whether this represents underprovision of necessary care remains uncertain.


Topic(s): coronary angiography, coronary artery bypass graft surgery

Objectives: To compare the types of patients selected for coronary angiography (CA) and coronary artery bypass graft (CABG) surgery, and the appropriateness of the procedures performed on these patients in a random sample of cases in British Columbia and Ontario. Design: Retrospective randomized medical record review. Setting: All hospitals performing CA and/or CABG in British Columbia and Ontario in fiscal year 1989/90. Patients: For CA, 395 randomly selected patients in Ontario and 139 randomly selected patients in British Columbia; for CABG, 431 randomly selected patients in Ontario and 125 randomly selected patients in British Columbia. Main outcome measures: Case selection was measured in terms of the demographic and clinical characteristics of patients undergoing the procedures. Appropriateness was measured by
comparing the clinical characteristics of patients undergoing the procedures with explicit criteria established by a panel of Canadian physicians. The yield from CA was measured as the proportion of patients who were found to have insignificant anatomical disease.

**Results:** Analysis of patients selected for CA showed that sample patients from Ontario were less likely than those from British Columbia to be female (25% versus 37%, respectively, P = 0.012) and less likely to have undergone a previous revascularization (12% versus 24%, respectively, P = 0.005). The distribution of main indications for CA differed between the two provinces (P = 0.002), with Ontario patients more likely to have chronic stable angina (45% versus 24%) and less likely to have unstable angina (16% versus 26%). For CABG, sample patients from Ontario were less likely to be 65 years of age or older (32% versus 45%, P = 0.016) and more likely to have an ejection fraction less than 35% (14% versus 5%, P = 0.006). The distribution of the main indications for CABG differed (P < 0.001), with Ontario patients more likely to have chronic stable angina (68% versus 38%) and less likely to have unstable angina (20% versus 43%). There was no statistically significant difference in CA cases rated as inappropriate (8.4% in Ontario versus 10.8% in British Columbia, P = 0.396) or CABG cases rated as inappropriate (3.9% in Ontario versus 2.4% in British Columbia, P = 0.393). There were no statistically significant differences in the proportion of CA that yielded insignificant anatomical disease (17.5% in Ontario versus 18.4% in British Columbia, P = 0.355).

**Conclusions:** There were differences between Ontario and British Columbia in the demographic and clinical characteristics of patients selected for CA and CABG. This may indicate differences in the referral process in the two provinces. Despite these differences the rates of inappropriate procedures and the yield from CA were similar.
D) Multiple panel comparisons

D.1) Across panelist characteristics


Topic(s): cholecystectomy

A consensus development approach was used to assess the extent to which doctors in the UK agreed about the appropriate indications for cholecystectomy. Two panels, one composed entirely of surgeons and one containing a mix of relevant specialists, were asked to rate a series of possible indications. A consensus was achieved for 61% (surgical panel) and 67% (mixed panel) of indications considered. The surgical panel considered more indications as being appropriate for cholecystectomy (29% versus 13%) and fewer indications as being inappropriate (27% versus 50%) than the mixed panel. From between one third and a half of all indications, the panels were unable to reach agreement, partly as a results of differences in views as to the role of endoscopic sphincterotomy.


Increasingly, expert panels are being used to determine whether or not a consensus exists about criteria of good practice. It is, however, unclear how sensitive the panels’ conclusion are to changes in the definitions of agreement and disagreement used. To explore this, two expert panels were established to assess the appropriate indications for cholecystectomy. Analyses of the results showed that the level of agreement depended on whether or not the views of outliers were included or eliminated. Exclusion of outliers increased the proportion of appropriate indications from about 40 per cent to 60 per cent. In contrast, the proportion of indications felt to be inappropriate was dependent on how strict the definitions employed were. Given that the principal purpose of expert panels is to inform quality assurance activities, the higher levels of agreement and disagreement achieved by eliminating outliers and employing more relaxed definitions are to be favored.


Topic(s): carotid endarterectomy

The current interest in the development of practice guidelines raises an important question about the effect of expert panel composition on the outcome of the guideline development process. We compared the ratings of appropriateness of indications for carotid endarterectomy produced by two panels: an all-surgical panel and a “balanced” panel composed of four surgeons, two neurologists, and one specialist each from family practice, internal medicine, and radiology. We then compared the effect of the two sets of ratings when used to evaluate 1302 patients who had undergone carotid endarterectomy. The all-surgical panel found more indications "appropriate" (24 versus 14%) and fewer
indications "inappropriate" (61 versus 70%) than the balanced panel (p less than 0.01). The all-surgical panel also more often reached agreement. When ratings were applied to patients, 70% were appropriate by the all-surgical criteria versus 38% by the balanced panel, while 19% of the operations were inappropriate by all-surgical criteria, versus 31% by the balanced panel ratings. However, the percentage of procedures judged "inappropriate with agreement" was 15% for all-surgical and 16% for the balanced panel. We conclude that the all-surgical panel was more likely to favor operative treatment for carotid disease than the multispecialty panel, but that consensus judgments of inappropriateness by the two panels were very similar.


**Topic(s):** cholecystectomy

Cholecystectomy is performed frequently and for relatively few indications. An all surgical panel agreed on more indications and fewer contra-indications for cholecystectomy than a mixed specialty panel but did not approve indications for patients with higher comorbidity. In mixed panels subspecialists (gastroenterologists) were more conservative than generalists and surgeons. Similar findings have been shown for carotid endarterectomy. Comparing the results of British and Israeli panels showed variations in rating appropriateness that indicate differences in approach between countries. We conclude that the composition of panels will influence the assessment of appropriateness.


**Topic(s):** coronary angiography, coronary revascularization

**Objectives:** To summarize the process and extent of interphysician agreement within two panels convened to derive indications for the appropriate use of coronary angiography and for coronary revascularization procedures. **Participants:** Two panels, each with nine practitioners. **Methods:** Panelists rated the appropriateness of intervention for a comprehensive set of indications for each procedure. Indications were brief profiles created by combining and permuting clinical characteristics pertinent to case selection for intervention. Ratings were first made at home, with a second round at the panel meeting following open discussion. Final rankings of indications as 'appropriate', 'uncertain' or 'inappropriate' were based on the pattern of panelists' responses on a nine-point scale, including the median rating and extent of agreement among panelists. Agreement was defined as at least seven panelists' ratings within the three-point region containing the median rating. Panelists were later mailed a much-reduced list of indications for which there was agreement on appropriateness. These were re-rated on a necessity scale. A procedure was rated 'necessary' only if a physician was ethically obligated to recommend it as the preferred treatment option. **Results:** For appropriateness of angiography, agreement occurred in 38.2% of indications in round 1 and 64.4% in round 2 (P < 0.0001). For coronary artery bypass graft (CABG) versus medical therapy, the
corresponding increase was from 43.5 to 54.0% (P < 0.0001). Agreement on necessity of angiography occurred for 44.3% of scenarios. For indications where CABG alone was appropriate, agreement on necessity was 56%. However, for indications where percutaneous transluminal coronary angioplasty (PTCA) could be regarded as the first-line intervention, agreement on necessity was only 5%. **Conclusions:** A two-step panel process permitted considerable convergence of panelists' ratings, highlighting the importance of formal panel methods in setting utilization management criteria. However, the extent of continuing disagreement on ratings underscores the need to avoid a forced consensus; instead, divergent opinions should be taken as indicative of uncertainty about the appropriateness of intervention. Interpanelist agreement on necessity ratings was modest, but may help in setting benchmarks to assess possible under provision of invasive cardiac services in Canada.


**Topic(s):** spinal manipulation for low-back pain

**Objective:** The objective of the study was to examine the appropriateness ratings for the use of spinal manipulation for low back pain of a multidisciplinary panel of medical and chiropractic physicians, and those of a panel composed only of chiropractic physicians. **Data sources:** The study analyzed data from two consensus panels conducted at RAND in 1990 and 1991. **Study design:** The study design followed that of the traditional RAND consensus panels. Nine individuals comprised each panel, and each panelist was asked to rate, on a nine-point scale, the indications for spinal manipulation twice, the first time alone and the second time jointly with the panel. **Data collection:** The ratings of the panelists from both groups, for both round one and round two, were collated and compared. **Principal findings:** While both panels were more likely to rate the indications as inappropriate than appropriate, the single disciplinary panel was more likely to rate an indication as appropriate than the multidisciplinary panel. **Conclusion:** The composition of a panel clearly influences the ratings and those who use a given procedure in practice, in this case manipulation, are more likely to rate it as appropriate than those who do not use the procedure.


**Topic(s):** chiropractic spinal manipulation for low-back pain

**Objective:** Spinal manipulation is an efficacious therapy for some patients with low back pain (LBP). In this pilot study, we tested the feasibility of assessing the appropriateness of chiropractic spinal manipulation for patients with LBP. **Methods:** Criteria for the appropriate and inappropriate use of spinal manipulation for low back pain were developed using the RAND/UCLA appropriateness method. Two separate expert panels, one multidisciplinary and one all chiropractic, each rated a comprehensive array of clinical scenarios for appropriateness. A random sample of practicing chiropractors was
selected, and data were collected from ten randomly selected office records from each participating clinician. Assessment of the appropriateness for the use of spinal manipulation was made by comparing the care delivered with the appropriateness criteria determined by each expert panel. **Results:** Eight of thirteen (62%) eligible chiropractors agreed to participate. For the remainder, by the multidisciplinary panel's criteria, 38% of care was appropriate and 26% of care was inappropriate. By the all-chiropractic panel's criteria, the same cases were judged 74% appropriate and 7% inappropriate. The two panel's appropriateness ratings were in agreement on 48% of all cases. **Conclusions:** In this geographic area, the rate of appropriate care is between 38% and 74% and the rate of inappropriate care is between 7% and 19%, depending on the criteria used to assess appropriateness. Data from other geographic areas of the U.S. will be needed before inferences to a larger population may be drawn, and we have demonstrated that such a study is feasible.


**Topic(s):** abdominal aortic aneurysm surgery, carotid endarterectomy, cataract surgery, coronary angiography, coronary artery bypass graft surgery, percutaneous transluminal coronary angioplasty

The authors compare the appropriateness ratings and mutual influence of panelists from different specialties rating a comprehensive set of indications for six surgical procedures. Nine-member panels rated each procedure: abdominal aortic aneurysm surgery, carotid endarterectomy, cataract surgery, coronary angiography, and coronary artery bypass graft surgery/percutaneous transluminal coronary angioplasty (common panel). Panelists individually rated the appropriateness of indications at home and then discussed and re-rated the indications during a 2-day meeting. Subsequently, they rated the necessity of those indications scored by the group as appropriate. There were 45 panelists, including specialists (either performers of the procedure or members of a related specialty) and primary care providers, all drawn from nominations by their respective specialty societies. Main outcome measures included: individual panelists' mean ratings over all indications, mean change and conformity scores between rounds of ratings, and the percentage of audited actual procedures rated appropriate or necessary. Performers had the highest mean ratings, followed by physicians in related specialties, trailed by primary care providers. More than one fifth of all actual procedures were for indications rated appropriate by performers and less than appropriate by primary care providers. At the panel meetings, primary care providers and related specialists showed no greater tendency to be influenced by other panelists than did performers. Multispecialty panels provide more divergent viewpoints than panels composed entirely of performers. This divergence means that fewer actual procedures are deemed performed for appropriate or necessary indications.

Topic(s): limited surgery in breast cancer

This study was aimed at assessing the extent to which different panels of physicians produce similar recommendations and whether their agreement is greater when members of panels practice in the same geographic area. A total of 34 physicians, organized in three multispecialty and one monospecialty (all surgeons) panels, working at both specialized and community hospitals, participated to the study. They were asked to rate the appropriateness of indication for limited surgery in breast cancer for 432 hypothetical patients described through clinical scenarios. The primary outcome of the study was the reliability of appropriateness ratings among pairs of panels, measured as the percent absolute agreement and Kappa statistic. Agreement of the ratings between pairs of panel of the same region was always quantitatively good – Panel I – II: Kappa = 0.58 (95%CI=0.50-0.67) and Panel III-IV: Kappa=0.65 (95%CI=0.56-0.75) and higher than when panels from different regions were compared. We conclude that in our study an acceptable level of agreement was reached when different panels of experts were asked to produce guidelines using a structured process which includes exposure to the relevant scientific literature. The fact that local (organizational, cultural, etc.) factors seemed to play a modulating role over scientific literature has implications that should be considered when deciding the level (local vs. central) where guidelines are produced.
D.2) Across methods of assessing appropriateness


**Topic(s):** coronary angiography

**Background:** Evaluations of the appropriateness of medical care are important to monitor the quality of care and to contain costs and enhance safety by reducing inappropriate care. Experts' views are usually incorporated into evaluations of appropriateness. However, practicing physicians may not concur with these views, and physicians' clinical backgrounds may influence their beliefs. **Methods:** We asked 1058 internists, family practitioners, and cardiologists in California, Florida, New York, Pennsylvania, and Texas to rate the appropriateness of coronary angiography after acute myocardial infarction for 20 common indications. Nine clinical experts also rated these indications using an established consensus method. **Results:** For 17 of the 20 indications, median ratings of surveyed physicians and the expert panel agreed within 1 unit on a 9-unit scale. Patients' older age had a negative effect on ratings by the expert panel but not on ratings by surveyed physicians. In multivariable analyses of surveyed physicians, cardiologists rated angiography as significantly more appropriate than did primary care physicians for complicated indications, and for uncomplicated indications cardiologists who performed invasive procedures gave higher appropriateness ratings for angiography than did cardiologists who did not perform such procedures and primary care physicians. For uncomplicated indications, physicians from hospitals providing coronary angioplasty and bypass surgery rated angiography as more appropriate than physicians from other hospitals. Physicians from New York and those employed by health maintenance organizations rated angiography as less appropriate than other physicians. **Conclusions:** Surveyed physicians agreed with clinical experts about the appropriateness of coronary angiography after myocardial infarction for most indications, indicating that well-designed expert panels can closely reflect the views of practicing physicians. Variations in beliefs among practicing physicians suggest that evaluations of medical practice should incorporate the views of a range of relevant types of physicians.


**Topic(s):** upper gastrointestinal endoscopy

Methods that combine information in the medical literature with expert clinical judgment are needed to determine the appropriateness of use of a procedure. The purpose of this study is to better understand the reliability and construct validity of this process by comparing ratings of appropriateness for diagnostic upper gastrointestinal endoscopy that were developed using different approaches by two independent groups. Both the RAND/UCLA Health Services Utilization Study (HSUS) and the American Society for Gastrointestinal Endoscopy (ASGE) combined scientific data with expert physician judgment to rate the appropriateness of specific clinical indications for the use of upper gastrointestinal endoscopy. This study applies the ratings developed by each group to a
nationally representative sample of 1,585 endoscopies performed on people 65 years of age and older in 1981. HSUS developed indications that could be used to rate all 1,585 procedures; ASGE indications were less comprehensive and applied to 70% (n = 1,115) of procedures. Of those rated by both groups, appropriateness category ratings agreed for 94% of the procedures. However, the procedures not rated by ASGE were unevenly distributed across HSUS appropriateness ratings. Twelve percent of procedures rated as appropriate by HSUS were not rated by ASGE, but 80% of procedures rated as equivocal by HSUS and 73% rated as inappropriate by HSUS were not rated by ASGE, for those procedures rated by both approaches there was good agreement; however, a more explicit and comprehensive method may be required if equivocal and inappropriate use of a procedure is to be identified.
D.3) Across countries/regions


**Topic(s):** coronary angiography, coronary artery bypass graft surgery

Two panels of doctors, one in the USA and one in the UK, were asked to indicate how appropriate they judged a series of possible indications for coronary angiography and coronary artery bypass graft (CABG) operations. "Appropriateness" was defined with respect to possible benefit to the patient and excluded considerations of cost. The indications were presented as a series of detailed clinical situations in which the procedure might be used, and for each indication individual panel members rated appropriateness on a scale of 1 to 9. The US panel judged more indications appropriate than did the UK panel, and there was more agreement among the members of the US panel than among those of the UK panel. Although the two panels tended to rate the appropriateness of the indications in the same order, the UK panel placed more emphasis than did the US panel on the importance of symptoms and the amount of medical treatment. Application of the panels' ratings to two groups of patients who had had coronary angiography showed that 17% and 27% of the investigations had been inappropriate by the standards of the US panel, whereas 42% and 60% were inappropriate by the UK panel ratings. 13% of the CABG operations studied were inappropriate by the US and 35% by the UK panel ratings.


**Topic(s):** coronary angiography, coronary artery bypass graft surgery

**Objective:** To compare the appropriateness of coronary angiography and coronary artery bypass graft (CABG) use between the United States and Canada. **Design:** Retrospective randomized medical record review. **Setting:** All hospitals performing coronary angiography and/or CABG surgery in two Canadian provinces (Ontario and British Columbia); in New York State, 15 randomly selected hospitals that provide coronary angiography and 15 randomly selected hospitals that provide CABG surgery. **Patients:** All patients were randomly selected. For coronary angiography, 533 patients in Canada and 1333 patients in New York were selected; for CABG, 556 patients in Canada and 1336 patients in New York were selected. **Main outcome measures:** Percentage of patients in each country who had coronary angiography or CABG for necessary, appropriate, uncertain, or inappropriate indications as rated by criteria developed separately in each country and the complications of those procedures. **Results:** For coronary angiography, 9% of Canadian cases and 10% of New York cases were rated inappropriate using Canadian criteria compared with 5% and 4%, respectively, using US criteria. For CABG, 4% of Canadian cases and 6% of New York cases were rated
inappropriate by Canadian criteria compared with 3% and 2%, respectively, using US criteria. A lower proportion of procedures were performed on persons aged 75 years or older in Canada than in New York for both coronary angiography (5% vs. 11%; P < .001) and CABG (6% vs. 14%; P < .001). Women were also represented in lower proportions among angiography cases in Canada than in New York (28% vs. 35%; P = .023). Canadian patients with left main coronary disease waited significantly longer between angiography and CABG than did New York patients (P < .0001).

Conclusions: Rates of inappropriate use of cardiac procedures were low in Canada and New York, which suggests that the regionalization of cardiac procedures that characterizes both health care systems contributes to better clinical decision making. Differences in the use of cardiac procedures among the elderly in the two countries merits further comparative examination.


Topic(s): upper gastrointestinal endoscopy

Objective: Examine the reproducibility of the RAND method for developing criteria for the appropriateness of medical procedures. Design: Comparison of two sets of explicit criteria for appropriateness of upper gastrointestinal (UGI) endoscopy, developed by separate expert panels from two countries. Setting: United States, Switzerland. Study participants: National experts from different medical specialties involved in the referral or application of UGI endoscopy. Interventions: Each panel was presented with about 500 clinical scenarios (indications) that were rated on a nine-point scale as to the appropriateness of performing UGI endoscopy for a patient with that clinical presentation. Main outcome measures: (1) distribution of appropriateness ratings and intrapanel agreement categories between the two panels, (2) between-panel agreement of assigning appropriateness for comparable indications and, (3) percentage of indications with major between-panel differences. Results: Ratings for 2/3 of indications could be compared. The Swiss panel showed higher intrapanel agreement (54.6% versus 46.2%, P = 0.002). Seventy-eight per cent of comparable indications were assigned to identical categories of appropriateness by both panels (kappa = 0.76, P < 0.001). For 93% of the 376 comparable indications, there were no major interpanel differences. Conclusion: Separate expert panels in different countries, using a standardized methodology, produce criteria for appropriateness of medical procedures that are similar. Given the resources being invested throughout the world in developing criteria and guidelines, international collaboration in seeking optimal use of limited health care resources should be intensified.


Topic(s): upper gastrointestinal endoscopy
This study examined the reproducibility of the RAND-UCLA multi-disciplinary panel method to determine appropriateness of indications for medical care. It compares ratings of appropriateness when two panels - one Swiss, the other American - rated the appropriateness of indications for upper digestive endoscopy in a large number of clinical scenarios. Inter-panel agreement for assigning appropriateness into three categories: appropriate, equivocal, or inappropriate, was 78%, and reached 92% when these criteria were applied to real patients. These results call for confirmation by other studies, but they lend further credibility to the use of this method to determine the appropriateness of medical care.


Topic(s): colonoscopy

Background: This study examined the reliability of explicit guidelines developed using the RAND-UCLA appropriateness method. Methods: The appropriateness of over 400 indications was rated by two multispecialty expert panels (United States and Switzerland). A nine-point scale was used, which was consolidated into three categories of appropriateness: appropriate, uncertain, inappropriate. The distribution of appropriateness ratings between the two panels and the intrapanel and interpanel agreement for categories of appropriateness were calculated for all possible indications. Similar statistics were calculated for a series of 577 primary care patients referred for colonoscopy in Switzerland. Results: Over 80% of all indications (348) could be directly compared. The proportions of indications classified as appropriate, uncertain, or inappropriate were 28,4%, 24,7%, 46,6% and 33,0%, 23,0%, 44,0% for the U.S. and the Swiss panels, respectively. Interpanel agreement was excellent for all the possible indications (kappa value: 0.75) and lower for actual cases (kappa value: 0.51) because of lower agreement for the most frequently encountered indications. Conclusions: Good agreement between the two sets of criteria was found, pointing to the reliability of the method. Partial disagreement, albeit frequently encountered, indications for use of colonoscopy in cases of uncomplicated lower abdominal pain or constipation.


Topic(s): low back surgery

Study design: Reliability study of guideline development. Objective: To compare criteria for low back surgery between two expert panels. Background: Reliability of expert panels for determining appropriateness of indications for surgical procedures heretofore received little attention. Methods: Two multi-disciplinary expert panels of similar composition were convened, in the United Stated (US) and in Switzerland (CH), to evaluate the appropriateness of 720 distinct clinical scenarios involving sciatica. Each indication was assigned to a category of Appropriate (A), Uncertain (U) and Inappropriate (I). The appropriateness of the 720 theoretical scenarios were compared between the two panels, and both sets of criteria were applied to a two series of actual cases. Results: Seventy-nine percent (N = 566) of the 720 theoretical indications were
assigned to identical categories of appropriateness by both panels (kappa 0.63; \( P < 0.001 \)). Only 2 of the 720 scenarios showed frank disagreement. The percentage of the 720 that were considered appropriate differed between the two panels (USA: 3%; CH: 11%, \( P < 0.001 \)), as did the percentage of intra-panel agreement for indications (USA: 51%, CH: 64%, \( P < 0.001 \)). When theoretical scenarios were matched with two series of actual cases agreement was moderate (kappa 0.46) to fair (kappa 0.30). **Conclusion:** There is substantial agreement on the appropriateness of surgery for theoretical cases of sciata.
E) Non-RAM studies related to appropriateness


Quality has become a vital component of health care decision making by consumers and third-party payers. Clinical indicators linked to patient outcomes are powerful tools to assess quality. When results of assessments are fed back to consumers and payers, they make better-informed decisions about the health care resources they use. This article discusses many ways of measuring and assessing clinical indicators to meet both internal and external needs.


The causes of geographic variations in the use of health care services continue to puzzle researchers. Some have proposed that physicians in geographic areas with high rates of use provide proportionally more unnecessary care than those in other areas. Available research does not support this hypothesis. Others contend that uncertainty about the effectiveness of health services leads physicians to differing conclusions about when to perform various services and is the primary cause of geographic variations. Available research also does not support this hypothesis. This article proposes a different explanation, i.e., the enthusiasm hypothesis. Currently, research data suggest that geographic differences in the use of health care services are caused by differences in the prevalence of physicians who are enthusiasts for particular services. This analysis explores the validity of the enthusiasm hypothesis using previously published data on carotid endarterectomy.


**Objective:** To gain insight in the volume of PTCAs in the region of Eindhoven and in the appropriateness of the intervention. **Design:** Descriptive. **Setting:** Department of Cardiology in the Catharina Hospital at Eindhoven. **Methods:** The number of PTCAs was counted in Eindhoven and surrounding area. Peer review was carried out of 50 randomly chosen patients. General practitioners were asked about their opinion of the appropriateness of the intervention in their patients. **Results:** In the region of Eindhoven 91.5 PTCAs are done per 100,000 inhabitants yearly (the national average in 1993 in the Netherlands is 71), 38 in the rural and 133 in the urban subregions. Peer review (by two (intervention) cardiologists, three public health doctors) showed that in 82% of the cases the indication for PTCAs was appropriate. According to the Working Group for Interventional Cardiology 34% of the cases were classic indications, 64% were indications in development, and 2% were wrong. In the urban region there were more
indications in development than in the rural regions. The general practitioners agreed with the cardiologists about the appropriateness of the PTCAs in their patients. **Conclusion:** Although the number of PTCAs is relatively large in the Eindhoven region, the indications for the intervention were in accordance with accepted rules.


**Background:** Since the institution of open access endoscopy units there has been a considerable increase of referrals for UGI examinations. Therefore, guidelines for the appropriate use of UGI endoscopy are needed. **Methods:** The outcome of first diagnostic UGI endoscopy was prospectively assessed for several referral indications in a consecutive series of 2900 patients. Indications were judged "appropriate" when significantly (p < 0.01) associated with clinically "relevant" endoscopic findings. **Results:** The proportion of relevant disease for various indications was as follows: signs of UGI bleeding (42.2%); history of peptic ulcer (40.5%); dysphagia (31.9%), short-term (24.4%), and without therapy (20.9%). Relevant endoscopic findings were observed in 21.0% of dyspeptic patients aged 45 years or less, and in 25.3% of those older than 45 years of age. **Conclusions:** The generally approved alarm symptoms should be a reason to perform endoscopy without hesitation. Dyspeptic symptoms, despite adequate empiric treatment, as well as first dyspeptic symptoms in patients older than 45 years should also be a reason for endoscopic investigation. Our results support the strategy to treat patients younger than 45 years who have isolated dyspepsia by a limited course of antipeptic agents, provided that they are seen for re-evaluation within 4 to 6 weeks.


**Objective:** To compare waiting times for percutaneous transluminal coronary angioplasty (PTCA) and coronary artery bypass graft (CABG) surgery in New York State, the Netherlands and Sweden and to determine whether queuing adversely affects patients’ health. **Methods:** We reviewed the medical records of 4487 chronic stable angina patients who underwent PTCA or CABG in one of 15 New York State hospitals (n=1021) or were referred for PTCA or CABG to one of 10 hospitals in the Netherlands (n=1980) or to one of 7 hospitals in Sweden (n=1486). We measured the median waiting time between coronary angiography and PTCA and CABG. **Results:** The median waiting time for PTCA in New York was 13 days compared with 35 and 42 days, respectively, in the Netherlands and Sweden (p<0.001). For CABG, New York patients waited 17 days, while Dutch and Swedish patients waited 72 and 59 days, respectively (p<0.001). The Swedish and Dutch waiting list mortality rate was 0.8% for CABG candidates and 0.15% for PTCA candidates. **Conclusions:** There were large variations in waiting time for coronary revascularization among these three sites. Patients waiting for CABG were at greatest risk of experiencing an adverse event. In both the Netherlands and Sweden, the capacity to perform coronary revascularization has been expanded since this study began. Further international cooperation may identify other areas where quality of care can be improved.

**Objectives:** To examine geographical variations in rates of coronary artery bypass grafting (CABG) and percutaneous transluminal coronary angiography (PTCA) in New York State, and to examine variations in the choice between these two procedures.

**Methods:** A retrospective analysis of data from the New York registries for CABG and angioplasty was conducted. Rates were compared for 12 different regions of the state to assess geographic variations. To assess variations in the choice of procedure, frequencies of each procedure by region were compared with expected frequencies obtained by a logistic regression model that related procedure performance to various patient risk factors.

**Results:** There was more than a three-fold variation in age/sex adjusted CABG rates and more than a two-fold variation in age/sex adjusted angioplasty rates among regions. The regional percentages of patients undergoing CABG rather than PTCA ranged from 49% to 70%, and most of the disparity was not related to patient risk factors. Instead, the disparity was largely a result of differences in racial composition and the hospitalization rate for myocardial infarctions.

**Conclusions:** There is considerable regional variation in New York in the tendency to use aggressive cardiac procedures and in the choice of which procedure to use, and these differences are mostly unrelated to patient need.


**Objectives:** To determine whether men in the community with lower urinary tract symptoms sought treatment, would choose to have a prostatectomy, and the factors that might influence their decision.

**Subjects and methods:** The study was a cross-sectional population survey using interviewers in the autonomous community of Madrid and comprised 2002 men aged > or = 50 years. The main outcome measures were self-reported International Prostate Symptom Scores (IPSS), treatment-seeking behaviour and the patients’ stated preference for prostatectomy.

**Results:** The response rate among eligible subjects was 68.1%; overall, 38.2% of men sought medical advice for their lower urinary tract symptoms. Whether a man sought medical advice was related to symptom severity, ‘bothersomeness’, interference in daily activities and his perception of the future; of these, bothersomeness and interference in activities were more likely to determine whether or not a man consulted his doctor. Most men in the sample (84.9%) reported that they would choose a prostatectomy, although this value depended on whether they had had a previous prostatectomy, were younger, and on the content of the information presented. Men were more likely to report that they would accept surgery if their doctor recommended it and less likely when presented with information on the outcomes of treatment.

**Conclusion:** Many Spanish men with lower urinary tract symptoms do not seek medical advice for their symptoms, although most stated that they would accept a prostatectomy on the recommendation of their doctor. Further research should examine whether reported patient preferences correspond to actual behaviour and what is the most appropriate type of information to give to potential patients.

A judicious combination of qualitative and quantitative methods can play a valuable role in health services research. This paper reviews the main reasons for combining methods: for different stages in a project; to compensate for the shortcomings of individual methods; and for the purpose of ‘triangulation’. It examines the potential for qualitative approaches to contribute to quantitative work - by providing insights into the process of data construction, identifying relevant variables to be studied, furnishing explanations for unexpected or anomalous findings, and generating hypotheses or research questions for further investigation. Similarly, qualitative work can be enhanced by using quantitative techniques - albeit often in a modified form – in analysing data, developing sampling strategies, and amalgamating findings from separate qualitative studies. Although there is potential to develop multi-method approaches, there remains an important role for rigorous studies employing either qualitative or quantitative methods.


This paper describes the development and testing of a European version of the Appropriateness Evaluation Protocol (AEP). It stemmed from the original U.S. version and the multiple adaptations and modifications made previously and separately by researchers in European countries. The group was particularly concerned with developing a common list of reasons for inappropriate admissions and days of stay, since the principal goal was to enable an understanding of inappropriate hospital use and potential solutions within local health and social care systems. Developing a common EU-AEP included several steps. First, each national instrument was translated from the national language to English. These back translations were compared with each other and with US-AEP. A working group analyzed the content of the lists of reasons published in the literature and proposed a novel conceptual approach. On the basis of workshop discussions, a draft of a common European version was circulated to each participant for agreement. In the EU-AEP, the clinical criteria for the appropriateness of admission include 10 related to patient condition and five to clinical services. The criteria for the appropriateness of days of care include 10 covering medical services, six for life support/nursing services, and eight related to patient condition. The proposed core list of reasons of inappropriateness distinguish clearly between two concepts: a) the level of care required by the patient; and b) the reason why this level of care was not used. The first list would thus refer to the nature of resources and facilities required, while the second would focus more on the efficient organization of those resources. A validated European tool to assess inappropriate hospital admissions and hospital days of stay and their causes might be used to assess the need for resources for inpatient care as well as for outpatient care. Assessing the reasons for inadequacies might lead also to the examination of organizational questions. Finally, a common tool allows comparisons between countries concerning the frequency of inappropriate admissions and days of stay and their reasons in relation to the different organizations of health care across Europe.

Objective: To test the ability of two different clinical practice guideline formats to influence physician ordering of electrodiagnostic tests in low back pain. Design: Randomized controlled trial of the effect of practice guidelines on self-reported physician test ordering behavior in response to a series of 12 clinical vignettes. Participants: A national random sample of 900 U.S. neurologists, physical medicine physicians, and general internists. Intervention: Two different versions of a practice guideline for the use of electrodiagnostic tests (EDT) developed by the U.S. Agency for Health Care Policy and Research Low Back Problems Panel. The two guidelines were similar in content but varied in the specificity of their recommendations. Main outcome measure: The proportion of clinical vignettes for which EDT’s were ordered for appropriate and inappropriate clinical indications in each of three physician groups randomly assigned to receive vignettes alone, vignettes plus the nonspecific version of the guideline, or vignettes plus the specific version of the guideline. Results: The survey response rate was 71%. The proportion of appropriate vignettes for which EDT’s were ordered averaged 77%, 71% and 79% for the no guideline, nonspecific guideline, and specific guideline groups, respectively (p = .002). The corresponding values for the number of EDT’s ordered for inappropriate vignettes were 32%, 32%, and 26% (p = .08). Pairwise comparisons showed that physicians receiving the nonspecific guidelines ordered fewer EDT’s for appropriate clinical vignettes than physicians receiving no guidelines (p = .02), and that compared to physicians receiving nonspecific guidelines physicians receiving specific guidelines ordered significantly more EDT’s for appropriate vignettes (p = .0007) and significantly fewer EDT’s for inappropriate vignettes (p = .04). Conclusions: The specificity of a guideline may be one important attribute contributing to the effects of practice guidelines.
# Chapter 3: Summary information  
(including explanation of abbreviations)

## Column 2: Country

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAN</td>
<td>Canada</td>
</tr>
<tr>
<td>CH</td>
<td>Switzerland</td>
</tr>
<tr>
<td>DE</td>
<td>Germany</td>
</tr>
<tr>
<td>ES</td>
<td>Spain</td>
</tr>
<tr>
<td>EU</td>
<td>European Union</td>
</tr>
<tr>
<td>FR</td>
<td>France</td>
</tr>
<tr>
<td>ISR</td>
<td>Israel</td>
</tr>
<tr>
<td>IT</td>
<td>Italy</td>
</tr>
<tr>
<td>NL</td>
<td>Netherlands</td>
</tr>
<tr>
<td>SE</td>
<td>Sweden</td>
</tr>
<tr>
<td>UK</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>US</td>
<td>United States</td>
</tr>
</tbody>
</table>

## Column 3: Topic

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>Abdominal Aortic Aneurysm Surgery</td>
</tr>
<tr>
<td>ACUTE ELD</td>
<td>Acute Medical Care for the Elderly</td>
</tr>
<tr>
<td>BC</td>
<td>Breast Cancer</td>
</tr>
<tr>
<td>BPH</td>
<td>Benign Prostatic Hyperplasia</td>
</tr>
<tr>
<td>COR ANGIO</td>
<td>Coronary Angiography</td>
</tr>
<tr>
<td>CABG</td>
<td>Coronary Artery Bypass Graft Surgery</td>
</tr>
<tr>
<td>CAROTID</td>
<td>Carotid Endarterectomy</td>
</tr>
<tr>
<td>CHOL</td>
<td>Cholecystectomy</td>
</tr>
<tr>
<td>CATARACT</td>
<td>Cataract Surgery</td>
</tr>
<tr>
<td>HOSP ADUL</td>
<td>Hospitalization of Adults</td>
</tr>
<tr>
<td>HYST</td>
<td>Hysterectomy</td>
</tr>
<tr>
<td>LAMIN</td>
<td>Laminectomy</td>
</tr>
<tr>
<td>LBP</td>
<td>Low Back Pain</td>
</tr>
<tr>
<td>LDH</td>
<td>Lumbar Disc Hernia</td>
</tr>
<tr>
<td>LGIE</td>
<td>Colonoscopy</td>
</tr>
<tr>
<td>MMCS</td>
<td>Manipulation and Mobilization of the Cervical Spine</td>
</tr>
<tr>
<td>OB/GYN</td>
<td>Gynecology and Obstetrics</td>
</tr>
<tr>
<td>PROS</td>
<td>Prostatectomy</td>
</tr>
<tr>
<td>PTCA</td>
<td>Percutaneous Transluminal Coronary Angioplasty</td>
</tr>
<tr>
<td>PREOP ERYTH</td>
<td>Preoperative Use of Recombinant Erythropoietin</td>
</tr>
<tr>
<td>TYMP TUBE</td>
<td>Tympanostomy Tubes</td>
</tr>
<tr>
<td>UGIE</td>
<td>Upper Gastrointestinal Endoscopy</td>
</tr>
</tbody>
</table>

## Column 4: Panel composition

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>Do-ers (people who perform the procedure)</td>
</tr>
<tr>
<td>R</td>
<td>Related (specialists in related areas, e.g. cardiologists for bypass surgery)</td>
</tr>
<tr>
<td>P</td>
<td>Primary care (includes general and family practitioners)</td>
</tr>
<tr>
<td>O</td>
<td>Other specialties (e.g. radiologists for surgical procedures)</td>
</tr>
</tbody>
</table>
Column 5: Additional ratings

NEC = Necessity rated
URG = Urgency rated
EVID = Importance of evidence rated

Column 6: Number of chapters

x = Number of chapters used in the panel study

Column 7: Number of indications

x = Number of indications rated
Chapter 4: Bibliography ordered by topic

**Abdominal Aortic Aneurysm Surgery**


**Acute Medical Care for the Elderly**


**Benign Prostatic Hyperplasia (see also Prostatectomy)**


Breast Cancer (Limited Surgery)


Carotid Endarterectomy


**Cataract Surgery**


**Cholecystectomy**


Coronary Angiography


**Coronary Artery Bypass Graft Surgery**

See Coronary Revascularization
Coronary Revascularization (includes both Coronary Artery Bypass Graft Surgery and Percutaneous Transluminal Coronary Angioplasty)


New York State”, *Journal of the American Medical Association*, Volume 269, No. 6, 1993, pp. 753 - 760.


B.3.10 Emanuelsson, H., B. Brorsson, S.J. Bernstein, and L. Werkö, *Coronary Artery Bypass Graft Surgery, Percutaneous Transluminal Coronary Angioplasty or Medical Therapy in


E.07 Hannan, E.L., D. Kumar, and the Ischemic Heart Disease Patient Outcomes Research Team (PORT), “Geographic Variation in the Utilization and Choice of Procedures for
Gastrointestinal Endoscopies


**Gynecology and Obstetrics (except Hysterectomy)**


**Hospitalization of Adults**


**Hysterectomy**


B.3.07 Wyss, P., J. Schilling, and U. Haller, Review of Recent Literature for First Consensus Panel: Swiss Indication Catalogue Appropriateness and Necessity of Hysterectomy, University Hospital of Zurich, Dept. OB/GYN and Institute for Public Health University of Zurich, August 1996.


**Laminectomy**


**Low Back Pain**


Lumbar Disc Hernia


Manipulation and Mobilization of the Cervical Spine


**Percutaneous Transluminal Coronary Angioplasty**

See Coronary Revascularization

**Preoperative Use of Recombinant Erythropoietin**


**Prostatectomy (see also Benign Prostatic Hyperplasia)**


**Tympanostomy Tubes**

Chapter 5: Alphabetical bibliography


Ballard, D.J., “Lo sviluppo di linee guida; l’esperienza del programma americano sull’approprietà clinica e le sue implicazioni per i sistemi sanitari europei”, l’Ospedale, Volume 47, No. 9, 1994, pp. 171-184. (A.2.06)


Brook, R.H., J. Kosecoff, R.E. Park, M.R. Chassin, C.M. Winslow, and J.R. Hampton, “Diagnosis and Treatment of Coronary Disease: Comparison of Doctors’ Attitudes in the USA and the UK”, The Lancet, April 2, 1988, pp. 750-753. (D.3.01)


Chassin, M.R., “Explaining Geographic Variations: The Enthusiasm Hypothesis”, Medical Care, Volume 31, No. 5, 1993, pp. YS37-YS44. (E.02)


Clivio, L., Sceadan (Scenarios Developer and Analyzer), November 1997. (A.5.02)


Hagenfeldt, K., and B. Brorsson, “När är det lämpligt att göra hysterektomi?: RAM-metoden en hjälp att hitta rätt behandlingsalternativ” (“When is it Justified to Perform Hysterectomy? The RAM Method Helps to Find a Right Therapeutic Alternative”), Lakartidningen, Volume 93, No. 4, 1996, pp. 240-245. (B.2.29)


Hunter, D.J.W., C.M. McKee, C.F.B. Sanderson, and N.A. Black, “Appropriate Indications for Prostatectomy in the UK - Results of a Consensus Panel”, Journal of Epidemiology and Community Health, Volume 48, No. 1, 1994, pp. 58 - 64. (B.2.22)


Kelly, J.T., and S.E. Kellie, “Appropriateness of Medical Care: Findings, Strategies”, Archives of Pathology and Laboratory Medicine, Volume 114, No. 11, November 1990, pp. 1119-1121. (A.2.02)


Members of the Swiss Society of Cardiology, Swiss Society of Internal Medicine, Swiss Society of Thoracic and Cardiovascular Surgery, “Results of a Swiss Consensus Conference on Coronary Angiography”, *Schweizerische Medizinische Wochenschrift*, Volume 127, 1997, pp. 1179-1190. (B.2.34)

Members of the Swiss Society of Cardiology, Swiss Society of Internal Medicine, Swiss Society of Thoracic and Cardiovascular Surgery, “Results of a Swiss Consensus Conference on Coronary Revascularization”, *Schweizerische Medizinische Wochenschrift*, Volume 127, 1997, pp. 1191-1210. (B.2.35)


Schilling, J., Qualitätssicherung in der Medizin: Methodik zur Indikationsforschung in der Interventionellen Kardiologie und in der Gynäkologie, Zürich, 1998. (A.2.18)


Williams, A.P., Malpractice, Outcomes, and Appropriateness of Care, Santa Monica: RAND Corporation, Publication No. P-7445, 1988. (A.2.01)


Working Group for the Director of Research and Development of the NHS Management Executive, “What do we Mean by Appropriate Health Care?”, Quality in Health Care, Volume 2, 1993, pp. 117 - 123. (A.2.05)

Wyss, P., J. Schilling, and U. Haller, Review of Recent Literature for First Consensus Panel: Swiss Indication Catalogue Appropriateness and Necessity of Hysterectomy, University Hospital of Zurich, Dept. OB/GYN and Institute for Public health University of Zurich, August 1996. (B.3.07)