with supersensitive jargon-detectors and complete authority to eliminate jargon and gobbledegook from all written or spoken material directed to the lay public. The problem of controversy and what to do when authorities disagree must be presented squarely to the public, and the public must learn that medical scientists believe that certain things are currently "for sure," that they are presently in disagreement on others, and that they have at the moment little secure knowledge on still others.

The public also deserves to be told the difference between factors that directly and with certainty cause (or prevent) disease, and risk factors that do not cause disease with certainty but do increase the risk that a disease will occur or become more severe. Then the public also deserves to be fully informed about causative factors and risk factors and (i) what the chances are (such as 9 in 10, 1 in 10, 1 in 1000, 1 in 1,000,000) that continuing to take the risk will result in earlier death or disability, and (ii) what the chances are that following a prescribed program will prevent disease, prolong life, and improve the quality of life.

The statement has often been made that more than 20 million people in this country have hypertension, that half of these do not know they have it, and that half of those who know they have it are inadequately treated, often because patients elect to discontinue treatment. Because hypertension by itself may produce no symptoms if mild and of recent origin, but increases the risk that the individual will acquire coronary artery disease or stroke, the National Heart and Lung Institute has conducted a Hypertension Information and Education program for several years. Its goal is to identify the 10 million individuals who are believed to have hypertension but do not know it, and to inform all of the 20 million with hypertension of the availability of drug treatment.

Considerable criticism has been directed against the public for not seeking diagnosis or, once a diagnosis has been made, for not following a recommended drug or dietary regime. It appears that we are approaching a national debate over compulsory diagnosis, treatment, and prevention of noncommunicable diseases (to eliminate hospital costs and unemployment caused by preventable or treatable illness) versus the right of an individual to know the risk factors involved and then be free to decide whether he prefers the treatment to the risk in no treatment. The historical point of view teaches us that much of what we once knew "for sure" was later disproved and suggests that some of what we now know "for sure" will one day be proved wrong (such as, regular exercise prolongs life, cancer of the bowel is caused by the food we put in it). It also suggests that in the long run a well-educated, well-informed citizenry will more often than not make the right decison. With physician education and public education going hand in hand, issues such as that discussed above may well be solved without first repeating the mistake of the prohibition amendment (which made one medical risk factor illegal but still available at a price).

#### References and Notes


1. J. H. Knowles, *Bull. Am. Acad. Arts Sci.* **31**, 26 (1978).
2. J. H. Comroe, Jr., and R. D. Dripps, *Science* **192**, 105 (1976); *The Top Ten Clinical Advances in Cardiovascular-Pulmonary Medicine and Surgery Between 1945 and 1975: How They Came About* (Government Printing Office, Washington, D.C., 1976), two volumes. Single copies of the latter are available on request from Public Inquiries and Reports Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Md. 20014.
3. J. H. Comroe, Jr., *Retrospectroscope: Insights into Medical Discovery* (Von Gehr Press, P.O. Box 7654 Menlo Park, Calif. 94025, 1977), p. 18.
4. Mrs. John Gibbon, personal communication.
5. J. H. Comroe, Jr., *J. Med. Educ.* **52**, 3 (1977).
6. A 12-member team, headed by R. Kirschstein, recently carried out a comprehensive evaluation of the peer review system for evaluating NIH research grant applications. It was completed in January 1977, and on 8 February 1978 the NIH director issued his decisions based on the Kirschstein report. He approved 42 of their recommendations. Among these were: "Special efforts must be made by reviewers and staff to identify unique or unorthodox research. The need to be alert to unique ideas should be stressed in the orientation of new members of initial review groups and councils and in the Guide for new members. . . . The responsible officials and staffs of the various initial review groups or Advisory Councils/Boards [shall] hold annual orientation sessions for their new members in order to place peer review in perspective and to inform reviewers of their functions, duties and responsibilities. . . . The peer review system must remain alert to innovation. . . . Training curricula [shall] be developed by NIH for extramural program and review staff in order to provide orientation and to refresh and reiterate principles concerning the philosophy, objectives, and procedures for peer review. . . . *Initial Review* Groups [shall] be requested to identify applications they consider to be especially creative or innovative."
7. S. Riva-Rocci, *Gazz. Med. Ital. Torino* **47**, 981 (1896); *ibid.*, p. 1001.


# Surgical Innovation and Its Evaluation


J. P. Bunker, D. Hinkley, W. V. McDermott


Many observers have called attention to the inconsistency of controls and regulations governing the introduction of new therapies in the United States (*1, 2*). New drugs are introduced in a manner conforming to strict federal regulations that require rigorous testing in animals according to careful experimental designs, followed by carefully controlled testing in humans with appropriate protocols and follow-up observation. In contrast, new surgical operations may or may not be tested in animals, may be introduced as human therapy with or without review by a human experimentation committee and with or without a formal experimental design, and may or may not be evaluated by long-term follow-up observation.

The question is asked: Why should operations not be subjected to testing and controls that are as timely and no less rigorous than those required for drugs? In an effort to answer this question, and to suggest solutions for problems found, we have reviewed the process by which four relatively new operations were introduced and evaluated. Three of the four were subjected to randomized clinical trials (RCT's), but only after the passage of much time and many procedures, and it was apparent that earlier trials would have speeded the process of evaluation in each case. Shortcomings in the evaluation process also included lack of systematic and comprehensive collection and reporting of clinical experience. Early clinical surveillance could have facilitated the design and early implementation of RCT's when necessary. Of


Dr. Bunker is professor of anesthesia and of family, community and preventive medicine, Stanford University School of Medicine, Stanford, California 94305. Dr. Hinkley is professor of statistics, University of Minnesota, Minneapolis 55455. Dr. McDermott is David W. and David Cheever Professor of Surgery, Harvard Medical School, Boston, Massachusetts 02215.




0036-8075/78/0526-0937$01.00/0

equal or greater importance, long-term surveillance of clinical experience would have allowed continuing evaluation of new procedures after their widespread dissemination into general practice.

Two of the four operations—shunt surgery for portal hypertension and coronary artery bypass graft for occlusive coronary artery disease—are intended to relieve conditions that threaten life. The other two—small bowel bypass for morbid obesity and total hip replacement—are intended to relieve conditions that profoundly impair the quality of life. All four operations entail substantial risk of mortality and morbidity.

*Summary.* Early clinical trials, observational or randomized, hasten the prompt evaluation of new operations. Early clinical surveillance facilitates the design and implementation of randomized clinical trials when they are necessary. Of equal or greater importance, long-term surveillance of operations allows continuing evaluation when their use becomes widespread. Standards, coordination, review, and funding of the evaluation of new operations we believe should be centralized in a single national agency, for which an Institute of Health Care Assessment might be created. Implementation and regulation of the evaluation we believe should remain at the local or regional level with existing mechanisms and agencies being used, such as institutional human research committees and local health systems agencies.

If the condition to be treated presents an immediate threat to life, if the proposed treatment appears to be physiologically sound, if the treatment appears to be dramatically successful, and if no reasonable therapeutic alternative is available, its efficacy is apt to be considered self-evident. These were the circumstances under which shunt surgery for portal hypertension was introduced (or, more accurately, reintroduced) in 1945. The description is equally applicable to the circumstances under which coronary artery bypass graft for occlusive disease was introduced in 1969.

### Shunt Surgery for Portal Hypertension

In 1877 Nikolai Eck (*3*) carried out the first experimental portacaval shunts in a series of eight dogs, of which one survived a number of weeks, thus establishing that the procedure was not incompatible with life. Over subsequent decades, the Eck fistula dog was used as an experimental model and, in the early part of this century, the procedure was carried out intermittently in attempts to control hemorrhage or ascites in patients suffering from cirrhosis of the liver. As there were no long-term survivors from these initial efforts, the procedure fell into disuse until it was reintroduced into clinical surgery by Whipple in 1945 (*4*). Subsequent series of shunts of various types reported by Blakemore, Linton, Rousselot, and others (*5*) established that hemorrhage from esophageal varices could be prevented in most patients. Operative mortality was high initially, but so also was the mortality of recurrent hemorrhage with medical management. Such, indeed, was the apparent success that it was many years before the value of the procedure came into serious question.

Success was initially measured in terms of short-term survival and absence of hemorrhage, and not until 1954 (*6*) was it recognized that portal systemic shunt in man is often followed by a severe intermittent encephalopathy that was given the name "episodic stupor" or post-shunt encephalopathy, now recognized to be related to disordered amine metabolism. This serious and often lethal complication led many clinicians to doubt the value of shunt surgery, but the unquestioned protection shunts provide against hemorrhage presented an argument that seemed difficult to challenge. There was, in fact, little serious debate concerning the appropriateness of shunt surgery in patients who had already bled from esophageal varices; the debate had become, rather, whether shunt surgery should be performed as a prophylactic procedure in patients with demonstrable portal hypertension and esophageal varices who had never bled, but who might, perhaps, bleed sometime in the future. Here, there was sufficient professional uncertainty to justify careful RCT's, and they were carried out (*7*).

The results revealed no difference in survival between the patients that had been selected at random for medical therapy and those similarly selected for surgery: very few of the surgical group had subsequent hemorrhage, whereas those under medical treatment often bled to an extent requiring emergency surgery; the surgical patients, on the other hand, suffered a high incidence of progressive hepatic failure and encephalopathy with a subsequent mortality closely approximating that from hemorrhage in the medically treated group.

With the prophylactic shunt discredited, the value of therapeutic shunts (in patients who had already bled from demonstrable esophageal varices) was reexamined, and RCT's were finally deemed necessary and appropriate (*8*). Many of these trials are still in progress and some differences in the statistical evaluation exist, but it now appears that the "therapeutic" shunt provides at best slight survival value (*9*), although it unquestionably protects against hemorrhage.

As one reviews the disorderly history of the innovation of shunting procedures, it is obvious that decades were wasted by the failure to introduce, standardize, and carry out RCT's from the beginning. That such trials were finally carried out resulted from two events: one was the recognition of an iatrogenic complication, encephalopathy, potentially as serious as the original condition; the second was the effort to extend the treatment beyond the original indications to situations where genuine uncertainty existed. Prior to these events, it could be argued—and was—that RCT's were neither necessary nor ethical.

### Coronary Artery Bypass Graft

The treatment of angina pectoris and myocardial ischemia by coronary artery bypass graft was reported independently by Favaloro (*10*) and by Johnson (*11*) in 1969. All of the circumstances favoring the operation that were listed earlier obtained: the condition is life-threatening and its associated anginal pain markedly diminishes the quality of life. No reasonable therapeutic alternative was available, medical or surgical, despite repeated and imaginative attempts in the past to find an effective treatment. The new operation made good sense anatomically and physiologically. And, most important, the new operation seemed from the outset dramatically effective in the relief of disabling angina.

Under such favorable circumstances, it is not surprising that randomized studies were not carried out by the surgical innovators or by the many who followed immediately on their heels. Indeed, Favaloro is said to have expressed the opinion shortly after his initial work that it would be unethical to withhold the new operation from a control group. The initial published results seemed to support the high expectations and led rapidly to a consensus which, though wide, was by no means unanimous. Indeed, there was from the beginning a small but vocal group of physicians who expressed doubts and called for RCT's. None were forthcoming, and by 1973–1974 many be-

lieved that it was too late to carry them out.

Coronary artery bypass graft had become the established treatment for patients with severe "stable" angina and radiologically demonstrable occlusive disease involving at least two of the three major coronary arteries. Many questions remained, however. Should surgery be performed on patients with single vessel disease, patients with mild angina—or, indeed, with no angina—or those with "unstable" angina (severe, rapidly changing angina with an expected very poor prognosis)? Even such a fundamental question as whether coronary artery bypass surgery prolongs life remained unanswered. Answers to these questions were sufficiently uncertain to justify randomized trials which were indeed instituted, and from which results are now beginning to be available. Thus, for example, in the Veterans Administration Cooperative Study of Coronary Arterial Surgery (*12*), an RCT of a subset of 113 patients with angina pectoris and reduction of the luminal diameter of the left main coronary artery 50 percent or greater has been completed, with statistically significant differences being found in life expectancy that favor surgical over medical treatment. With the exception of the patients with significant left main coronary artery disease, however, there is, to date, no difference in survival between patients receiving medical treatment and those treated surgically. Relief of pain was achieved in a very high proportion of cases, however, reinforcing the widespread belief that surgery is the treatment of choice in severe, disabling angina.

Further evidence in support of coronary bypass surgery has come from decision analyses, such as those of Weinstein *et al.* (*13*). On the basis of an analysis of a variety of hypothetical clinical situations, these authors reported "that surgery was the optimal course of action for almost all patient types considered." Yet, even as evidence of the efficacy of coronary artery bypass surgery accumulates, public and professional controversy also mounts. This controversy is related at least as much to the enormous public costs as to possible medical shortcomings, with much of the concern focused on the widespread duplication of facilities, including institutions which may perform as few as one or two open heart procedures a month. Such institutions cannot be expected to have as good results as more experienced institutions (*1*) and they will certainly experience greater expenses, particularly in capital investment. Unfortunately, there are no routinely collected, comprehensive outcome data on the basis of which the relative success of all institutions can be judged. We believe that the current greatest need in the area of coronary artery bypass surgery is for surveillance of all such operations as they are carried out in greater numbers and at an increasing number of institutions that are less experienced and perhaps less well equipped.

### Jejunocolic and Jejunoileal Bypass

The 15-year experience with small intestinal bypass surgery as a treatment for morbid obesity has been very different from that of coronary artery bypass graft. The circumstances of the surgical innovation were equally favorable—the medical condition was serious, even life-threatening; no effective alternative treatment was available; and the new operation's potential for success appeared dramatic. The operation itself was quickly found to entail a substantial immediate risk to life (8 percent in the largest series reported to date) (*14*), so that a favorable long-term effect on life expectancy was by no means certain, and the justification appeared to rest largely on an improved quality of life. One clear difference between intestinal bypass for obesity and coronary artery bypass graft for angina lies in the frequency of use. While it is estimated that 70,000 coronary bypass operations were performed in 1976 in the United States, it is estimated (*15*) that 5000 jejunoileal bypass operations have been performed in the 10 years since its introduction—a smaller but still not inconsiderable volume.

In contrast to the continuing doubts and clamor for RCT's of coronary bypass graft surgery, the critical response to jejunoileal bypass has been largely favorable. This might be viewed at least in part as a question of scale. The enormous volume of papers devoted to jejunoileal bypass surgery belies any lack of interest, however. Is the operation so clearly favorable in its effects that no doubts need to be entertained? This does not appear to be the case; not only is operative mortality large, but there are a host of serious side effects (liver failure, arthritis, and a variety of moderate to severely debilitating metabolic disorders).

The explanation seems to be that no alternative treatment offers hope of relief. Candidates for jejunoileal bypass are limited to those in whom extended medical and psychiatric treatment has failed, and the seriousness of the condition is considered by patient and physician alike to justify the risks of the treatment. Yet many questions remain, among the most prominent being which operation to perform, jejunoileal bypass or the more recently introduced gastric bypass (*16*). Malt and Guggenheim (*15*) have suggested that reliable answers might require RCT's and the first report of such a trial comparing operations for morbid obesity has just been published (*17*).

At least as important as the need to carry out RCT's of alternative surgical approaches to the treatment of morbid obesity is the need for comprehensive surveillance (*18*). As with coronary artery bypass surgery, it is essential that we know the results of surgery for obesity in all hospitals, those in which such surgery is occasional, as well as those in which it has been made a specialty.

### Total Hip Replacement

For people who would like to see new operations introduced in a manner closely similar to that of drugs, the recent introduction of total hip replacement offers a nearly perfect paradigm. Orthopedic surgeons had attempted to reconstruct hips damaged by disease or injury for over a century. Some form of prosthesis, usually metal, had been used in most instances, but the perfect metallic alloy and a fully satisfactory result had not been achieved.

Mechanical loosening of the prosthesis had been the most frequent cause of failure of partial or total hip replacement. This complication has been almost entirely overcome by the use of self-curing acrylic cements to anchor both the femoral replacement and the acetabular acrylic or metal socket, a technique first described by Charnley (*19*) at the University of Manchester in 1961. For the first operations involving this technique, perhaps several hundred in number, Charnley used polytetrafluoroethylene for the acetabular cup, but these operations resulted in failure because of excessive wear of the plastic (*20*). The subsequent use of a high-density polyethylene, methylmethacrylate, corrected the deficiency.

The clinical introduction of the Charnley total hip replacement in this country in the late 1960's came under the aegis of the Food and Drug Administration (FDA), apparently somewhat by chance. A casual inquiry as to the FDA's possible interest or concern with the use of methylmethacrylate elicited the prompt institution of FDA regulations as for any other investigational new drug (IND), including a requirement to submit full particulars of proposed experimental

protocols, before the orthopedic surgeons were allowed to proceed with the new operation.

Considered at first a regrettable nuisance, the procedures required by the FDA turned out to be a blessing for all concerned. As a result of the IND process, total hip replacement was introduced in this country in an orderly way, including the requirement for complete and continually updated information. Nationwide data were available from the outset and provided a knowledge of surgical outcome which was comprehensive and almost unique. Only for organ transplants have such complete data been routinely available. For most other operations, new or old, individual institutions may collect and publish their experience, but the experience of most institutions, and therefore any opportunity to project the national experience for that particular operation, is not known.

By late 1971, clinical experiments were considered to have established the safety and efficacy of methylmethacrylate, and in October that year the FDA released it for use in total hip replacement (*21*). Institutions were thus relieved of the requirement to obtain an IND license and were no longer required to collect and submit outcome data, and efforts to maintain a national registry stopped. Most institutions have continued to evaluate their individual experience with care, but knowledge of the total national experience with this important and still relatively new operation is today little better than that of any other surgical procedure.

## Discussion

The reader will no doubt agree that a new treatment should be introduced in a manner that allows prompt and reliable evaluation of its efficacy and safety. When the new treatment is an operation, its introduction is almost always uncontrolled. Is it possible, under such uncontrolled conditions, to evaluate new operations promptly and reliably? The foregoing brief case studies of recently introduced operations, we believe, provide some answers.

From our review of portacaval shunts, it is clear that their uncontrolled use led to conflicting evidence and growing uncertainty as to beneifts that originally seemed self-evident. Unfortunately, but not atypically, when a point of uncertainty is reached, the accumulated evidence from past use of a procedure is often not adequate to resolve such uncertainty—or even clearly to recognize that it exists. Thus, some physicians questioned the efficacy of therapeutic portacaval shunts, but professional uncertainty was not considered sufficient to justify an RCT until attempts were made to extend indications to conditions where uncertainty was recognized (that is, its prophylactic use in patients who had never bled). Only when prophylactic shunts were subjected to RCT and found not to prolong life were therapeutic shunts recognized as sufficiently uncertain to require RCT's.

Increasing and unrestricted use of a new operation may simply add to the uncertainty. In the absence of an explicit and detailed experimental protocol, institutions and surgeons are apt to vary considerably in their selection of patients and choice of operative technique, as well as in their criteria for subsequent evaluation. These factors appear the most likely explanations of the wide variability in reported mortality, morbidity, and weight loss following jejunoileal bypass, a variability far exceeding that which can be explained on the basis of chance alone. The critical importance of a standardized protocol has recently been examined in detail by Cochran and coauthors (*22*).

It is apparent that reliable information documenting the benefits of a surgical procedure is often not obtained, and that a contributing cause is the uncontrolled variation in its application. What is required is a well-supervised collaborative observational study of a new procedure, a study in which documented protocols are followed and in which all relevant quantitative evidence is collected and analyzed according to predetermined statistical criteria. The ongoing statistical analysis of this observational study should have as one purpose the differentiation of outcome by prognostic factors. These should include, at the minimum, age and sex, diagnostic categories, severity of condition, and major procedure-specific factors, such as the length of jejunoileal bypass and the number and identity of occluded coronary arteries bypassed. The identification of therapeutic uncertainty justifying randomized trials, and the design of such trials, would be based on such carefully designed and analyzed observational studies. For jejunoileal bypass, if further RCT's of end-to-end and end-to-side jejunoileal and gastric bypass are to be carried out, it will be important to build points of consensus (such as agreed-upon prognostic factors) into appropriate stratification of patient groups; the treatment groups should be balanced by age and weight, ventilatory function, and psychiatric history; diet and other treatment variables should be standardized, and so on.

In focusing attention on procedures for evaluating the efficacy of a feasible new operation or the comparative efficacies of competing operations, we must not overlook the central role of innovation. Here history speaks against initially rigid control, inasmuch as such control may lead to premature evaluation. It is quite possible, for example, that today's human experimentation committees would not have condoned the more than 50 percent mortality of the first operations for mitral stenosis (*23*), the 33 percent mortality of the first portacaval shunts at the Massachusetts General Hospital (*24*), or the large failure rate in Charnley's first series of total hip replacements.

One might properly regard the initial innovation as the beginning of a feasibility study, during which the physician-investigator develops and refines the new procedure and defines diagnostic criteria for its application. Independent review of the results would, when favorable, lead to collaborative trials. In the presence of a "learning curve" phenomenon, the timing of the shift from feasibility study to multicenter trials may be crucial. For this to be accomplished, there must be a central reviewing authority capable of sophisticated statistical and economic analysis and empowered with authority and resources necessary to initiate and coordinate appropriate trials.

Both the feasibility study and subsequent clinical trials should clearly be carried out under the most favorable circumstances and concentrated in a relatively few institutions. There is a double advantage arising from such concentration: the improvement in treatment which results from greater experience, and the more reliable statistical information that can be achieved in larger series. Fragmented results from short series of trials have less direct and indirect (scientific) human benefit. This point has been elaborated in a recent British report on RCT's (*25*) that discusses many of the practical and theoretical issues involved in such trials.

Once the efficacy of a procedure is established by collaborative study, wider use of the procedure should follow. This should not, however, mean an end of evaluation. Indeed, there is a clear need for continuing observation for many procedures, including all four described above, to determine long-term results. There is, in addition, a need to determine the outcome of surgery carried out by surgeons with varying degrees of training and experience and in a variety of clinical settings. In the case of total hip re-

placement, the demonstration of efficacy led to removal of FDA controls, abandonment of the data registry, and loss of opportunities for comprehensive assessment of long-term effects of surgery. A recent editorial in *Lancet* (*26*) expresses concern for the long-term results of hip replacement in Great Britain, stating that most "orthopaedic surgeons have no time to follow patients beyond the immediate postoperative period, so they do not know"—a situation that may apply to the United States as well, at least for some institutions.

In urging that a system of comprehensive surveillance be established, we acknowledge the formidable logistic difficulties to be encountered. Clearly, it is not cost-effective to collect and collate all possibly relevant data on all patients who undergo surgery in this country. But some data can and should be collected on all patients, and such data should include not less than age, sex, operation, and discharge status. Indeed, such information is currently available from many private medical data corporations, such as the Commission on Professional and Hospital Activities in Ann Arbor, and Forrest and his colleagues (*27*) have recently reported the use of these data, with appropriate adjustments for patient mix, as the basis for assessing institutional differences in surgical mortality and severe morbidity. For greater detail of observation and discrimination, the initial observational studies, such as the hip registry, can be continued over a matter of years or indefinitely. For this purpose, the modern computer-based data bank (*28*) is versatile and appropriate. We believe that long-term surveillance of both types—broad and shallow, to detect patterns of major outcomes; narrow and deep, to allow greater discrimination of detail—is urgently needed and may indeed be as important in the overall evaluation process as how the operation is initially introduced.

## Conclusions and Recommendations

We believe that implementation of the foregoing program can be achieved by the fuller use of existing procedures and institutions. Toward this end we recommend:

1) That an "Institute of Health Care Assessment" be charged and adequately funded to provide independent evaluation of surgical procedures—including old procedures where professional uncertainty persists. These functions might best be assumed by an independent, private agency such as the Institute of Medicine or the American College of Surgeons. Alternatively, a government agency such as the National Center for Health Services Research or a new agency, such as the "Health Outcomes Commission" proposed by Ellwood *et al.* (*30*), might serve this role.

2) That new procedures be introduced and feasibility studies carried out by physician-investigators at institutions with the approval of, and under the surveillance of, the local institutional review board, in accordance with nationally formulated policies, and with funding awarded competitively from federal sources. The local agency might appropriately be an existing human experimentation committee, and the central reviewing and funding agency the designated "Institute of Health Care Assessment."

3) That after feasibility of the new procedure has been established to the satisfaction of the Institute, clinical trials (randomized or observational, as appropriate) be carried out at selected institutions, such trials to be approved and funded by the Institute on the basis of competitive applications, and again with local review and surveillance by the local institutional review board. (During this phase of evaluation, it would seem wise to prohibit funding from existing third-party insuring agencies.)

4) That after efficacy of the new procedure has been established, again to the satisfaction of the Institute, the new procedure be released for more general use, but with continuing surveillance and compilation of data in registry form, and with restriction to institutions demonstrating personnel and facilities appropriate to the complexity of the procedure. Regulation and controls over regional distribution among institutions may reasonably be assigned to local health systems agencies. Funding for this phase (comparable to phase III of drug testing) might appropriately come from the usual third-party insuring bodies. Reinforcement of regional control may be afforded by designation or withholding of reimbursement for specific procedures on an institutional basis.

5) That continuing surveillance be maintained at a level appropriate to the procedure to allow long-term evaluation, and that old or established procedures be included as well as new. Such surveillance should be carried out at the minimum level necessary. This would include not less than age- and sex-specific procedure rates and associated mortality. Such data should be reported routinely to the relevant responsible agency or agencies (for example, hospital board of trustees or regional health systems agency).

### References and Notes


1. T. C. Chalmers, *Ann. Thorac. Surg.* **14**, 323 (1972).
2. D. H. Spodick, *Am. Heart J.* **85**, 579 (1973); *J. Am. Med. Assoc.* **232**, 35 (1975); J. P. Swazey, *Daedalus* **107** (No. 2), 129 (1978). See also part 2, "Surgical Innovation and Its Evaluation," in *Costs, Risks and Benefits of Surgery*, J. P. Bunker, B. A. Barnes, F. Mosteller, Eds. (Oxford Univ. Press, New York, 1977), pp. 107–220.
3. N. V. Eck, *Voyenno-med* **130**, 1 (1877). Translation by C. G. Child, III, *Surg. Gynecol. Obstet.* **96**, 375 (1952).
4. A. O. Whipple, *Ann. Surg.* **122**, 449 (1945).
5. A. H. Blakemore, *Bull. N.Y. Acad. Med.* **22**, 254 (1946); *Surg. Gynecol. Obstet.* **94**, 443 (1952); R. R. Linton, *N. Engl. J. Med.* **238**, 723 (1948); L. M. Rousselot, *J. Am. Med. Assoc.* **140**, 282 (1949); C. G. Child, III, and M. A. Payne, *Monographs in Medicine* (Williams & Wilkins, Baltimore, 1952); C. G. Child, III, *Am. J. Gastroenterol.* **25**, 148 (1956); W. V. McDermott, H. Palazzi, G. L. Nardi, A. Mondet, *N. Engl. J. Med.* **264**, 419 (1961); L. M. Rousselot, W. F. Panke, R. F. Bono, A. H. Moreno, *Am. J. Med.* **34**, 297 (1963).
6. W. V. McDermott and R. D. Adams, *J. Clin. Invest.* **33**, 1 (1954).
7. F. C. Jackson, E. B. Perrin, A. Degradi, *et al.*, *Arch. Surg.* **91**, 43 (1965); H. O. Conn and W. W. Lindenmuth, *N. Engl. J. Med.* **279**, 725 (1968); R. H. Resnick, T. C. Chalmers, A. Ishihara, *et al.*, *Ann. Intern. Med.* **70**, 675 (1969).
8. F. C. Jackson, E. B. Perrin, W. R. Felix, A. G. Smith, *Ann. Surg.* **174**, 672 (1971); R. H. Resnick, F. L. Iber, A. M. Ishihara, T. C. Chalmers, H. Zimmerman, and the Boston Inter-Hospital Liver Group, *Gastroenterology* **67**, 843 (1974).
9. H. O. Conn, *Gastroenterology* **67**, 1065 (1974).
10. R. G. Favaloro, *J. Thorac. Cardiovasc. Surg.* **58**, 178 (1969).
11. W. D. Johnson *et al.*, *Ann. Surg.* **170**, 460 (1969).
12. M. L. Murphy, H. N. Hultgren, K. Detre, J. Thomsen, T. Takaro, and participants of the Veterans Administration Cooperative Study, *N. Engl. J. Med.* **297** (No. 12), 621 (1977); T. Takaro, H. N. Hultgren, M. J. Lipton, K. M. Detre, and participants in the study group, *Circulation* **54** (Suppl. 3), 107 (1976).
13. M. C. Weinstein, J. S. Pliskin, W. B. Stason, in *Costs, Risks, and Benefits of Surgery*, J. P. Bunker, B. A. Barnes, F. Mosteller, Eds. (Oxford Univ. Press, New York, 1977), pp. 342–371.
14. L. T. DeWind and J. H. Payne, *J. Am. Med. Assoc.* **236**(20), 2298 (1976).
15. R. A. Malt and F. G. Guggenheim, *N. Engl. J. Med.* **295**, 43 (1976).
16. K. J. Printen and E. E. Mason, *Arch. Surg.* **106**, 428 (1973).
17. W. O. Griffen, Jr., V. L. Young, C. C. Stevenson, *Ann. Surg.* **186**, 500 (1977).
18. E. E. Mason, quoted by J. F. Alden, in *Arch. Surg.* **112**, 799 (1977).
19. J. Charnley, *Lancet* **1961-I**, 1129 (1961).
20. H. C. Amstutz, in *Surgery of the Hip Joint*, R. G. Tronzo, Ed. (Lea and Febiger, Philadelphia, 1973), p. 656.
21. D. S. Burton, *The Stanford M.D.* **11**(2), 17 (1972).
22. W. G. Cochran, P. Diaconis, A. P. Donner, *et al.*, in *Costs, Risks, and Benefits of Surgery*, J. P. Bunker, B. A. Barnes, F. Mosteller, Eds. (Oxford Univ. Press, New York, 1977), pp. 176–197.
23. J. P. Swazey and R. C. Fox, in *Experimentation with Human Subjects*, P. A. Freund, Ed. (Braziller, New York, 1970), pp. 315–357.
24. R. R. Linton, I. B. Hardy, Jr., W. Volwiler, *Surg. Gynecol. Obstet.* **87**, 129 (1948).
25. R. Peto, M. C. Pike, P. Armitage, N. E. Breslow, D. R. Cox, S. V. Howard, N. Mantel, K. McPherson, J. Peto, P. G. Smith, *Br. J. Cancer* **34**, 585 (1976); *ibid.* **35**, 1 (1977).
26. Editorial, *Lancet* **1976-I**, 234 (1976).
27. Staff of the Stanford Center for Health Care Research, *Health Serv. Res.* **11** (No. 2), 112 (1976).
28. R. A. Rosati, A. G. Wallace, E. A. Stead, Jr., *Arch. Intern. Med.* **131**, 285 (1973); J. F. Fries, *N. Engl. J. Med.* **294**, 1400 (1976).
30. P. M. Ellwood, P. O'Donoghue, W. McClure, *et al.*, *Assuring the Quality of Health Care* (Interstudy, Minneapolis, 1973).


31. We wish to acknowledge with thanks the constructive advice of Alain C. Enthoven, William H. Forrest, James F. Fries, and David J. Shurman and of the many other colleagues who reviewed earlier drafts of this article.