When Is Surgery Indicated? A Book Review

EUGENE VAYDA

Department of Health Administration,
Faculty of Medicine, University of Toronto

Costs, Risks and Benefits of Surgery, edited by John P. Bunker, M.D., Benjamin A. Barnes, M.D., and Frederick Mosteller, Ph.D. (New York: Oxford University Press, 1977), $22.50

Many studies have identified wide variations in surgical rates, most marked for discretionary operations. The variations which have been found among countries and between large and small areas within countries have raised questions of unnecessary surgery and of treatment effectiveness. The costs of existing services as well as new and expensive high technology measures are being scrutinized because of the rapid escalation of health care costs and the resulting implications for national health insurance. Universal accessibility and coverage are financially feasible and politically palatable when there is some control of health care expenditures. Public funding places health care spending in the same category as other publicly funded services (such as education, housing, and environmental control). In a situation of finite resources and financial constraints, competition among and within publicly funded programs is inevitable. A health care dollar is one dollar less for education, and a dollar for surgery is one less for preventive or geriatric services.

Surgery is particularly vulnerable to question and analysis because it is discrete, quantifiable, and highly visible. A hospital admission for a medical diagnosis, say peptic ulcer disease or infectious hepatitis, is more difficult to assess than one for a surgical procedure, such as tonsillectomy or cholecystectomy. However, the questions posed regarding surgery must be applied to all medical care. Many studies have already done so. The psychological effects of coronary care units are as well documented as the emotional

effects of hysterectomy; home versus hospital care of myocardial infarction raises the same questions of efficacy and unnecessary treatment that surgery does.

Differences in discretionary surgery rates pose a number of fundamental questions. The first is the *effectiveness* of established and frequently performed operations, from tonsillectomy to radical breast surgery. The failure to determine effectiveness makes it difficult to agree on definitive indications of need for many operations. Since, in many instances, effectiveness is uncertain and indications are unsettled, the causes of differing operative rates have been attributed to treatment styles, methods of organizing and paying for services, available resources (dollars, insurance, surgical personnel, and hospitals), and patient expectations. Disease incidence is less important as a cause of the differing rates, especially for conditions of unknown significance and high prevalence (e.g., hypertrophied tonsils, asymptomatic gall stones, enlarged uteruses), or for serious diseases whose prevalence is similar in the areas under comparison (e.g., breast cancer).

Although the issues of if, when, and why to perform established operations are still unresolved, they have been joined by another problem—that of new, innovative, and highly expensive procedures. Here differing rates may be a function of accessibility of beds and surgeons, but the effectiveness of many new procedures is rarely resolved before the procedures are widely adopted. New procedures such as coronary artery bypass surgery or computerized tomography entail high technology and substantial expense. Economists estimate a projected cost of between 1.3 and 3.0 billion dollars per year to make coronary artery bypass surgery available for the adult male population of the United States who might presumably benefit from this unproven procedure. If the procedure were demonstrated to be effective, the system would either face this sum as an add-on cost or be forced to eliminate or reduce other expenditures. If the procedure proved ineffective (historically many new procedures have failed to withstand rigorous testing) the increased expenditures or reductions in other programs would be indefensible, as they would result in costs and risks but no benefits.

Costs, Risks and Benefits of Surgery, edited by John P. Bunker, Benjamin A. Barnes, and Frederick Mosteller, uses surgery as a model to consider the generic issues of effectiveness and benefits and then examines, in detail, a number of surgical procedures as

well as several medical treatments. The book, a compendium growing from a joint Seminar in Health and Medicine (between Harvard's Center for the Analysis of Health Practices and the Department of Statistics) is an important contribution. Although its methods, conclusions, and recommendations can stand alone, the long-term benefits of this collection of twenty-three critical essays may well lie in the questions raised and the identification of data and research needs to assess health care services. This book sets forth those critical agendas which must inform decisions in clinical therapy as well as those for public policy.

The initial section examines the general principles involved in cost benefit and decision analyses. These methods are described in comprehensible fashion, but, still more important, their inherent limitations are made specific. Beyond the problems of inadequate or unreliable data, the difficulties in computing social costs and assigning dollar values to the quality of life are explored. Furthermore, a number of caveats highlight the difficulties in applying economic models to medical care systems and decision making. The economic concept of rationality assumes that each individual is paying for care from his own income. Health insurance or a universal payment system would nullify the utility of such economic models.

Cost benefit and cost effectiveness analyses are important in that they raise the issue of value for money in the context of fixed or limited resources and offer an additional parameter for public policy decision making. However, the results of cost benefit analyses do not determine policy decisions. Elsewhere, Creese has said, "The economist cannot make the decisions nor eliminate the element of judgment from planning decisions. He can only widen the planners' information base, to take some account of the value-for-money implications of decisions. At its most basic level, the value of his contribution rests upon the acceptance that information is preferable to ignorance in decision making and, particularly, that economic information is important" (Creese, 1976:23)

Cost benefit studies cannot be used as substitutes for clinical judgment. However, such studies, like controlled trials of treatment efficacy, should form part of the basis on which particular operations or procedures are adopted, continued, or dismissed. Decision analysis, on the other hand, does provide a potential algorithm for clinical decision making based on multiple probabilities; as such, it could be an aid to the surgeon in his assessment of an individual

patient. The reliability of the probability determination is enhanced by the quality of the data on which it is based and the number of essential variables considered.

The second section of *Costs, Risks and Benefits of Surgery* deals with the evaluation of surgical innovation. This section begins with a chapter by Benjamin A. Barnes, which reminds us that a number of surgical procedures have been discarded over time because experience proved them to be ineffective. Included are operations for ptosis of abdominal organs, surgery for constipation, extirpation of endocrine glands and section (or removal) of peripheral or autonomic nerves. In these instances the procedures had been initially adopted because of the enthusiasm of a few surgeons, and discarded over time because the beneficial claims could not be replicated by others performing these operations.

In their contribution, "Progress in surgery and anesthesia," Gilbert, McPeek, and Mosteller analyze a number of surgical innovations, which were found to be preferable to standard treatments in about half the studies reviewed. (This chapter's weakness is that it does not assess the methods employed in the cited studies.) The authors identify the problem of "Type II errors" (i.e., erroneously failing to reject the null hypothesis) in those instances where only small numerical gains or losses occur. Careful design, appropriate beta levels, and sufficiently large sample sizes are proposed to overcome this problem.

Lin Miao's essay describes the rise and fall of gastric freezing for peptic ulcer disease, and one by Barsamian describes a similar fate for internal mammary artery ligation for angina pectoris. The rapid demise of these two procedures was made possible by the results of randomized controlled clinical trials. New drugs require demonstration of their effectiveness before they may be accepted and marketed. A compelling case is made that similar regulations should be applied to surgical (and medical) innovations.

Section III is an "Assessment of costs, risks, and benefits of established surgical procedures." A number of time-revered justifications for surgery are considered and questioned. Duncan Neuhauser examines elective inguinal herniorraphy in the elderly as a means of avoiding the high mortality frequently attributed to obstruction and strangulation and of improving the quality of life by relieving pain or avoiding an uncomfortable truss. Using decision analysis, he finds a small effect on expected days of life lost. Because

of the operative mortality of the elective operation, herniorraphy has a slightly greater (but negligible) loss of days of life than conservative therapy. Reliable data are not available to assess quality of life with and without a truss. Since anecdotal evidence is conflicting, there is a disarming suggestion of letting the patient try a truss so he can make his own decision regarding surgery.

In a later essay based upon data from the National Halothane Study, Fitzpatrick, Neutra, and Gilbert conclude that a small benefit may be expected from removing silent gallstones in young low-risk patients (an average prolongation of life with surgery of one to two weeks). This benefit disappears as risk or age are increased. Only about 3.4 percent of patients with gallstones are operated on in the United States each year. If cholecystectomy for silent stones in younger low-risk patients were to be adopted on a national basis in the U.S., a program of screening and surgery would cost about 1.3 billion dollars per year. Clearly, cholecystectomy for silent gallstones cannot be done on a universal basis, but neither can the decision to operate be based solely on increased life expectancy.

Certain discretionary operations have wide variations in their rates, as is noted by Gittelsohn and Wennberg (whose contribution to Section I, "On the incidence of tonsillectomy and other common surgical procedures," is even more cogent in the present context). Tonsillectomy, varicose vein stripping, hemorrhoidectomy, elective hysterectomy, and cholecystectomy are prime examples. These procedures appear to be resource-sensitive—that is, their rates depend on the availability of facilities and surgeons rather than on the prevalence of abnormal conditions.

Related evidence from Ontario (not cited in this book) indicates that despite the fact that the rates for certain discretionary procedures are associated with numbers of beds and surgical personnel, communities served by medical school-related hospitals have lower rates of elective surgery despite their higher ratios of beds and surgeons (Stockwell, 1977). Community standards and peer pressures are apparently potent forces. Dyck in Saskatchewan showed that hysterectomy rates could be reduced by retrospective audit of hospital records and publication of the results (Dyck et al., 1977).

Bunker, McPherson, and Henneman examine elective hysterectomy (for causes other than cancer, large, painful, or bleeding fibroids, or uterine prolapse). They find a possible increase of 2 to 4

weeks in life expectancy after elective hysterectomy and oophorectomy in young low-risk patients, primarily due to precluding cancer of the uterus and ovaries. Again, however, as with cholecystectomy for silent stones, increased age or risk reduced average life expectancy among this group largely as a result of the increased risks of surgery. Reduced mortality is not the issue; any benefits of elective hysterectomy will be due to an improved quality of life. On the other hand, some of these benefits may be offset by the psychological sequelae of surgery. As a result, the decision to do elective hysterectomy should be made only by the informed patient. Under universal health insurance, society also has a stake in the decision process. It must decide whether to use public funds for this or any elective procedure whose "qualitative" benefits cannot be readily quantified.

Section III ends with an analysis of breast cancer and concludes that radical surgery has not emerged as the most effective means of treatment. Because that operation is so firmly established, McPherson and Fox realistically but unhappily conclude, "The burden of proof has been shifted from the demonstration that a procedure is effective to a demonstration that a procedure provides no benefit. Proof of innocence is indeed formidable!" (p. 319).

In the next group of essays the costs, risks, and benefits of a number of new procedures are considered. The dearth of controlled trials in their assessment is noteworthy. The analysis which examines coronary artery bypass surgery, by Weinstein, Pliskin, and Stason, is predicated on the effectiveness of that procedure, at least for certain patients. Decision analysis in this instance is a valid method only if bypass surgery relieves angina and improves the quality and length of life. Although many reports, and only a few controlled trials, have reported efficacy, the decision to do bypass surgery would be confirmed by positive results from additional definitive studies of effectiveness. In the absence of rigorous investigation another expensive procedure may be prematurely accepted only to be proven ineffective later.

While ethical considerations are examined throughout this book, they are especially pertinent to those essays in which randomized controlled clinical trials are advocated. Drawing on evidence presented in the assessment of surgical innovations, the authors remind us that about half of the innovations considered were less satisfactory than existing treatments. We cannot predict, ex ante, in which instances experimental or control groups will fare better. Thus, on average, no greater risks accrue to either group. Statistical

techniques such as sequential analysis make it possible to terminate trials if the innovation proves to be either clearly effective or useless. The authors conclude that clinical trials facilitate prompt acceptance or rejection of new therapies.

In general, published papers rarely offer full descriptions of the ethical measures employed. This lack may be the result of an oversight in the report or, as I suspect, in the project design. The contributors to this volume suggest that greater attention and more specification are in order; ethical behavior in human research requires a provider-patient partnership predicated on disclosure and truly informed consent.

This compendium of essays concludes with a helpful set of recommendations. First, appropriate studies of the effectiveness of surgical treatment should be carried out for selected conditions, particularly those where uncertainty leads to professional disagreement. Effectiveness is the key question. Difficult as the cost elements are to determine, cost-benefit analysis is predicated equally on the determination of the beneficial effects of a procedure. Unless the beneficial effects are known with certainty, the benefit part of the equation is, at best, a guess. The application of cost-benefit analysis to some surgical procedures or medical treatments may be premature unless effectiveness is established.

One unequivocal "benefit" with low risk and low cost of this important book is the questions it asks. Do intensive care units, hysterectomies, or cholecystectomies do any good and—if they do—in what circumstances are they useful? If they are effective, what are the financial and personal costs? Which presently accepted treatments should be discarded or sharply limited, and which new procedures should be adopted? Until now, these questions have been rarely asked and even less frequently answered.

The controversy and lack of agreement regarding efficacy are measures of our knowledge gaps, intensified and perpetuated by the attitudes of some physicians. To be sure, in many instances the physician is forced to make decisions in the absence of clear-cut evidence, and must treat patients who are ill and demanding immediate relief. The physician's self image makes conservative measures and "expectant treatment" less acceptable. In the absence of proven measures he does the best he can, frequently by accepting unproven statements and conclusions by "experts" or by relying upon his own experience in lieu of scientific evidence.

This lack of knowledge has also afforded license to the less

critical physician to carry out unproven but highly profitable procedures in large numbers. Fee-for-service payments and patient pressures have intensified this behavior. In assessing treatments, this book suggests more conservative treatment or no treatment as alternatives to particular interventions. However, neither the methods of payment nor the patient expectations which are prevalent in the United States or Canada support reduced medical intervention. As a result the "doer" is rewarded and the "waiter" criticized and penalized financially.

The second recommendation states: "Our grasp of the components of cost benefit analysis and their interrelations, the values of the various data gathering techniques, and our understanding of the ethics of data gathering must be improved by theoretical and empirical work and by continued discussions in the public forums" (p. 393). The third recommendation is prefaced by a call to assess the efficacy of existing treatments and to develop a policy for the introduction of new medical and surgical technology. Recommendation III declares: "These principles of cost benefit evaluation should be included as an integral part of the medical school curriculum; and their application to the assessment of the efficacy of medical care should be incorporated into clinical practice and continuing medical education" (p. 394).

The order of the book's first three recommendations is not random. Medical students and physicians must learn to distinguish effective from non-effective treatments. Such skills require understanding of the methods used to assess effectiveness so that evidence can be properly evaluated, useful procedures adopted, and ineffective ones discarded.

Recommendation IV deals with public knowledge and consumer expectations: "Information on outcomes as well as costs of medical care should be routinely formulated in a manner suitable for presentation to the public" (p. 394). Critical attitudes on the part of physicians must be coupled with changes in consumer attitudes. Unrealistic expectations must be replaced by the knowledge of what medical treatments can and cannot do.

Bunker, Barnes, and Mosteller are aware that their book, like Cochrane's (1972) *Effectiveness and Efficiency,* is an important step in rationalizing medical and surgical treatment. They state in the

preface (p. xvi), "All told, then, the methods presented in this book are intended to create the atmosphere and provide a set of approaches that may be of continued value in reaching solutions for many surgical and medical problems. Specific chapters may be quickly outmoded, or be found to need revision. It is the thrust of the methods and the value of joining these ways of thinking to those commonly used by physicians that form a large part of what the reader should gain from this book. Other aspects, it is hoped, will contribute in concrete ways to the design of surgical research generally and to our understanding of medical issues in single diseases such as diagnosis and evaluation of therapy for the individual patient."

The book's issues with regard to all medical care are generic. Surgery has been used because of its importance and because it is readily amenable to assessment. However, the examples cited include a few non-surgical treatments (e.g., gastric freezing, intensive care units). We must ask the same questions about *all* treatments, about cholecystectomy for asymptomatic gallstones as well as benzodiazepines for stress and anxiety.

Only with better studies of effectiveness, increased training in clinical epidemiology for practitioners, and greater discernment by users can physicians and patients make joint and more rational therapeutic decisions. Public policy on medical care could then be both better informed in its inception, and more readily understood and accepted in its implementation. This book reviews present methods of assessment, the state and shortcomings of our knowledge, and the gaps in existing data. As such, it is an essential step in the process of closing the door on unquestioning acceptance and entering the realm of critical judgment and demystification.

References

Cochrane, A.L. 1972. *Effectiveness and Efficiency: Random Reflections on Health Services.* London: Nuffield Provincial Hospitals Trust.

Creese, A. 1976. Economics and health planning. In *Health Services Planning,* edited by K. Dunnell. London: King Edward Fund.

Dyck, F.J., et al. 1977. Effect of surveillance on hysterectomies in Saskatchewan. *New England Journal of Medicine* 296 (June 9): 1326–1328.

Stockwell, H.G. 1977. *Variations in surgical rates in Ontario.* M.Sc. thesis. Hamilton: McMaster University.

Address reprint requests to: Eugene Vayda, M.D., Professor and Chairman, Department of Health Administration, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada