

McPherson K, Bunker JP (2006). Costs, risks and benefits of surgery: a milestone in the development of health services research.



© Klim McPherson, Nuffield Dept of Obstetrics & Gynaecology, Research Institute, Churchill Hospital, Roosevelt Drive, Headington, Oxford OX3 7BN. Email: klim.mcpherson@obstetrics-gynaecology.oxford.ac.uk.

Origins

In 2006, the *Journal of Health Services Research and Health Policy* included *Costs, risks and benefits of surgery* in a list of 26 books considered to have been most influential in changing health services and health care policy over the previous century and a half (Black and Neuhauser 2006). How did *Costs, risks and benefits of surgery* come to be written?

The early 1970s was characterised by a growing acknowledgement in medicine, public health and the social sciences that contemporary health provision was neither evidence-based nor necessarily what people would prefer, given true choices. The great variation in health care utilisation rates observed across countries and communities gave rise to legitimate questions about macro-economic effectiveness and efficiency. Archie Cochrane's book *Effectiveness and Efficiency: random reflections on health services* (Cochrane 1972) asserted and justified the view that much of modern health care was not based on reasonable evidence about efficacy or safety. This came as a shock to many. Cochrane's book was especially popular in the USA, exemplified by Cochrane's recollection of meeting an American who said "So you are Archie Cochrane. I bought fifty copies of your book as Xmas cards this year."

Fifty years previously, Codman had proposed and implemented systematic assessment of the end results of surgery at his hospital in Boston (Codman 1916). Although questions about the use of discretionary surgery had been raised in the 1930s by Glover's (1938) demonstration of widely varying rates of tonsillectomy, the late 1960s and early 1970s witnessed an explosion of studies documenting unexplained international and intra-national variations in surgical rates (Lewis 1969; Bunker 1970; Lichner and Pflanz 1971; Wennberg and Gittelsohn 1973; Vayda 1973).

Each of these studies raised fundamental questions about what health care was for, and ultimately for whose benefit. When such different amounts of it were being delivered to apparently similar patient populations with similar outcomes, it was not possible any longer to entertain the notion that all of it was unambiguously beneficial and appropriate. The variations in cost were also fairly clear (surgery is expensive), but the benefits and the consequent risks of these variations were much less obvious. It seemed irrational to remain so ignorant when it was so unclear that greater expenditure resulted in greater aggregate benefit.

Given the general level of uncertainty evident throughout health provision, both about morbidity levels and the effects of most treatments, questions about how much health care is best for communities cannot generally be resolved by knowing whether treatment A is better (or not) than treatment B, among a specific group of patients. The justification offered for high rates of surgery was often that prophylactic procedures (cholecystectomy and hysterectomy, for example) prevented problems in the longer term. However, although prophylactic surgery with modern infection control and anaesthesia might well fulfil the needs of individuals, what were the costs for communities? Moreover did the community want to use its resources in this way, given other options? Why was there no strong evidence that, within developed countries at least, more health care resulted in measurably better health? Was health care running away with itself in places, on essentially spurious implied clinical benefit grounds?

The move to Harvard University of two of the researchers who had contributed to the surgical variations evidence (John Bunker and Jack Wennberg) provided a stimulus to address some of these quandaries. Fred Mosteller (Professor of Statistics at Harvard) and Howard Hiatt (Dean of the Harvard School of Public Health) had already planned to have multidisciplinary seminars, accompanied by a meal at the Faculty Club, and Bunker persuaded them that the questions raised by the surgical variations data should be addressed in these. The basis for this kind of question and indeed many of the methodologies to address them had already received impetus from the creation, in the School of Public Health, of the Centre for the Analysis of Health Practices (CAHP) in the early 1970s. People like Milt Weinstein and Don Shephard under the direction of Hiatt had progressed with the hard nosed assessment of

Cite as:

McPherson K, Bunker JP (2006). Costs, risks and benefits of surgery: a milestone in the development of health services research. JLL Bulletin: Commentaries on the history of treatment evaluation (www.jameslindlibrary.org).

comparing treatments for chronic disease using models and other methods - clearly the nature of the question already had a resonance. Fortnightly seminars were held between 1973 and 1976, and, in all, about fifty people presented completed work, discussed work in progress, proposed future work, and argued through uncertainties and methodological issues.

During the course of the series it became apparent that although randomized trials were essential for assessing the relative merits of some specific procedures, other research methods were needed as well, particularly as most of the issues seemed to concern well established, but differing, surgical practice styles about which individual surgeons were not sufficiently uncertain to tolerate random allocation. Furthermore, the questions raised went beyond merely clinical issues, and were thus only partially amenable to better evidence about attributable clinical outcomes. For example, what was the role of health extra care provision compared to other services? It was felt that adequate clinical outcome information might come from routinely collected data, reviews, modelling, or observational studies - at least to suggest hypotheses.

The contributors

The end product of the seminar series was the book *Costs, risks, and benefits of surgery*, edited by John Bunker, Benjamin Barnes and Frederic Mosteller. Bunker (an anaesthetist) had collaborated previously with Mosteller in a study to assess the risks of fatal hepatic necrosis following administration of the anaesthetic halothane. The most important observation of the study turned out to be the very large variations in overall surgical mortality among the 34 hospitals that participated (Bunker et al. 1969). This led to a search for an explanation of the high rates of surgery in the United States when compared to rates in Great Britain (Bunker 1970). Bunker and Wennberg (1973) subsequently reported a positive association of rates of surgery and iatrogenic mortality - a generalised version of the hypothesis that the positive relationship between high appendectomy rates and deaths attributed to appendicitis might reflect a larger number of normal appendices being removed with a procedure that had some inherent risks (Lembcke 1952). Barnes, the third member of the editorial team, was a general surgeon at the Brigham Hospital in Boston.

Contributors to the seminars and books included epidemiologists, statisticians, policy researchers, economists, anaesthetists, surgeons, internists and aspiring health service researchers, members of Mosteller's department at Harvard being particularly active in the work. The meeting of these disparate disciplines, on a novel but specific question, was exciting. Each was variously perplexed by the phenomenon under consideration, and each anxious to make a significant contribution from his or her perspective. Everyone began with a different lexicon and different prior beliefs. It was miraculous that, with the faintly possible exception of the economists (who, it turned out, were probably the most naïve group), no discipline tried to dominate the proceedings. The complexity of the questions, and the related uncertainties, soon subdued even the most experienced participants to a shared humility - a stunning achievement in the rarefied atmosphere of Harvard Medical School! For a young researcher (KMCP) on sabbatical from the UK at the time, the whole, egalitarian experience was utterly transfixing, even if a certain misplaced credibility seemed to attach to someone familiar with a system of implicitly rationed health care, which could be seen as a worthy control arm.

During the course of the seminar series, the American College of Surgeons had established a Study of Surgical Services in the US (SOSSUS), partly in response to the implied accusations of overprovision. The study showed that workload was strongly related to surgeons' income, but contributed little to any understanding of the outcomes of surgical services in the US in a context of genuine social choice.

Many of the contributors to the Harvard seminars attended working meetings of various sections of the American College of Surgeons' Study of Surgical Services. Indeed KMCP was part supported by this endeavour and was responsible for the analyses of a large questionnaire survey to US surgeons. SOSSUS gatherings seemed to symbolise the dominant complacency within the surgical community, and they thus influenced the chapters being drafted for the book. Each seminar participant began to write chapters based on their researches and talk about current progress and thoughts. Sometimes colleagues felt some of these ideas to be too heretical and said so. People thus had either to stand their ground, or concede to stronger evidence - but this was really the only judgement criterion. Attempts to pull rank - on the basis of age, experience or discipline - were very rare.

Two core members of the group kept up the intellectual pressure by publishing an editorial in the *New England Journal of Medicine* suggesting that part of the explanation for higher surgical mortality rates in the US compared with the UK might be attributable to much higher discretionary surgical rates in the US, and the consequent mortality (Bunker and Wennberg 1973).

The book itself

The book itself was eventually published in 1977. *Costs, risks, and benefits of surgery* has 23 chapters, most the subject of individual seminars, written by 35 authors. There are four sections addressing, respectively: general principles of evaluation; some accounts of specific surgical innovations and their evaluations; an attempt to assess the costs, risks and benefits of established procedures; and finally, assessment of new procedures. The book ends with a summary and recommendations.

The first section of the book raised important questions about cost-benefit methodology. Acknowledging quality of life as a legitimate outcome of surgery – especially when operations could not be expected to prolong life - and understanding the importance of patient preferences in making decisions, the benefits of surgical intervention had to be compared with the effects of other socially beneficial interventions. The book emphasised the transitory nature of many of these phenomena, as well as the difficulty in measuring them. For example, it drew attention to the need for analysis and interpretation of surgical rates in populations to take account of the extent of organ loss attributable to previous surgery.

The second section of the book began with an account of important aspects of surgical evaluation in recent history, to demonstrate the particular nature of the challenges in this field of research. It considered discarded operations and how they came and went, the implications of progress in surgical and anaesthetic technology, improved quality of life due to surgery, and seminal examples of surgical innovation. This section is notable for containing one of the earliest examples of the use of meta-analysis in medicine, to assess the effects of surgical treatments for duodenal ulcer (Cochran et al. 1977).

The book's third section used examples of situations in which variation in surgical practice styles had prompted uncertainties about benefits and risks. Elective hernia repair, cholecystectomy for 'silent' gallstones, elective hysterectomy, surgery for appendicitis and treatment for breast cancer were examined, demonstrating that a variety of methodological approaches were needed. In the breast cancer chapter, for example, McPherson and Fox (1977) examined the evidence on the unexplained variation in response to the extremes of treatment options, from nothing to super-radical mastectomy. The authors argued that the then dominant treatment modality in the USA, radical mastectomy, could not be supported by the evidence and that the biological model of disease progression on which it was based was, at best, merely the most plausible of many alternative models. They went on to suggest that ignorance of the complex natural history of disease progression, and what might influence it, was so remarkable that any claim for intrinsic plausibility of the dominant model was suspect. The evidence, such as it was, suggested that treatments available then were all of limited utility for many breast cancer sufferers. Similarly plausible models of disease progression might lead to much more conservative surgery with similar, or even better, outcomes.

In the fourth and final section of the book, analyses of new surgical procedures and intensive care were presented, using the knowledge already reviewed. The costs of treating end-stage renal disease, coronary by-pass surgery, and intensive care were analysed, assessing the implications of different assumptions about the type of patient treated and the possible benefits.

In summary, although *Costs, risks, and benefits of surgery* concerned only one aspect of health services, it can be seen now as an early basic text of health services research, covering both principles and methods.

The book's legacy

As it had done with Archie Cochrane's book *Effectiveness and Efficiency: random reflections on health services* (Cochrane 1972) five years previously, the lay press took a great deal of interest in *Costs, risks and benefits of surgery*.

Articles usually seized on a few issues highlighted by the book, quite frequently a critical evaluation on treating silent gallstones, or of radical surgery for breast cancer, which was then a sensitive public issue in the United States. For example, the chapter on the treatment of breast cancer (McPherson and Fox 1977) was featured on the front page of the *San Francisco Chronicle*. It was big news at that time that the Halsted radical mastectomy might not be the best treatment for breast cancer.

The New York Times (24 May 1977) ran a headline "Harvard Group Questions Cost and Value of Much Surgery", and went on to review the principles of proper evaluation which the book had sought to cover. It concluded that increasing costs would inexorably force such analyses and these should necessarily be part of any future health plan. Quoting Barnes they wrote "... not performing operations might save many millions of dollars at a cost of relatively few lives, but there is no way of knowing that now..".

A headline in *The Palo Alto Times* (23 May 1977) read "New rules urged for judging surgery", and quoted Bunker: "public and professional attitudes towards surgical care will have to change. There is a belief that a human life is priceless and no costs should be spared... but small benefits or risks associated with much surgery require sophistication to measure them and they might not be worth the money."

What was the reaction from the professions? The night that the article questioning the need for radical surgical treatment of breast cancer appeared in the *San Francisco Chronicle*, Bunker was called by an outraged pathologist, who had trained at the institution at which radical mastectomy had been introduced, to complain about the chapter on breast cancer. *American Medical News* (30 May 1977) used the banner headline "Calm voices join the 'unnecessary surgery' debate", and reported extensively on the chapter on treating silent gallstones (Fitzpatrick et al. 1977). Bunker,

Barnes, Mosteller and Hiatt were all cited in the newspaper. Barnes emphasised the need to avoid using pejorative and vague terms like 'unnecessary' surgical operations, but instead, to increase understanding of their actual benefits and risks.

At the end of the year in which the book was published, *The Milbank Quarterly* devoted an entire issue of the journal to the Study of Surgical Services in the US (SOSSUS) and *Costs, risks and benefits of surgery*. Commenting on the former, Edward Hughes, of Northwestern University, concluded that "Despite the considerable amount of money, time, and energy lavished on SOSSUS, the report does little more than support concerns initially voiced by others regarding a possible malfunction in the delivery of surgical services in the United States." (Hughes 1977). The same issue contained a favourable review of *Costs, risks and benefits of surgery* by Eugene Vayda (1977).

There were a dozen or so largely favourable references to the book reviews in the professional press, usually in the context of an article about cost-effectiveness research. An article in the *New England Journal of Medicine* stated: "For many elective operations such as a repair of a symptomatic hernia or hysterectomy for excessive uterine bleeding, the major concerns are relative benefits and costs of alternative strategies. The methods involved in answering these questions are well described in an excellent book by Bunker and his colleagues." (LoGerfo 1977).

An anonymous editorial in the *Annals of Internal Medicine* (Anonymous 1978) commented "... the introduction of the randomized clinical trial into medical science by Bradford Hill... provided clinicians...with a uniquely powerful tool for determining efficacy of medical treatments....there has been a growing interest in the application of cost-benefit techniques to medical treatment. A recent monograph, *Costs, risks and benefits of surgery*, contains a number of essays illustrating the importance of the latter two scientific tools." The editorial goes on to describe in considerable detail chapters on breast cancer by McPherson and Fox (1977) and on inguinal hernia by Neuhauser (1977).

Writing in the *Journal of the Royal Society of Medicine*, a British surgeon wrote as follows: "The craft of surgery consists of easily recognisable items of service which, because of their essentially mechanistic nature, can be fairly readily quantified. Thus it is hardly surprising that a good deal of attention has recently been paid to what could be termed the economics of surgery – the costs incurred and the benefits that might ensue (Bunker et al. 1977). To some this approach is quite unacceptable, but it does seem inevitable that with a definite financial ceiling ... we should at least have some crude ideas of what value we get for the money spent." (Dudley 1978).

Although the book appears to have been formally cited in other publications on less than 200 occasions over the quarter century since it was published (Bunker, unpublished data), it does seem to have helped to challenge notions of clinical certainty based on plausibility, but without actual evidence, and to promote recognition of the way that public health and clinical research can complement one another.

A 1978 review of the book in *Science* by a senior statistician at Johns Hopkins University (Shapiro 1978) identifies the probable reasons that *Costs, risks and benefits of surgery* was judged, a quarter of a century later, to be one of the most influential books on health services and health care policy. The review suggested that "... the volume will have great durability. The interplay of conceptual, analytical, and mathematical approaches to the subject and the extensive use of examples should make it useful to academicians and policy makers... and should attract researchers to the difficult areas of investigation mapped out." He thus predicted the effect of the book on health services research as an academic discipline, and this is probably its most enduring legacy.

This *James Lind Library* commentary has been republished in the *Journal of the Royal Society of Medicine* 2007;100:387-390.

References

Anonymous (1978). Myocardial infarction: unit care or home care? *Annals of Internal Medicine* 88:259-61.

Black N, Neuhauser D (2006). *Journal of Health Services Research and Policy* 11:180-183.

Bunker JP (1970). Surgical Manpower: a comparison of operations and surgeons in the United States and in England and Wales. *New England Journal of Medicine* 282:135-44.

Bunker JP, Wennberg JE (1973). Operation rates, mortality statistics and the quality of life. *New England Journal of Medicine* 289:1249-51.

Bunker JP, Forrest WH, Mosteller F, Vandam LD (1969). *The National Halothane Study*. Bethesda, Md: National Institutes of Health.

Cochran WG, Diakonik P, Donner AP, Hoaglin DC, O'Connor NE, Peterson OL, Rosenoer VM (1977). *Experiments in*

surgical treatments of duodenal ulcer. In: Bunker JP, Barnes BA, Mosteller F, eds. Costs, risks and benefits of surgery. Oxford: Oxford University Press, p 176-197.

Cochrane AL (1972). Effectiveness and Efficiency: random reflections on health services. London: Nuffield Provincial Hospitals Trust.

Codman EA (1916). A study in hospital efficiency. Boston: privately printed.

Dudley H (1978). Economics and surgery. *Journal of the Royal Society of Medicine* 71:397-8.

Fitzpatrick G, Neutra R, Gilbert JP (1977). Cost-effectiveness of cholecystectomy for silent gallstones. In: Bunker JP, Barnes BA, Mosteller F, eds. Costs, risks and benefits of surgery. Oxford: Oxford University Press, p 246-261.

Hughes EFX, Lewit EM, Pauly MV (1977). The "Study on surgical services for the United States": A valid prescription for American surgery? *Milbank Memorial Fund Quarterly* 55:465-84.

Lewis CE (1969). Variations in the incidence of surgery. *New England Journal of Medicine* 281:880-4.

Lichtner s, Pflanz M (1971). Appendectomy in the Federal Republic of Germany: epidemiology and medical care patterns. *Medical Care* 9:311-30.

LoGerfo JF (1977). Variations in surgical rates: fact vs fantasy. *New England Journal of Medicine* 297:387-89.

McPherson K, Fox MS (1977). Treatment of breast cancer. In: Bunker JP, Barnes BA, Mosteller F, eds. Costs, risks and benefits of surgery. Oxford: Oxford University Press, p 308-322.

Neuhauser D (1977). Elective inguinal herniorrhaphy versus truss in the elderly. In: Bunker JP, Barnes BA, Mosteller F, eds. Costs, risks and benefits of surgery. Oxford: Oxford University Press, p 223-245.

Shapiro S (1978). Medical care: issues of evaluation. *Science* 199:964-66.

Vayda E (1973). A comparison of surgical rates in Canada and in England and Wales. *New England Journal of Medicine* 289:1224-9.

Vayda E (1973). When is surgery indicated? *Milbank Memorial Fund Quarterly* 51:493-503.

Wennberg J, Gittelsohn A (1973). Small area variations in health care delivery. *Science* 182:1102-08.

[Home](#)

[Contents](#)