To Subscribe:  https://fulfillment.healthaffairs.org

Health Affairs is published monthly by Project HOPE at 7500 Old Georgetown Road, Suite 600, Bethesda, MD 20814-6133. Copyright © by Project HOPE - The People-to-People Health Foundation. As provided by United States copyright law (Title 17, U.S. Code), no part of may be reproduced, displayed, or transmitted in any form or by any means, electronic or mechanical, including photocopying or by information storage or retrieval systems, without prior written permission from the Publisher. All rights reserved.

Not for commercial use or unauthorized distribution
A DIFFERENT VIEW OF QUEUES IN ONTARIO

by C. David Naylor

Prologue: Every health Care system, regardless of how rich the country in which it operates, rations medical services, because no nation has the resources to match the insatiable demand for services. In the variety of approaches to rationing that nations employ, as David Naylor points out in this paper, the United Kingdom and the United States represent the extremes. Britain’s National Health Service (NHS), which offers patients medical Care that is free at the point of service, practices queue-based rationing. People face time delays before medical problems are addressed. In the United States, those with health insurance rarely have to wait long for treatment. But those without insurance have no ready access to care and must fend for themselves in public hospitals and other institutions prepared to accept charity cases. Canada prides itself on developing a health Care system that strikes a middle ground, it is publicly funded and universally available, but care is privately provided. Administration and delivery of Care are decentralized. But as demands for service have increased, Canada has resorted to T-a-tioning by queue for some procedures, such as cardiac surgery, for which demand exceeds supply, Naylor, an assistant professor on the faculty of medicine, University of Toronto (UT), holds a medical degree from UT and a doctoral degree in social and administrative sciences from Oxford University, where he was a Rhodes Scholar (1979-1983). In this paper, Naylor describes the rationing of coronary care in Ontario, Canada’s richest and most populous province. He recounts the evolution of Ontario’s crisis over growing waiting lists and how the system has addressed the problem. One fascinating dimension of rationing, Canadian style, is the interface between the media, who publicize the fate of patients who fall victim to the queues, and politicians, who respond to the resulting hue and cry in a way that underscores their accountability to the electorate.
On 12 December 1990, The Wall Street Journal carried a story with the dramatic headline: “Canadians Cross Border to Save Their Lives.”¹ The story used anecdotes from the Canadian province of Ontario (population 9.5 million) to suggest that dangerously long waiting times for open-heart surgery were the norm and gave favorable coverage to an organization that facilitates movement of patients from southwestern Ontario to cardiac surgery centers in the United States.

In fact, long waits in Ontario for open-heart surgery were most pronounced between mid-1987 and late 1989; by the time of the Wall Street Journal report, they had been ameliorated. Nonetheless, queues persist for other procedures, and the cardiac surgery situation has served as a bellwether for emerging access problems in Canada’s federal/provincial Medicare system. The existence of waiting lists for cardiac and other surgical procedures has also been noted by the American Medical Association (AMA) in advertising campaigns warning against rationing that might result from U.S. adoption of universal health insurance patterned on the Canadian model.²

In this article, I focus on the open-heart surgery situation in Ontario, particularly coronary artery bypass surgery (CABS) queues, which have represented the largest part of the delayed caseload.³ I use the case of queues for coronary surgery to develop two basic arguments. First, queue-based allocation of services, particularly if it is predicated on explicit and objective criteria with selective delay, is potentially superior to price-based rationing, because the latter will-almost invariably lead to implicit and arbitrary denial. Second, and as a corollary, the real issue for any health care system dedicated to universal access is not that queues exist for some services, but rather how best to measure, monitor, and manage them. In this latter respect, the CABS crisis may actually indicate less about rationing and more about inadequate information systems and central management in the Canadian provincial health insurance plans.

Context: The Ontario Provincial System

Ontario offers universal first-dollar coverage for all basic medical and hospital services. General hospitals are private nonprofit corporations, funded predominantly by direct government allocations that are increased annually by uniform amounts. These prospective payments, known as global budgets, serve as a fiscal envelope within which each hospital must function for the following year. Major capital allocations are made separately. Global operating budgets are supplemented by various categories of add-on payments. Add-on funds can flow to permit establishment or expansion of specialized programs such as open-heart
surgery, transplantation, and dialysis. These include one-time capital renewal grants and ongoing operating funds. Add-on funding formulae also exist to help offset the extra costs imposed by rapid growth in approved high-technology services.

Intraregional duplication is common for services funded through hospital global budgets. However, by providing extraglobal supplements only to designated referral centers, the health ministry can create strong fiscal pressures for regionalization of highly specialized services. One corollary is that only for programs initially given funds outside the global budget does the provincial government have a potential “purchaser” position, with clear caseload targets to be met by hospitals. Even that position is weak, because the ministry quickly allows new program funds to roll back into the global budget without continued programmatic funding and line-item caseload accountability.

General hospitals are administered by professional managers who report to boards of public trustees. Although district health councils exist, their authority is minimal, and regional planning is rudimentary. The various funding formulae also tend to limit central accountability or provincial planning, since most funds are transferred as a bloc without clear utilization targets.

Most physicians are in private practice and paid predominantly fee for service. Direct charges to patients above and beyond the benefits schedule have been forbidden by law since 1986. Administrative overheads in the fee-for-service sector are much lower than in the United States because of the single payer with automated claim-card processing and direct batched payments to providers. Savings result for both doctors and the provincial prepayment agency. Despite the regulated payment system, clinical autonomy has been preserved, and managed care mechanisms are virtually nonexistent.

Global budgets for hospitals also permit large overhead savings, since bills are only submitted to patients or insurance agencies for such items as semiprivate or private rooms. Partly in consequence, most Ontario hospitals lag at least a decade behind their U.S. counterparts in expenditure tracking and management information systems. Central information tracking is similarly limited. Indeed, the provincial ministry of health has repeatedly been unable to produce data about what actually happens on the front lines of health care.

### The Problem

**Regionalization.** Adult cardiovascular surgery in Ontario is one of only a few services that are highly regionalized. This has occurred partly by
historical happenstance but has been reinforced by patterns of provincial funding through the years. Until late 1989, nine hospitals offered CABS: eight in southern Ontario, where the population is concentrated, and one in the north. All eight southern centers were teaching hospitals with salaried housestaff; supervising cardiovascular surgeons hold medical faculty appointments, but their incomes depend on fee-for-service billings by their group practice.

Attention began to focus on open-heart surgery waiting lists during 1988. In May of that year, a senior Toronto cardiovascular surgeon gained front-page newspaper coverage by announcing that there had been twelve or thirteen deaths on his personal waiting list since 1987. Queues for open-heart surgery had actually been lengthening steadily from the mid-1980s. In Toronto, with annual caseloads of about 2,500, one estimate is that in 1984 there were 356 patients waiting an average of two to three weeks, as compared with 848 patients waiting three to nine months in 1989. (These data; from informal reports by surgical chiefs, presumably reflect waiting periods for stable patients without urgent indications for revascularization who had been seen and accepted by a cardiac surgeon.)

**Efforts to prioritize.** Resources normally permitted some operating room slots to be set aside for unexpected cases needing rapid attention. However, increased demand and long waiting times reduced willingness to leave such openings. Elective patients were therefore more likely to be “bumped” not only by the usual volume of urgent cases but also by previously elective patients who had destabilized while waiting. A few surgeons moved canceled patients to the bottom of their lists to avoid having to cancel and reshuffle bookings for all other patients—a practice that penalized those unfortunate enough to be deferred. Others tried to reinsert the canceled patient near the top of the queue, increasing the likelihood of multiple cancellations and booking chaos. Practitioners at some hospitals wisely stopped offering firm elective bookings, instead suggesting only a range of dates.

With adverse publicity and political pressure mounting, the provincial health ministry funded a multidisciplinary provider group to develop better mechanisms for managing urgent referrals and waiting lists in the Toronto region. Among the group’s objectives was to develop consensus criteria for determining which patients deserved priority for CABS. To this end, an expert panel process, patterned on RAND Corporation work on the appropriateness of care, was initiated in fall 1988. Sixteen cardiologists and surgeons agreed on practice guidelines, a waiting time scale, and a clinical multifactorial scoring system to assess patient priority for revascularization.
At the end of 1988, the deaths of two cardiac surgery patients in Toronto attracted national media attention. One death was actually postoperative, but both patients had suffered long waits and multiple cancellations. In mid-January 1989, the Ontario Medical Association launched a poster campaign in doctors’ offices highlighting access problems, including the following: “Some people wait so long for surgery, it’s no wonder they’re called patient. Most heart patients are now waiting six months or more for triple bypass surgery—if they make it that long.”

**Effect of demographics on supply and demand.** What led to the imbalance between supply and demand? CABS rates among Ontario adults increased by 39 percent between 1979 and 1983 and then showed minimal change through to 1988, stabilizing at forty to forty-five procedures per 100,000 (Exhibit 1). Major caseload growth was constrained in part because resources were fixed. Although extraglobal funding arrangements were in place to reimburse hospitals for minor overruns, each hospital’s resources were subject to intense internal competition. Hence, competing priorities tended to preclude institutional reallocation of funds from other services supported by the global budget.

Note, however, that the plateau in Ontario rates was also associated with a demographic shift: the proportion of procedures in the over-sixty-five population rose from 12.8 percent in 1979 to 27.4 percent by 1985. Since overall rates were more or less stable, one then wonders why demand increased, particularly among the elderly. Coronary artery disease incidence appears to be falling, or at least migrating to the later stages of

<table>
<thead>
<tr>
<th>Operations per 100,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>---</td>
</tr>
<tr>
<td>28.6</td>
</tr>
</tbody>
</table>

**Exhibit 1**

Rates Of Coronary Artery Bypass Surgery Per 100,000 Population In Ontario, 1979-1988

life; prevalence is not falling, due to lower death rates from the disease at all stages. However, the shift to more elderly patients undergoing CABS far outstripped that which would be expected from changes in patterns of the underlying disease.

The role of patients’ expectations is uncertain, but with respect to providers, one can conceive of a medical funnel and a surgical spout. Increasing activity in the medical funnel is explicable. Canadian physician supply has expanded rapidly; recent data show increasing application of all services, and especially diagnostic tests, to elderly patients. Coronary atherosclerosis increases with age. If internists and family physicians are more aggressive about noninvasive tests and referrals for coronary angiography, and if cardiologists in turn respond with a higher proportion of angiograms among referred patients, more and more patients—elderly and otherwise—will meet symptomatic and anatomical criteria for CABS. As to the surgical spout, it appears that as CABS became accepted as a routine procedure, greater confidence in the technical safety of the procedure led surgeons to accept more elderly and higher-risk cases.

This increased demand also affected supply, because elderly and high-risk patients have prolonged lengths-of-stay in postoperative intensive care units (ICUs). In one hospital, it was estimated that there was a reduction of 130 open-heart cases per annum (about 15 percent of the 700 cases funded) because complications after cardiovascular surgery slowed ICU bed turnover and delayed transfer to conventional wards.14

### The Solution

Continuing publicity and political pressure led the Ontario government to mount a special investigation into cardiac surgery. Nominally focused on one hospital, the investigation actually reviewed the entire provincial situation. The investigative team’s final report, released in February 1989, highlighted differences in waiting times among institutions and physicians; inconsistent approaches to inserting urgent patients into the operating rota; variation in assessing patient priorities for revascularization; and a general failure to monitor adequately the status of patients on waiting lists.15

About the time of the report’s release, data from chiefs of cardiac surgery showed that the total open-heart waiting list (CABS plus other procedures) for Ontario was 1,700 persons—more than 25 percent of the annual caseload of 6,400 procedures performed in 1989. Waiting times for elective cases ranged from as little as four to eight weeks in some centers to six months or more in others. Discrepancies existed within the same city, and even in the same hospital depending on the surgeon.
A multipronged response from providers and government was launched. With one new unit and expanded capacity elsewhere, a fixed increase of about 800 open-heart cases per year was attained, the majority of which would be CABS procedures. Spot funding was provided to address specific bottlenecks that were impeding some hospitals from meeting their previous caseload targets (for example, to expand ICU bed capacity and increase the complement of cardiovascular ICU nurses). Designated nurses were also funded for each hospital to assist in managing and monitoring waiting lists. Surgeons responded by changing booking procedures so that many more slots were held open for urgent cases, thereby avoiding repeated cancellations and dislocated schedules and, to some extent, using formal queue-ordering criteria to abet cooperation among providers. Tighter case-selection criteria were also followed in some centers to avoid problems with “blocked” ICU beds.

In January 1990, a survey sponsored by the Canadian Cardiovascular Society found that patients undergoing elective procedures still waited at least three months in about half of the Ontario open-heart surgery centers, although shorter queues were reported in the rest.\textsuperscript{16} Claims of deaths of patients on waiting lists continued: Heartbeat Windsor, an advocacy group pressing for an open-heart surgery unit in a city bordering Detroit, stated there had been at least fifteen deaths in Windsor and surrounding Essex County during the twelve months before July 1990.\textsuperscript{17}

However, as 1990 progressed, the combination of efficiency-enhancing maneuvers and increased capacity had a dramatic impact on the queues. By January 1991, the cardiac surgery waiting list in Ontario was estimated to have fallen from 1,800 persons the previous year to about 1,000. Average waiting times were also down dramatically, with elective patients waiting only a few weeks in the hands of most surgeons. Total 1990 caseloads had risen to 7,100, with further increases projected for 1991. Queues were also eased by growing use of percutaneous transluminal coronary angioplasty (PTCA)-the balloon dilatation technique that can serve as an alternative to CABS for some patients-and by movement of up to 300 patients across the border to U.S. heart surgery centers. However, given current capacity and waiting times, it seems doubtful that the Ontario government will continue reimbursement for elective open-heart surgery performed in the United States.

Lessons And Issues

Genesis and management of queues. Why do queues occur? The simple answer is a mismatch between supply and demand. As noted above, Ontario CABS queues lengthened due to an interaction of factors
on both the demand and supply sides of the equation and had been growing for years prior to the actual crisis.

Queuing theory is well developed in operations research and commonly applied to many enterprises.\textsuperscript{18} However, because no system for monitoring waiting lists was in place in Ontario, mathematical modeling and an optimal response were largely precluded. For example, even with a perfect match of supply and demand, a queue could persist as a result of past backlogs that remain uncleared. A queue that is more or less stable should therefore be attacked first by a short-term rather than a fixed increase in supply. In Ontario, only a brief attempt was made to reduce the queue through a transient supply boost in summer 1989. This strategy was not used again, even though the provincial queue was more or less stable during early 1989 to early 1990.

Instead, the primary and ongoing response has been to provide funding that would permit a fixed increase in caseload capacity of up to 20 percent, well over half of which has now been achieved. This increase was not based on any planning process that took account of regional queues or CABS rates in counties surrounding a given referral center, even though Ontario CABS rates vary as much as fivefold between highest and lowest counties. Instead, ministry funds were channeled wherever the fastest and largest caseload increments could be attained. Since a 10-12 percent increase in capacity has already greatly improved the situation, is funding for further expansion really needed?

A related question is how many cases should never have been in the queue in the first place. A major cooperative study is now under way to compare the appropriateness of CABS utilization in Canada and the United States. From my reviews of charts and registry reports, I find Canadian CABS case selection to be conservative; reasonable anatomic and symptomatic criteria appear to be fulfilled. Nonetheless, the dramatic increase in procedures among the elderly leads inevitably to concerns about risk/benefit ratios and the appropriateness of surgery for some in this subgroup. The demographics of Ontario CABS patients are not unique; the rate of growth among the elderly has been similar in the United States (Exhibit 2).\textsuperscript{19} Indeed, during the 1980s there was continued expansion of total U.S. caseloads, with Ontario rates ending up less than half those in the United States. CABS rates in the United Kingdom are about half those in Canada.\textsuperscript{20} The contentious issue, as always, becomes: Which rate is right, and for which age group? In sum, growth in waiting lists represents an opportunity to address appropriateness issues and other micro- and macromanagement concerns—provided, of course, that one has the information systems in place to preempt a crisis. If not, as was the case in Ontario, providers and patients are alienated, the media
Exhibit 2
Rates Of Coronary Artery Bypass Surgery Per 100,000 Elderly Persons, United States And Two Canadian Provinces, 1981 And 1985

<table>
<thead>
<tr>
<th>Operations per 100,000 elderly</th>
<th>U.S.</th>
<th>Canada</th>
</tr>
</thead>
<tbody>
<tr>
<td>300</td>
<td></td>
<td></td>
</tr>
<tr>
<td>250</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td></td>
<td></td>
</tr>
<tr>
<td>150</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age 65–74</th>
<th>Age 75 and over</th>
<th>Total</th>
<th>Age 65–74</th>
<th>Age 75 and over</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>146</td>
<td>76</td>
<td>222</td>
<td>101</td>
<td>50</td>
<td>151</td>
</tr>
<tr>
<td>32</td>
<td>9</td>
<td>41</td>
<td>141</td>
<td>25</td>
<td>166</td>
</tr>
<tr>
<td>100</td>
<td>100</td>
<td>200</td>
<td>100</td>
<td>25</td>
<td>125</td>
</tr>
<tr>
<td>94</td>
<td></td>
<td></td>
<td>94</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


a The two Canadian provinces are Ontario and Manitoba.

b This represents a 215 percent increase over 1981 for the United States, and a 188 percent increase for Canada.

seize on the “outliers” who become identifiable victims of resource shortages, and an atmosphere is created in which political pressure leads the funding agency to palliate the situation with visible infusions of capital and operating funds. That caveat aside, at least the palliation was relatively rapid when one considers that long queues have persisted year after year in Britain’s National Health Service (NHS). Canadian providers and government are also working together now to bring about fundamental change in the way queues are managed and resources allocated. The cornerstone reform will be a provincial information system to monitor cardiac surgery queues in all referral hospitals, so that community internists and cardiologists can consider queue length in deciding where to send a given patient. The system will assist in monitoring case-mix as suggested by standardized urgency-rating criteria. Among the proposals is that case-mix weighting of queues be used to help determine allocation of additional funds for cardiac surgery programs.

Defining the queue and its burden. The CABS crisis in Ontario has highlighted the need for better information about waiting lists. Among the issues is deciding when waiting times truly start. Ontario data thus far have tended to start the clock on the day the patient is seen by the
surgeon and the mutual decision is made for surgery. One alternative is to use the date on which an angiographic result suggests the need for a surgical consultation. Moreover, considering overall access to a system of specialized cardiovascular services, the clock could start from the time that an internist or community cardiologist first considers the patient a candidate for surgery and refers him or her for coronary angiography.

Preliminary data are available from a review of consecutive charts of patients undergoing coronary angiography for ischemic heart disease at each of four major catheterization centers in Toronto. The review began in October 1987, when the waiting list crisis intensified, and continued until at least 100 charts per center were assessed. Three of the four centers also offered revascularization procedures. Patients were tracked through medical records and procedure registries for timing of revascularization. Specialist nurses with experience in cardiovascular intensive care performed the reviews, with physician checks and extremely high agreement on chart parameters among abstracters. In Hospital A, the mean waiting time was 73.11 days (standard error, 6.71); Hospital B, 55.45 days (9.52); Hospital C, 30.46 days (4.87); and Hospital D, 64.77 days (7.86). The mean CABS waiting times were much lower than expected on the basis of reports predicated on elective cases, largely because of rapid service for patients with unstable angina, disabling symptoms, and clinical features suggesting high ischemic risk. The latter phenomenon reflects the role of the doctor as triage agent—an important safeguard against the adverse consequences of resource limitations.22

Obviously, retrospective studies do not provide dynamic data for management purposes. Similarly, cross-sectional surveys may not accurately capture the natural history of a disease. For example, the Canadian Cardiovascular Society survey of waiting lists asked respondents when “elective” patients were typically being booked for open-heart surgery. That is, if the date of initial visitation is X, and the patient is eventually booked for date y, waiting time is defined as (y - x). This approach cannot capture the inherent variability in the mean waiting time and does not allow for differing definitions of “elective.” Cross-sectional analyses are also subject to a form of prevalence/incidence bias, since the most severe cases may either die or be moved to the head of the queue. More important still, they cannot adequately capture or calculate hazard rates: risk of death, myocardial infarction, or other discrete adverse events per unit of time waiting. Thus, the need for prospective registry data becomes compelling to reflect adequately the patient’s natural history from the time of coronary angiography to the completion of the bypass procedure.

Deaths, fortunately, are rare even when waiting times are long or when large numbers of persons are awaiting procedures. From the above-noted
chart review and prospective data collected since 1989 through a regional referral office, data are available on the course of over 400 patients who were scheduled for revascularization. Less than 1 percent of patients died prior to CABS or PTCA, despite the crisis in waiting times. Since preferential streaming of urgent patients already occurs, and since ischemic heart disease remains somewhat unpredictable in its course, it seems likely that queues would have to be slashed drastically to reduce deaths much below the current levels.

On the other hand, the rarity of discrete adverse events simply emphasizes the need to delineate the burden of the queue in other ways. Most procedures are, after all, performed primarily to abate symptoms. Three issues therefore arise. First, to what extent is quality of life impaired on the waiting list? Is it simply the continuation of symptoms that could be relieved, or is it compounded by a “labeling” effect, wherein patients develop an exaggerated perception of their risks or grow intolerant of symptoms and/or drug side effects when potential relief through surgery is delayed? Some empirical data suggest the latter, but we simply know too little about the experience of persons waiting for surgery.

Second, since death or myocardial infarction on the waiting list for CABS appears to be rare, shortening patients’ duration of symptoms must be taken as a key reason for shortening waiting lists. But in that instance, it seems only fair to consider competing demands for the same limited resources. How long are the waiting lists for hip replacement, for cataract surgery, or for many other procedures undertaken primarily to enhance quality of life? The extra dollars spent to reduce waiting lists and waiting times could be weighed against potential improvements in quality of life and deaths prevented (that is, life years added), with costs per quality-adjusted life year or healthy life-year equivalents calculated and compared as is done for other programs and procedures.

Third, and as an extension of any economic analysis, one cannot ignore the hidden economic costs of waiting times. The waiting list experience may be a “disabling” one. Data exist to support the concept that the longer the waiting time for persons who are out of work due to coronary symptoms, the lower the probability that such a person will return to work and remain in the work force after CABS. Thus, a societal perspective on the economics of the queue seems appropriate, considering both present time off work and penalties for subsequent disability.

The lesson is simple: The experience of persons waiting for a health service can and should be studied, analyzed, and summarized, as is done for any disease. Just as coronary artery disease has a certain morbidity, mortality, economic burden, and natural history, so does a queue of patients seeking surgical therapy for that disease. Once one accepts this analogy, rigorous
epidemiological, psychological, and economic analyses are both applicable and mandatory. Couple this with a management-oriented view of queues, and a paradigm shift is possible.

**The doctor as triage agent.** The Toronto charts from the crisis period also show a striking correlation (p<0.0001) between individual multifactorial urgency scores and actual waiting times.\(^\text{27}\) Sorting the average waiting time by the most important determinant of urgency—patients’ symptomatic status or angina class—revealed that, on a unifactorial and aggregate basis, Class 1 or 2 patients (least severity) waited 76.32 days (standard error, 8.08); Class 3 patients, 66.4 days (7.15); Class 4a patients, 59.64 days (8.30); Class 4b patients, 33.56 days (5.90); and Class 4c patients, 7.29 days (2.09). Further analyses and second-look chart checks are still in progress. However, these preliminary findings support the concept that, even prior to adoption of explicit queue-ordering criteria, Toronto physicians and surgeons acted as effective triage agents, with a clear tendency to allocate limited resources to those in greatest need.

In a review of charts from a single center in Winnipeg, Andrew Morris and colleagues performed a similar analysis for coronary angiography rather than surgery.\(^\text{28}\) Their review was for 1981-1982 and drew on individual clinical factors rather than a composite multifactorial urgency score, but it shows that cardiovascular practitioners assign sensible priorities when resources are limited.

On the other hand, the Toronto charts reveal many instances of relatively short waits for elective cases, while more urgent cases waited inappropriately long periods of time.\(^\text{29}\) Differences in waiting times among centers were highly significant (p<0.0001) and remained so after controlling for potential differences in urgency profiles among centers (p<0.0001, with analysis of covariance procedures). The use of explicit criteria and intercenter networking to assist decision making and improve resource allocation will therefore be important if queue-based management is to be more successful.\(^\text{30}\)

Beyond scheduling, there is the issue of case selection itself. Some surgical chiefs have reported that their services used tighter case-selection criteria during 1989-1990, in part because high-risk patients with a protracted stay in the postoperative ICU frequently block beds that could be used to support surgery for other patients. This form of utilitarianism poses ethical challenges. Yet, when one considers resource constraints in health care, such policies may well be defensible, if not mandatory.

**The doctor’s dilemma: One queue or many?** While the burden of waiting falls primarily on the patient, account must be taken of the toll on providers who triage patients, treat the adverse effects of long waiting times, and face criticism from various stakeholders when adverse events
occur. Medicolegal consequences are also possible. A lawsuit is now under way in British Columbia, where the hospital and government are being sued in connection with a patient whose surgery was delayed and who died postoperatively after having become unstable on the cardiac surgery waiting list. A potential benefit of guidelines for assessing patient priority is to reduce providers’ uncertainty and medicolegal risk.

Implicit criticisms of medical conduct were made by the government inquiry that highlighted interpractitioner and intercenter discrepancies in waiting times. Yet, for the first time, Ontario’s cardiovascular practitioners were facing emergence of queue-based allocation of care in a specialty sector where facilities had been adequate to meet demand within reasonable time frames. Cardiac surgeons are trained to exacting technical standards; assuming that they should be experts in resource management is hardly sensible. Indeed, the Canadian Medicare model—public funding, but with private fee-for-service practitioners as entrepreneurs working in private nonprofit hospitals—is not conducive to a cooperative attack on waiting list problems. All surgeons clearly made efforts to stream urgent patients to the head of their personal queues; only a minority were part of group practices that routinely redistribute patients as priorities, preferences, and skills demand.

When caseload redistribution was first suggested in Ontario, some practitioners objected on the grounds that regression of waiting times to the mean was a shell game that would not actually reduce the overall suffering of patients. That is, thirty months’ cumulative wait is incurred, be it with five outpatients waiting either six months each or, say, two, three, four, nine, and twelve months. However, judging from the Ontario experience with CABS, the outliers—persons waiting nine or twelve months, and particularly the identifiable victims who experience adverse events after waiting excessive periods—are seized upon by the media and generate the perception of crisis. Indeed, on occasion, surgeons with long queues have played to these perceptions and resisted redistribution of patients on the grounds that the queues illustrate the inadequacy of government resources for cardiac surgery.

Beyond perceptions, there are more compelling reasons for trying to reduce interpractitioner and intercenter discrepancies in queues. The first is an equity concern. Why should a patient referred to a given surgeon with a long queue be obliged to wait if an equally competent practitioner down the hall can do the procedure in half the time? It is also conceivable that the risks of illness and death for patients on waiting lists are not monotonic as a function of time. Some risks may plateau, while others increase nonlinearly; some patients may adapt to their symptoms, while others develop an increasing intolerance of the same basic symptom
complex. These points add weight to the arguments for scientific study of disease processes in the queue.

Another point of resistance to redistribution was the private and fiduciary relationship between surgeon and patient, and the related concept of the cardiologists’ freedom to act as patients’ advocate in selecting the best surgeon. Yet even in the same hospital, cardiologists seldom know the relevant severity-adjusted outcomes data for each surgeon. In fact, where all are more or less equally competent, preferential streaming to the surgeon with the shortest queue is not unreasonable. And, where clear differences in outcomes emerge, then the risk of the private practice network is that the wrong surgeon will accept the complicated patient, with unnecessary risks. Only by empowering the cardiac surgery service chiefs to play a larger role in assigning cases (and managing queues) can these risks be avoided.

For redistribution between institutions, the Ontario model will be to facilitate regional cross-referral by providing cardiologists with more information about variation in waiting times for differing categories of surgical patients. But unless reallocation within institutions is the norm, these waiting list data may need to be provided by center rather than by surgeon, for two reasons. First, since long waiting lists are often taken as indicating surgical prowess, some surgeons might be embarrassed by publication of their (short) waiting times. The corollary is that referring physicians could respond paradoxically—a phenomenon perhaps analogous to the “prestige good” in economic theory, where raising the price is reputed to increase demand by creating a perception of better quality.

Finally, one returns to the concept of proper forward planning. If some redistribution of patients were guaranteed within the same center, operations researchers could contribute to better management of queues and even more rational use of operating theaters. It is almost axiomatic that a single queue will be more efficient than multiple queues. In fact, prospective data accumulated during 1989-1990 at our regional referral office show that if the referring doctor prespecifics a surgeon of choice and thereby sidesteps attempts to find the most readily available slot, patients wait significantly longer for CABS—an average of ten days (p<0.0001 after controlling for differences in urgency scores). These differences are more notable because the majority of patients referred to the central office had urgent indications for surgery or PTCA.

**Implications For Reform**

In an ideal world, appropriate health services would be available without charge or delay. The reality, however, is that allocative decisions
must be made. Students of health policy often view Great Britain and the United States as the two poles for allocative approaches. With over a million Britons awaiting hospital procedures as of late 1990, the NHS queues are probably the largest and longest of any nation. It is arguable that queue-based allocation in Britain is no more unfair than is rationing by price and income in the United States, particularly since the growth of a parallel British private sector adds new elements of rationing by income. Yet to couch the argument in such terms is to presuppose that Britain and the United States are somehow the competing paradigms for allocation of services. Britain ranks near the bottom of the list for per capita expenditures on health care in industrialized nations. Conversely, the United States combines allocation by income and insurance and ranks highest in per capita expenditures on health care in the world. Clearly, the goal must be to find a middle ground between these two extremes.

For years, Canadians prided themselves on a health care system that appeared to be just such a middle ground: public funding and prepayment administration, largely private and decentralized arrangements for delivery of services, and first-dollar coverage that seemed likely to enhance equity. However, the system's success can be questioned on three levels. First, it is expensive. Although Canadian per capita expenditures on health care are about $700 (U.S.) lower than U.S. expenditures, Canadian costs per capita are the second-highest in the world. Second, equity goals have not been fully achieved. With first-dollar coverage and nominal abolition of price rationing, one would like to see evidence of leveling out of differentials in health service consumption by socioeconomic class or even a positive bias in favor of the poor due to their greater burden of illness. Although some studies are positive, the evidence is far from conclusive and suggests partial success at best. Furthermore, major differences persist in life expectancy and health status by socioeconomic class. Third, waiting lists now exist for a plethora of services and procedures. Apart from high-profile services such as cardiac surgery, adequate study and documentation of these queues have been lacking. Nonetheless, one has only to practice on a very part-time basis (as does this internist) to run into the queuing phenomenon in major cities for diagnostic and therapeutic services.

Denial versus delay. If queues exist in an expensive, publicly administered system such as Canada's, additional point is given to the debate between the “rationalizers” and the “rationers.” The former group argue, in a nutshell, that systematic reform of health care can permit all citizens of a given industrialized country to receive all necessary services at an overall cost that most societies can shoulder without undue impingement.
on other public and private expenditures that contribute to quality of life. The latter argue either that such reforms cannot be implemented or that, if implemented, they will not suffice to avert the need for some persons to be denied some services that would prolong or improve their lives.

Framing the debate in these terms helps move one beyond the spurious dichotomy of British versus American paradigms and instead raises the question: What combination of private and public financing and administration will best promote both efficiency and equity, and in a way that minimizes the amount of suffering from denial of potentially useful services? Regardless of how one answers that question, it could be argued that some elements of allocation by queue are defensible. Indeed, if rationing is understood to be an implicit or explicit mechanism whereby services are denied to some persons who would benefit from them, then queues may not necessarily be a form of rationing. In the United Kingdom, waiting lists have become less an explicit strategy of efficiency-enhancing delay than a tool to-palliate an implicit policy of denial. Yet in theory, all a queue should do is put off the time that a service is delivered.

As noted above, with limited resources and waiting lists, physicians and surgeons act as triage agents, ensuring that those with the most pressing needs receive attention first. A policy of managed delay may therefore have advantages over systems of price rationing, which are based on the demonstrably questionable axiom that consumers can ration their own demands according to tradeoffs between perceived need and deterrent effects of a user fee or other price to be borne for the service.\textsuperscript{38}

Waiting lists do carry a political price. Persons dying on waiting lists are truly identifiable victims. The risk, then, is that resources will be infused preferentially to services with these high-profile deaths, such that many other victims, identified and otherwise, are created by resource shifts and by the opportunity costs of public dollars no longer available for other socially beneficial purposes. The sobering cost/utility analyses suggested above are therefore critical to rationalization of queue-based allocative mechanisms.

**Excess capacity versus queues.** On the other hand, the efficiency-enhancing aspects of cooperative queuing for specialized services are obvious. No queues imply excess capacity. That is, one could only offer more or less immediate surgery to all cases, elective and urgent, if there were idle operating rooms and surgical staff, empty intensive care unit and ward beds, and so on. Indeed, delay may be beneficial for some elective cases. It offers the patient time for sober second thoughts, including a further trial of medical therapy and time to put affairs in order, given the minor risk of death at surgery and the certainty of some postoperative recovery time. Lastly, there is an enormous variability in decision thresholds
among practitioners. Even if volume-related income incentives were reduced or eliminated, it seems plausible that cardiologists and surgeons would interact to lower the thresholds for surgical intervention if unused capacity was available. Combine excess capacity with fee-for-service payment of both surgeons and hospitals, as in the United States, and it is no surprise to find that U.S. CABS rates are extraordinarily high.

The lesson is plain. Dismissing health service queues as some inevitable and evil by-product of “socialized medicine” is as fallacious as the argument that U.K.-style queue-based rationing, with delays for hundreds of thousands of citizens, is automatically legitimate because it occurs in a universal system with appearances of egalitarian access. The respective burdens of denial versus delay must be measured, and societal values—above all, the interplay between liberty, equality, and equity—must be taken into account. The key issues are measurement and management, using explicit criteria and sound methodology.

The author acknowledges financial support from Career Scientist Award 02377 of the Ontario Ministry of Health, from Sunnybrook Research, and from the Department of Medicine of the Sunnybrook Health Science Centre, Toronto, Ontario. I am grateful to many Ontario cardiovascular surgeons and cardiologists and to Ministry of Health officials for their frank discussion of these issues and comments on the manuscript. Special thanks for advice from Ronald S. Baigrie and Bernard S. Goldman, chiefs of cardiology and cardiovascular surgery, respectively, and many colleagues in the clinical epidemiology program at the Sunnybrook Health Science Centre. Preliminary analyses of the chart review data were performed by Carey Levinton, biostatistician at the Sunnybrook Clinical Epidemiology Unit (full manuscript in preparation); chart reviews were performed by Linda Hunter, Susan Wheeler, and Carol Klupsch.

NOTES

3. CABS accounts for about 75 percent of the provincial open-heart surgery caseload. The remainder are valve procedures and adult congenital heart lesion repairs. Some valve procedures are accompanied by incidental bypass.
4. As an example, let us say that a CABS center has been doing about 500 cases per year. An expansion grant is made for renovations, and operating funds are provided to increase caseloads to 750 per year. In the third year after expansion, the program does 800 cases. The hospital then receives post hoc reimbursement for the extra fifty cases. In the fourth year, there are 650 cases because the hospital reallocates some operating room time to hip replacements and a cardiac surgeon moves away. The per case overrun funding is lost, as is only fair, but there also is no penalty for going under target by 100 cases. Clearly, then, the ministry has set up a situation in which it has given away a large part of its potential fiscal leverage.
6. Ibid.
15. Ibid.
16. Canadian Cardiovascular Society Questionnaire on Cardiac Surgery in Canada (Mimeo, 1990). Send requests for information to Lyall Higginson, CCS Secretary, Heart Institute, Ottawa Civic Hospital, 1053 Carling Avenue, Ottawa, Ontario, Canada K1Y 4E9.


27. Naylor et al., “Forming the Queue for Coronary Surgery in Toronto.”


29. Teaching requirements are a hidden factor in scheduling, since trainees should ideally face a reasonable mix of cases to enhance their adaptability. Also, stable cases may be slotted in simply to avoid a situation of unremitting pressure on surgeons, residents, and nurses.


32. Kaminski et al., “Investigation of Cardiac Surgery at St. Michael’s Hospital.”


35. R.F. Badgley and S. Wolfe, “Equity and Health Care” (Working paper, Department of Behavioural Science, University of Toronto), see particularly 20-31. Request copies from R.F. Badgley, Department of Behavioural Science, University of Toronto, Ground Floor, McMurrich Building, 12 Queen’s Park Crescent West, Toronto, Ontario M5S 1A8.

