A FINGER ON THE PULSE:
Comparative Models for Reporting the Quality of Hospital Care

JULIA WITT

June 2005
The Atlantic Institute for Market Studies (AIMS) is an independent, non-partisan, social and economic policy think tank based in Halifax. The Institute was founded by a group of Atlantic Canadians to broaden the debate about the realistic options available to build our economy.

AIMS was incorporated as a non-profit corporation under Part II of the Canada Corporations Act and was granted charitable registration by Revenue Canada as of October 3, 1994; it recently received US charitable recognition under 501(c)(3) effective the same date.

The Institute’s chief objectives include:

a) initiating and conducting research identifying current and emerging economic and public policy issues facing Atlantic Canadians and Canadians more generally, including research into the economic and social characteristics and potentials of Atlantic Canada and its four constituent provinces;

b) investigating and analyzing the full range of options for public and private sector responses to the issues identified and acting as a catalyst for informed debate on those options, with a particular focus on strategies for overcoming Atlantic Canada’s economic challenges in terms of regional disparities;

c) communicating the conclusions of its research to a regional and national audience in a clear, non-partisan way; and

d) sponsoring or organizing conferences, meetings, seminars, lectures, training programs, and publications, using all media of communication (including, without restriction, the electronic media) for the purpose of achieving these objectives.

Board of Directors
Chair: David McD. Mann; Vice-Chairs: John F. Irving, Peter C. Godsoe, John C. Walker
Chairman Emeritus: Purdy Crawford
President: Brian Lee Crowley

Advisory Council

Board of Research Advisors
Chair: Professor Robin F. Neill, University of Prince Edward Island
Professor Charles S. Colgan, Edmund S. Muskie School of Public Service, University of Southern Maine; Professor Jim Feehan, Memorial University of Newfoundland; Professor Doug May, Memorial University of Newfoundland; Professor James D. McNiven, Dalhousie University; Professor Robert A. Mundell, Nobel Laureate in Economics, 1999; Professor David Murrell, University of New Brunswick

2000 Barrington Street, Suite 1006, Halifax, Nova Scotia B3J 3K1
Telephone: (902) 429-1143; fax: (902) 425-1393
E-mail: aims@aims.ca; web site: www.aims.ca
A FINGER ON THE PULSE:
Comparative Models for Reporting the Quality of Hospital Care

JULIA WITT

Atlantic Institute for Market Studies
Halifax, Nova Scotia

June 2005
Julia Witt has a PhD in Economics from the University of Guelph. She earned an Honours BA in Economics and Environmental Management at the University of Toronto and an MA in Economics at the University of Guelph. Her fields of specialization include health economics, social welfare analysis, applied microeconomics, and cost-effectiveness analysis.
Reporting on the quality of care that patients receive in Canadian hospitals is the responsibility of governments, hospital boards, and professional accreditation agencies. Accreditation and quality measurement of hospitals and health systems are controversial issues, but the popularity of and demand for such measurement are growing — now usually with the voluntary, but in the future perhaps the compulsory, participation of hospitals.

Unfortunately, there is a lack of understanding about many accreditation and quality measurement issues. Problems exist with the quality and usefulness of available information, the lack of standardized measures, and a sense that hospitals all too often attempt to manipulate the numbers in order to look better in comparative rankings rather than institute genuine reforms that might improve patient outcomes. Many reports are also oriented more toward hospital management than to other stakeholder groups in the health care community.

This paper looks at the US, European, and Canadian experience with accreditation mechanisms and quality assessment systems, to give the reader a sense of the complexity of quality assessment issues. Indicators — arguably the most important measure of quality from the perspective of patients — are the focus of a second paper in this series, to be followed by the release of the Atlantic Institute for Market Studies’ own Hospital Report Card.

Accreditation agencies are broadly of two types. First, there are organizational quality assessment schemes, which look at whether the service does what it is supposed to do and which assess the quality of a health care service, but with a focus more on the provider than on the consumer. Examples of such schemes include the European Foundation for Quality Management, the International Organization for Standardization, the Joint Commission on Accreditation of Healthcare Organizations (JCAHO, a US body), and the Health Quality Service (a UK body formerly known as the King’s Fund Organisational Audit). Second, there are comparative quality assessment schemes, which shift the focus to the consumer, or patient, and compare indicators of the quality of service both to other services at the same time and to the same service at another time. Examples include the Maryland Hospital Association’s Quality Indicator Project, the Pennsylvania Health Care Cost Containment Council, and the indicator scheme run by JCAHO.

Several different types of measures can be used to determine the quality of a hospital. One type, structural measures, look at issues such as the conditions under which care is provided, the equipment used, the quality of the staff, and the credentials of clinicians. Such measures are inconclusive and difficult to interpret, however, and the direction of causality is often lacking. Process measures describe the content of health care and physicians’ treatment of their patients. Here, however, there are issues concerning treatment guidelines and the appropriate treatment that patients should
receive. Finally, outcome measures look at changes in the patient’s health that occurred because of the treatment received. Recovery, as a measure of health care quality, is a concept both consumers and providers easily understand. Outcome measures, however, are complicated by the patient’s health status and by side-effects that might not reflect the quality of care the patient received.

Whatever indicators are used to measure quality, they should be comparable, in order to avoid ambiguity as to what they represent and to set common criteria that different institutions can use. In Canada, some progress was made toward establishing such criteria with the implementation by the Canadian Institute of Health Information of a (flawed) set of 14 quality measures in the areas of health status, health outcomes, and quality of health care services. In the United States, organizations such as JCAHO are concerned with the measurement of health care quality, but such accreditations are often based on little more than guidelines, which may not be appropriate for all patients. A shortage of reliable empirical studies of hospital quality indicators has also hampered the emergence of comparable standards of measurement outcomes in many US and Canadian hospitals. The United Kingdom has experimented with the use of controversial “league tables” to compare the performances of its hospitals. Proponents argue that league tables encourage hospitals with low rankings to improve and consumers to seek out the best hospitals. Opponents charge that rankings can be manipulated by the choice of indicator, such as mortality rates, or even by a hospital’s choice of patients, and that it is difficult to compare health care providers over time.

Finally, another way of measuring the quality of hospital care is to shift the focus from indicators to information about “adverse events” such as medical errors and drug reactions. Estimates of the incidence of such events vary widely, but they are blamed for tens of thousands of deaths each year in the United States; the problem of preventable adverse events is a large one in Canada as well. In addition, better statistical methods need to be adopted to measure and reduce such errors.

Consumers need to be able to compare the quality of health care provided by different hospitals if they are to make an informed choice. Thus, the quality and usefulness of the measures used to make those comparisons must be high, reliable, and, above all, relevant to patients. Finding the right mix will help establish quality assurance and comparability among health care institutions, to the benefit of all stakeholders.
Reporting on the quality of care that patients receive in Canadian hospitals is the responsibility of governments, hospital boards, and professional organizations such as the Canadian Council on Health Services Accreditation. The Canadian Institute for Health Information (CIHI) formally gathers and reports comparative information, which may be in whole or in part used by the various organizations responsible for such reporting.

Although much controversy surrounds accreditation and quality measurement of hospitals and health systems, such measurement is undertaken increasingly frequently — usually with the voluntary participation of hospitals. It is not unthinkable that, in the near future, all hospitals will have to take part in some form of external assessment. Before such a requirement comes into force, however, accreditation and quality measurement need to be studied in much more detail. Although demand is increasing for the information that external assessments provide, studies that support the validity and comparability of these reports have lagged behind. As a result, accreditation and quality measurement surveys are not well standardized and there is no mechanism to ensure the reports provide useful information. Indeed, the Auditor General of Canada and several provincial auditors general have noted their inability to comment on the quality of information generated by CIHI, due to inadequacy in the quality assurance processes that ensure the data are accurate, a lack of data standards and a lack of data definition for certain indicators. There is evidence, in fact, that such reports occasionally may be counter-productive. Instead of encouraging “low-quality” hospitals to improve, they may lead institutions to manipulate the numbers to obtain better ratings without improving outcomes. In this sense, such reports are misleading, particularly since a large part of their audience may rely solely on the “bottom line”, without knowledge of potential biases.

In addition, although a tremendous amount of information is produced, the usefulness of much of it for patients is at best questionable, since the outcomes of these studies are oriented toward management and, often, efficiency issues that result in cost savings, but not necessarily toward better outcomes for patients. Having said that, while patients and the general public are important stakeholders, interested in the relative safety and quality of hospitals, waiting-time information, the incidence of adverse events (especially preventable ones), and the likelihood that a particular intervention will produce a benefit, they are not the only stakeholders. Others include governments (as insurers, administrators and regulators), statutory bodies, health care providers, private insurers, membership societies, professions,1 and consumer organizations (Shaw 2000). These stakeholders are interested

---

1 Membership societies and professions include any corporate stakeholders that are concerned with external quality mechanisms.
in health policy, resource allocation decisions, registration issues, management of hospitals, funding, quality improvements, self-regulation, and public information. Identifying for whom a particular report will be useful and how to define usefulness for that particular audience is thus the first important step in ensuring good reporting on quality issues.

This paper, the first in a series, outlines the accreditation mechanisms and quality assessment systems currently in use in the United States, Europe, and Canada, and briefly illustrates some of the problems associated with each. The purpose of the paper is simply to familiarize the reader with these mechanisms and to give a sense of the multitude of aspects that must be considered in quality assessment. Indicators, arguably the most important measure of quality from the perspective of patients, are discussed in only a very general sense here; the second paper in this series will focus entirely on these measures. Together, this review of quality assessment systems and the second paper’s in-depth look at indicators will set the stage for the release by the Atlantic Institute for Market Studies of a Hospital Report Card.
The first step in assessing whether the quality of a hospital is “good” or “poor” is to define what encompasses “quality” and how to measure it. Different components contribute to the quality of hospital care a patient receives, including the quality of the equipment being used, the ratio of patients to nurses and physicians, whether the patient receives the right diagnosis and treatment, and how long the patient must wait for diagnosis. Placing these quality measures into different categories makes it easier to identify who will benefit from the information contained in these measures and how useful they are in measuring quality in the first place.

Romano and Mutter (2004) describe three broad measures of quality. The first measure is structural, which looks at the conditions under which care is provided, including such factors as the equipment being used, the quality of the staff, the credentials of clinicians, and so on. It is relatively easy to quantify these measures: counting the number of magnetic resonance imaging (MRI) machines or assessing clinicians’ credentials is fairly straightforward. The findings are, however, inconclusive and difficult to interpret, and there is no evidence of the direction of causality. Do better structures lead to better outcomes or do better outcomes lead to better structures? Are more MRI machines necessarily associated with better health outcomes or just with more effort and more worry from false alarms? Do patients choose hospitals with better outcomes, and if so, how can patients identify such hospitals? Are hospitals with many patients able to purchase better structures?

Romano and Mutter’s second measure of quality relates to process, which describes the content of health care and reveals how physicians treat their patients. The information process measures provide is useful because it reflects the intensity of care a patient receives — for example, after surgery, are proper medications prescribed and are regular follow-up appointments scheduled? Such data are costly to collect, however, since they require reviewing individual medical charts to assess whether the treatment provided was the best possible. Additionally, there is not always agreement on what treatment is “best” for a particular individual. Many factors — such as co-morbidities, age, weight, and allergies, among others — have to be taken into consideration, and not all patients qualify to receive the “standard” treatment that is appropriate for most other patients.

Another concern is that guidelines (rather than standards) are meant to apply only to a portion of patients, and that conceptual problems exist in converting practice guidelines into quality measures. For example, Walter et al. (2004) find that the federal Department of Veterans’ Affairs’ adherence to performance measures in its assessment of how it screens for colorectal cancer is vulnerable to

---

2 Co-morbidities are diseases that the patient may have in addition to the condition for which he or she was initially diagnosed.
several pitfalls, including the selection of the appropriate target population, the determination of target screening rates, and the measurement of screening performance. Yet Veterans’ Affairs relies heavily on these guidelines and exerts pressure on medical centres that fail to meet its target screening rate, to the point of penalizing financially those that repeatedly fall short, thereby potentially compromising the discretion that is expressed at the referral of individual patients. Although health administrators often rely on target screening rates, patients would clearly be better served if an organization could say that 100 percent of its procedures are done appropriately and for appropriate indications.

The third type of quality measure that Romano and Mutter describe relates to outcome, or the changes in a patient’s health that are a direct consequence of the treatment received. The biggest advantage of outcome measures is that they quantify what is important to the patient — namely, recovery. Moreover, such measures are easily understood by both consumers and providers to mean the attempt to improve comfort, function, and life span, and to provide information about the patient’s health status. Outcome measures may be difficult to interpret, however, because expected outcomes vary depending on how sick a person was in the first place, and health organizations do not routinely and systematically capture information about health status before and after treatment. Hence, outcomes are affected not just by the quality of treatment, but also by factors not related to treatment quality. Outcome measures, therefore, are not always a good measure of quality — a “side-effect” that is often obscured by the reports that use or endorse them.

Although ambiguities in measuring outcomes or making diagnoses can be a clear disadvantage for hospitals that administer to very ill patients, there are also concerns that hospitals can use these measures to work in their favour. For example, in 1983, the US government switched from a cost-based reimbursement system to a prospective payment system, in which hospitals are reimbursed according to the patient’s diagnosis rather than to the treatment provided. In response, hospitals changed their coding of conditions, evidence of which Cutler (1995) discovered in changes hospitals made to the timing of deaths and increases in readmission rates. A study of Mount Sinai Hospital in Toronto also found that a move to concurrent coding from conventional chart coding3 changed the resource intensity weighting of patients on a particular unit, even though there was no change in clinical or administrative behaviour (Odell and Young 2001).

**Defining Indicators**

The problem with poorly defined indicators and outcome measures is that their flexibility permits statistics to be manipulated to render a favourable verdict. Finding a clear definition of “indicator” is, however, difficult.

---

3 Under “concurrent coding”, patients’ charts are coded in the relevant hospital unit while the patient is still in the hospital. Under “conventional coding”, charts are coded after the patient has left the hospital, usually in the hospital’s coding room by staff who have little or no communication with clinical staff.
One important characteristic that an indicator should possess is comparability. When indicators are comparable, there should be no ambiguity as to what they represent since all providers define them according to the same criteria. Øvretveit (2001, 231) defines an outcome indicator as “a quantified representation of any change in a person’s health status before and after receiving care, which may be caused by the care they received.” Although this is a very general definition, it clearly states that the indicator should quantify the change in a person’s health, which takes into account the health status of the individual before treatment. In other words, they measure the amount of something per something.

In health, there are also sentinel indicators, which are adverse events that should never occur and would be preventable if proper systems were in place. In environmental studies, sentinel indicators typically consist of a species that is quickly and severely affected by pollution, signifying that environmental damage has been done. The drawback of sentinel indicators is that adverse events have to happen before one becomes aware that they might take place and takes measure to prevent a recurrence.

**Box 1: CIHI Indicators**

<table>
<thead>
<tr>
<th>Health Status Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Life expectancy</td>
</tr>
<tr>
<td>(2) Infant mortality</td>
</tr>
<tr>
<td>(3) Low birth weight</td>
</tr>
<tr>
<td>(4) Self-reported health</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Health Outcomes Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>(5) Change in life expectancy</td>
</tr>
<tr>
<td>(6) Improved quality of life</td>
</tr>
<tr>
<td>(7) Reduced burden of disease, illness, and injury</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quality of Service Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>(8) Waiting times for key diagnostic and treatment services</td>
</tr>
<tr>
<td>(9) Patient satisfaction</td>
</tr>
<tr>
<td>(10) Hospital re-admissions for selected conditions</td>
</tr>
<tr>
<td>(11) Access to 24/7 first-contact health services</td>
</tr>
<tr>
<td>(12) Home and community care services</td>
</tr>
<tr>
<td>(13) Public health surveillance and protection</td>
</tr>
<tr>
<td>(14) Health promotion and disease prevention</td>
</tr>
</tbody>
</table>

Source: See the CIHI web site: <http://secure.cihi.ca/cihiweb/dispPage.jsp?cw_page=pirc_e>.

In September 2002, Canada’s First Ministers agreed to the provision of clear accountability in reporting to Canadians on the performance of provincial health systems. Following that agreement, CIHI implemented a set of 14 quality measures in the areas of health status, health outcomes and quality of health care services, with the goal of achieving comparable reporting (see Box 1). Each indicator has a particular measure associated with it. For example, change in life expectancy (indicator 5) is measured as age-standardized mortality rates for lung, prostate, breast, and colorectal cancer, while hospital readmission for selected conditions (indicator 10) is measured by readmissions for acute myocardial infarction (AMI) and for pneumonia. In its report on these indicators, Health Canada notes that those on health outcomes “attempt to track the effects of policy, program or clinical interventions on quality of life” (Canada 2002).
Unfortunately, however, these new indicators are still flawed. For example, readmission to hospital does not necessarily reflect quality of life, whether long or short term. Moreover, readmissions are higher for sicker patients, thus potentially underestimating the quality of care of a hospital that, by chance, happens to have relatively more of such patients. And, as Health Canada itself admits, these indicators are difficult to use without a baseline and a systematic approach of reporting results. Moreover, the audits that each province and the federal government issues on reports on the 14 indicators show that currently available quality assurance processes to support the indicators are inadequate, with a lack of accurate data, data standards and even data definition for certain indicators. For example, health outcomes include a measure of the frequency of hip and knee surgery without any associated information about how the surgery improved or impaired the individual’s quality of life.

The topic of indicators involves many other aspects. These, however, are the focus of another paper in this series. For now, this brief introduction will suffice to carry the discussion of indicators throughout the rest of this paper, the focus of which is methods for comparing outcomes and the agencies responsible for carrying out comparisons.
In the United States, a number of different organizations are concerned with the measurement of health care quality, including the National Committee for Quality Assurance, the Joint Commission on Accreditation of Healthcare Organizations (JCAHO), the National Quality Forum, and the Centers for Medicare and Medicaid Services. The first two are provider oriented and thus focus on process measures; the latter two bring together measures from the perspectives of both providers and consumers/employers. Together, these organizations are working toward consensus on a growing list of quality indicators for in-patient care, depending on the condition for which patients require a hospital stay. For AMI patients, for example, measures include “Aspirin at discharge”, “Beta-blocker at arrival (within 24 hours)”, and “in-hospital mortality (risk-adjusted)”. The advantage of such indicators is that they can be used by individual studies; indeed, JCAHO requires hospitals to use some of these indicators in order to receive accreditation.

In issuing hospital accreditations, JCAHO looks at three calculations that affect a hospital’s rating (JCAHO 2004). First, it determines the number of times a hospital has met a particular measure — such as giving Aspirin to heart attack patients — divided by the number of opportunities the hospital had to use that measure. Second, JCAHO expresses the results of the first calculation as a percentage, and calculates 99 percent confidence intervals around the national average. These computations show how significantly different a particular organization’s results are from the national average and whether or not the organization measures up. Finally, to provide consumers with information “at a glance”, JCAHO aggregates the results of individual measures (such as the dispensing of Aspirin at arrival and at discharge) to arrive at an overall measure of patient care for a particular condition (such as heart attack care). In doing so, the commission uses “roll-up” methodology, which sums the number of opportunities that the hospital had to meet the performance expectations, divided by the number of times such expectations were actually met.

JCAHO’s accreditations, however, are based on indicators that might more appropriately be viewed as guidelines rather than as outcome measures. Aspirin, for example, may not be appropriate in all circumstances, but in the accreditation procedure, there is no “appropriateness measure” that captures this. It would be interesting to look at how many heart attack patients received Aspirin even though they should not have — that is to say, by highlighting the importance of this particular performance measure, is JCAHO encouraging hospitals to give Aspirin to all heart attack patients even when it is unnecessary?

At the same time, indicators of the kind JCAHO uses in its accreditation process do have other useful purposes. In a survey of the US literature on the use of indicators of hospital competition and quality, Romano and Mutter (2004) find that AMI is the most frequently analysed condition and that
mortality and readmission, in that order, are the two most commonly used outcome measures. The advantage of studying AMI patients is that, since they usually need immediate care, such patients do not choose their medical care facility on the basis of quality, so there is no selection bias on the part of patients. On the other hand, this same feature may make hospitals complacent in improving AMI care — since patients do not choose their facility based on quality in any case, there is no reason to enhance this service to attract more patients. At the same time, Romano and Mutter argue, since the infrastructure needed to treat AMI patients is used for other elective procedures as well, it is probably well maintained and frequently upgraded, making AMI mortality a potentially unbiased overall proxy for hospital quality.

There is, however, a shortage of empirical studies of hospital quality indicators and some of the few that do exist are of questionable validity. For example, many studies fail to take account of the possibility that hospitals can “doctor” their numbers to produce more favourable outcomes: in-hospital mortality can be manipulated by the simple expediency of releasing moribund patients and transferring the record of their deaths to outpatient settings. A report on hospitals in western Pennsylvania, for example, indicates significantly higher readmission rates for stroke victims and decreases in average lengths of stay for many procedures and types of treatment from 2002 to 2003 (PHC4 2004), which is suggestive of attempts to manipulate outcomes. Such trends have been observed in Canadian hospitals as well, but research is hampered by the lack of formal Canadian studies relating the effect of decreasing time spent in hospital and decreasing number of hospital deaths. Clearly, more solid empirical evidence is necessary on the use of these indicators of hospital care in order to isolate problems.

Many other accreditation agencies also exist, and these agencies use different schemes to assess organization, management, and quality. Øvretveit (2001) classifies these agencies under two broad definitions. First, there are organization and management assessment schemes, which look at whether the service does what it is supposed to do and which assess the quality of a health care service; these include the European Foundation for Quality Management, the International Organisation for Standardisation (ISO 9000), JCAHO, and the Health Quality Service (a UK body). Second, there are schemes that collect and compare indicators of the quality of service both to other services at the same time and to the same service at another time; examples include the Maryland Hospital Association’s Quality Indicator Project, the Pennsylvania Health Care Cost Containment Council, performance indicators used in the United Kingdom, and the JCAHO indicator scheme.

Organizational Quality Assessment Schemes

The use of organizational quality assessment schemes is voluntary in most of Europe and in Canada; in the United States, however, hospitals must have JCAHO accreditation in order to obtain federal funding for patients. These schemes focus on whether the service has the resources, structures, and processes that experts deem to be necessary to produce good-quality outcomes and experiences for patients. But does a good organizational structure result in better outcomes or do “good” hospitals
attract more patients and hence can pay for better organizational structures? While these schemes guarantee only that a hospital meets specific requirements, not that high-quality care is provided at the same time, and there is no basis for comparing hospitals to one another, there are clear advantages to the information they provide. Knowing which hospitals meet “minimum” standards is useful for patients, and it provides an incentive to hospitals to meet these standards. The following sections examine some of these assessment schemes in more detail.

ISO Certification

ISO 9000 is a widely applicable international standard for quality systems, first introduced in 1987. Since then, the ISO system has been revised twice, largely because the early versions did not place much weight on the final product. The latest revision, ISO9001:2000, places more emphasis on the consumer and is less partial to manufacturing companies.

ISO9001:2000 is based on eight quality management principles: customer focus, leadership, involvement of people, process approach, system approach to management, continual improvement, factual approach to decisionmaking, and mutually beneficial supplier relationships (Moullin 2002). There is a cyclical relationship between these aspects, with key inputs and outputs being defined by the customer.

In order to obtain ISO 9001:2000 certification, an organization must pass an initial evaluation, and, to maintain certification, continue to pass unannounced inspections. As part of the certification process, the organization must meet requirements in five areas:

- It must have developed and implemented a quality management system, with documented procedures and process interactions.
- It must implement a process of management responsibility, which includes support and involvement of senior management for the quality systems, meeting customer demands, and clearly defining responsibilities and authorities.
- Its resource management must ensure a capable staff, well-maintained buildings and equipment, and management of the working environment.
- Its product and/or service realization must ensure continued evaluation and upgrading of production processes and customer satisfaction, and that the monitoring devices are accurate.
- It must implement procedures for measurement analysis and improvement, which guarantees back-up plans if something goes wrong, and it must have plans for improvements.

The advantages of ISO certification are numerous, but one of the most important is that organizations are forced to maintain standards of quality. Moreover, since certification also involves substantial detail in the quality process, the organization itself plays a major role in determining explicit descriptions of its objectives.
Staines (2000) examines ISO 9001:2000 certification at a regional hospital in Switzerland and finds that it was forced to deal with philosophical and operational quality issues. In other words, as part of the certification process, the hospital was obliged to deal with issues that might otherwise have been considered less important, such as keeping records up to date. Certification also forced the hospital to remain in compliance with new legislation, even if other health care facilities lagged behind. Another perceived benefit of ISO certification was that the establishment of meticulous maintenance procedures and documentation made it easier to trace the cause of problems and helped staff, particularly newcomers, to determine what role each of them played in the overall quality process. Finally, satisfaction on part of hospital management and staff of having obtained a prestigious and internationally recognized credit motivated them to maintain the required quality standards.

The drawback of such “benefits”, however, is that they focus mostly on process, and fail to provide what Staines calls the “cultural dimension”. ISO certification is a very technical and formal process, but such assessment systems fail the health care sector by being provider oriented, not consumer or patient oriented. Although it is important to ensure that physicians and staff also believe in a quality system and deliver quality care in a quality-certified environment, new procedures, meticulous record keeping, and regulations alone do not guarantee that the patient receives appropriate care.

The European Foundation for Quality Management

The European Foundation for Quality Management represents another important quality assessment model. The rationale behind the EFQM model is “Excellent results with respect to Performance, Customers, People and Society are achieved through Leadership driving Policy and Strategy, People, Partnerships and Resources, and Processes.” The model is based on the idea that striving for excellence in one different aspect of an organization’s activities does not have conflict with the goal of excellence in other areas — a hospital can provide excellent clinical care, excellent financial performance, excellent employee motivation, and so on.

The EFQM model uses nine elements to assess an organization’s progress toward excellence: five “enablers” (leadership, people, policy and strategy, partnerships and resources, and processes) and four “results” (people results, customer results, society results, and key performance results). An organization that adopts the EFQM model would start by identifying the results; it would then evaluate how well it was doing based on performance criteria. In doing so, it would identify areas in need of improvement and avoid simplistic solutions.

Staines (2000) looks at how the EFQM model was implemented at a hospital in the United Kingdom and notes that, among its advantages, it provided a common language for a comprehensive performance management system, a common framework, quality improvement initiatives, and the opportunity to involve all staff in setting priorities. Importantly, the model has proved to be beneficial not

4 See the web site: <http://www.efqm.org>.
just for the providers but also for patients in terms of improved outcomes. The South Tees Acute Hospitals National Health Service Trust, for example, determined that although its colposcopy unit was efficient from a clinical perspective, with approximately 750 referrals per year, staff and patients were frustrated with how the system worked. In implementing the EFQM model, the Trust got together all staff involved to review and redesign the process with the result that the time needed to generate an appointment was reduced from 13.5 days to 24 hours for the general practitioner and 48 hours for the patient, average consultation time was increased from 10 minutes to 30 minutes, and the missed appointment rate was reduced from 20 percent to 10 percent.

Unfortunately, however, such efficiency improvements do not necessarily guarantee high-quality care for patients. The problem with organization quality assessment schemes such as the ISO and EFQM models is that they were originally developed for use in an industrial setting, with quality standards that focus on for example, infrastructure and equipment. It is not easy or even valid to transfer such provider-oriented approaches to health care, with its emphasis on the delivery of quality care and health outcomes. For example, Moeller (2001) compares the EFQM approach in industrial settings to its use in hospitals in Germany. He finds that, while the maximum score that an organization can earn under the EFQM approach is 1000, more than 50 percent of German hospitals scored between 200 and 300 points, and none scored higher than 450; in contrast, German industrial organizations scored as high as 750 points.

### Comparative Quality Assessment Schemes

One big problem with organizational quality assessment schemes is that they focus mostly on the provider and guarantee little to the consumer. They do this by design, because the standards that define quality focus on requirements around, for example, infrastructure and equipment, not on the delivery of quality care, which is measured by health outcomes. To shift the focus to the consumer, the measure of quality has to reflect this goal; moreover, since quality is easier defined in relative terms, the measurements need to be comparable.

In contrast to organizational quality assessment schemes, comparative quality assessment schemes shift the focus to what matters to the consumer/patient. To achieve that end, these schemes use various types of quality indicators as units of measurement, including such proxy measures as ratios of nurses to patients and such direct measures as waiting times, medication errors, and so on. In addition, the choice of indicators is based on who will use the data and for what purpose. The indicators are then used to make comparisons in the areas of research, clinical improvement, support referrer and patient choice, resource management, and transparency.

### Examples of Comparative Quality Assessment Schemes

In 1985, seven US hospitals started the Maryland Hospital Association Quality Indicator Project, whereby they agreed to collect ten indicators of in-patient care — including rates of hospital-
acquired infections, surgical wound infections, mortality, caesarean sections, and unplanned readmissions — in order to compare services among them. Today, the project includes more than 1000 acute care hospitals and other health care facilities.

Another example of a comparative quality assessment scheme is the Nordic Registries. In Norway, quality of care in hospitals is assessed according to five indicators: infection rates, turnover rates of health workers, the length of time from discharge to the mailing of a discharge letter to the general practitioner, rates of cancelled elective operations, and rates of adverse drug events. In Denmark, which set up a quality indicators project in 1999, assessments initially focused on health care for stroke, hip fracture, schizophrenia, acute surgery, heart insufficiency, and cancer. Doctors and nurses decided which process and outcome indicators to use for each disease and what kinds of data to collect about patients.

In the United States, JCAHO has required accredited health care organizations to compile and submit comparative data on various quality indicators since 1999. These indicators, most of which measure patient clinical outcomes, include data on peri-operative complications and mortality, caesarean section rates, low birth weight rates, and indicators for cardiovascular care, oncology, trauma, medication use, infection control, and ventilated in-patients who develop pneumonia.

Issues Surrounding Comparative Indicator Schemes

Comparative indicator schemes have many advantages over organizational quality assessment schemes, particularly their applicability to outcomes, rather than inputs. Unfortunately, however, as Øvretveit (2001) notes, indicator systems that try to serve too many purposes often end up not serving any one purpose very well. Moreover, the use of indicators such as mortality or survival rates is problematic since they are affected not only by quality, but also by the patient’s age and sex, among other factors, which can lead to changes in reporting methods that render comparability moot. In addition, the definition of indicators depends on who will use them. Governments and politicians, for example, will judge a system on its ability to identify dangerous providers or to encourage quality improvements, while clinicians will judge it on its scientific validity and reliability.

This latter point leads to questions about whose interests are actually represented by the agencies that rely on quality indicators to accredit health care institutions. Accreditation agencies such as the Canadian Council on Health Services Accreditation (CCHSA), the Australian Council on Healthcare Standards, JCAHO in the United States, and Quality Health New Zealand all have governing bodies formed by representatives of professional organizations representing health professionals and health institutions (Bohigas et al. 1996). JCAHO’s governing body also has six seats for public members, to represent the general public’s health interest; the CCHSA has public representatives as well. In Australia and New Zealand, governing boards also have seats for representatives of those countries’ health ministries.
In the United Kingdom, in contrast, the Health Quality Service and the Hospital Accreditation Programme are oriented more toward research and innovation, and the interests of the people and organizations in these fields naturally will be reflected in the accreditation process. In other countries, such as Italy, Spain, Switzerland, and Taiwan, quality assurance in health care is the responsibility of government. In Taiwan, hospitals are denied health insurance payments unless they are accredited, a system similar to that in the United States, where only hospitals that are accredited by JCAHO are eligible for federal funding.

The principal source of revenues for these agencies is fees for their accreditation services. For its part, the CCHSA also charges an annual membership fee, that is offset by low accreditation fees. Annual fees to the CCHSA over a three-year cycle (in 1994 dollars) are $9264 for a 50-bed hospital and $21,618 for a 200-bed hospital. Additional revenues derive from related activities, such as education, publications, and consulting. Concerns exist, however, about accreditation agencies’ offering consulting as a separate service, since it raises the possibility that health care institutions that pay for a consultation will expect to be accredited more easily. In the United States, to allay such concerns, consultations are provided by a company called Quality Healthcare Resources Inc., which is wholly separate from any accreditation agencies (Bohigas et al. 1996).

It is not clear what the specific problems are that play a role in each of these systems. However, it is worth mentioning a couple of points by way of caution, particularly for those systems where no governing agency oversees the accreditation activities. First, it is plausible that hospitals that choose to participate in an accreditation scheme are those that are likely to be accredited or that are planning to improve in any case. Second, since accreditation is costly, only hospitals with sufficient funds will be able to participate, and these could be the ones that have better structures to begin with. Thus, self-selection probably plays a considerable role in the accreditation process. As evidence to support this possibility, it is informative to note that accreditation is uncontroversial for those who choose to participate.

In the United Kingdom, in contrast, hospitals were compared in “league tables” without their consent, which gave rise to a great deal of controversy stemming not so much from the actual rankings as from perceived unfairness inherent in such measures. The next section of the paper discusses that controversy in more detail.
Traditionally, “league tables” are used to compare performances in commerce and sports. Indeed, the frequency with which they appear in the popular press suggests that they are highly valued by consumers, who are able to compare particular performance measures at a glance. Arguments in favour of league tables are based largely on their potential to encourage the emergence of a competitive market structure. When it comes to health care, it is argued that publication of performance measures should stimulate hospitals with low rankings to improve their standing and become competitive with higher-ranked hospitals. League tables should also improve consumer choice by encouraging consumers to seek out the best hospitals, which, in turn, should reinforce competition among providers that wish to attract customers (patients). In addition, league tables ought to change the focus of providers from low cost to high quality.

Whether any of these factors actually come into play, however, is dubious. In 2001, in what Kmietowicz (2001) calls a “courageous attempt” to compare the quality of hospital health care, the National Health Service (NHS) published league tables of mortality rates in hospital trusts throughout the United Kingdom. The NHS adopted league tables for two main reasons (see Adab et al. 2002). First, identifying hospitals whose scores were considerably higher or lower than expected should make it easier to target those with low scores for improvement. Second, illustrating the range in variation among all hospitals could help to explain the reasons for such variation.

Criticisms of League Tables

The publication of hospital league tables has been met with considerable hostility from both the medical community and the public. One of the most problematic issues is the mortality rates used to construct the league tables, which indicated that hospitals with fewer doctors per hospital bed had a higher death rate than hospitals with more favourable ratios of hospital beds to doctors. At first glance, these mortality rates might seem to be a useful comparative indicator, but critics argue, for example, that the association of hospital beds and the supply of doctors with mortality rates fails to take into account such factors as “30-day mortality”, multiple admissions, or hospice transfers, all of which affect mortality rates but do not necessarily reflect the quality of health care available at a particular hospital. Critics also point out that huge differences exist among hospital departments and that, in counting physicians, the NHS omits junior doctors, which likely leads to a gross underestimation of the actual number of physicians available (Bamji 2001).

5 In other words, the mortality rate that is observed within 30 days of the onset of a particular disease or the occurrence of an event such as an operation.
There are other fundamental problems with using mortality as a measure of quality. Rao (2001) notes that mortality depends on the patients who are admitted — people who are sicker upon admission are more likely to die. *A priori,* mortality applies to statistically “average” patients but does not help patients to estimate a successful outcome. Indeed, it might lead to complacency with respect to “good” hospitals and to lawsuits and hysteria with respect to “bad” ones. Castledine (2001) writes that “the emphasis [of hospital league tables] is very much on the patient and outpatient waiting times, financial results and meeting Government agendas, under broad headings such as treatment of staff and hygiene.” Hence, hospital league tables do not convey useful information about the quality of medical services that patients receive at a particular location. Besides, as Shiu (2002) points out, the hospitals with higher mortality rates tend to be larger hospitals that accept more cases, as well as more high-risk cases. The fact is that statistical distortions occur when attempts are made to compare disparate populations. How meaningful are hospital rankings when different-sized hospitals are providing care to patients with different group characteristics?

Other difficulties also exist with respect to the use of league tables. One is that hospitals are in a position to influence the way they report statistics, which is an easier way to gain a higher ranking than actually improving the quality of care they provide. Dranove et al. (2002) suggest that league tables, or report cards, encourage providers to “game” the system by avoiding sick patients, seeking healthy patients, and using their more detailed knowledge of the patients under their care to manipulate statistical outcomes.

Another problem is that, even though they may be armed with considerable knowledge about the quality of particular hospitals, patients have a limited choice of providers they may use. Furthermore, although, as Adab et al. (2002) note, outside the competitive market structure, league tables help system regulators monitor providers and make sure they are accountable for their outcomes, league tables represent only a snapshot in time, which obscures overall quality ratings. For example, over a long period of time, two hospitals may have the same mortality rate for a particular procedure, but their rates might differ substantially in any given year. A snapshot taken in that year might then present a distorted comparison of the two institutions for reasons not necessarily related to the relative quality of care they provide.

One remedy for the possibly distorted information that league tables provide is, then, to take account of the fact that hospitals may show substantial variations in outcomes over the short term and to compare providers over time. As Parry et al. (1998) relate, however, even comparisons over time may be misleading. In their study of mortality rates in neonatal intensive care units over a six-year period, they find that “[i]t is impossible to assess whether a single significant result reflects a difference in the quality of care or chance” (p. 1933). Moreover, they argue, variations in rankings from year to year can be explained by random events, rather than by systematic differences among hospitals.
The difficulty of using indicators to compare health care providers in a meaningful way has led, of late, to an increased focus on using information about “adverse events” — that is, events that should not happen in a hospital — as a measure of quality.

Unfortunately for patients, adverse events are all too common, and data on them are readily available. The US Food and Drug Administration (FDA) cites adverse drug reactions (ADRs) as one of the leading causes of morbidity and mortality in health care. The Institutes of Medicine reported in January 2000 that, of the 44,000 to 98,000 deaths annually in the United States from medical errors, about 7000 are attributable to ADRs.\textsuperscript{6} Other US studies estimate that as many as 6.7 percent of hospitalized patients experience serious ADRs and claim an annual fatality rate of about 0.32 percent, which would translate into more than 106,000 deaths annually from ADRs.

According to the FDA, the reason so many ADRs occur is that more drugs are being prescribed than ever before: 2.8 billion prescriptions were filled in the United States in 2000 (equivalent to ten prescriptions per person). Indeed, 64 percent of all patient visits to a physician result in the writing of a prescription. Moreover, ADRs increase exponentially when the patient takes four or more medications at the same time. Such multiple drug ADRs represent between 3 and 5 percent of all in-hospital medication errors, and they are an important cause of patient visits to emergency departments.

The problem of adverse medical events is a serious one in Canada as well. According to Forster et al. (2004), almost one-quarter of all patients discharged from Canadian teaching hospitals experience an adverse event, half of which are preventable or improvable. Baker et al. (2004) offer a more conservative estimate of the incidence of adverse events: 7.5 percent in their sample of Canadian hospitals; given 2.5 million annual hospital admissions, that translates into 185,000 adverse events, 70,000 of which are judged to be preventable.

**Six Sigma**

Despite the relatively high incidence of preventable adverse events and drug reactions, however, most accreditation models do not gather and report on their frequency. Only one formal approach to improving quality focuses on reducing the number of errors: an industrial approach called “Six Sigma”

\textsuperscript{6} See the FDA’s web site: <http://www.fda.gov/cder/drug/drugReactions/default.htm#Sample%20Case>.
(a Motorola trademark). The system was invented to reduce the number of manufacturing defects, and involves a six-level scale whereby the lower the sigma level the higher the number of defects per million opportunities. Thus, for example, a sigma of 0.5 represents 841,345 defects per million opportunities, and a sigma of 6 represents 3.4 defects per million opportunities. Adherents to Six Sigma therefore strive to achieve the highest sigma level possible.

Six Sigma can also be used to assess the quality of care in the health sector, where defects represent medical errors. Hospitals can receive a high sigma score by reducing the number of adverse events that occur in their facility, pharmacies can receive a high sigma score by reducing medication errors, and laboratories can receive a high sigma score by delivering accurate results more frequently. Thus, for the health care provider, the quest for a high sigma score reduces costs, sometimes significantly, while for the patient it reduces the likelihood that an adverse event will occur.

It is not just the attainment of the highest possible sigma value that is important, however, but also the development of processes to detect and evaluate error rates and to make systematic changes to improve dependability. Johnstone et al. (2003) suggest a four-phase approach to implementing a Six Sigma strategy in a health care setting. Phase 1 makes sure that everyone in the organization supports the drive for high-quality care. Phase 2 implements “second checks” on the processes. Phase 3 mandates automation wherever possible to reduce human errors. Phase 4 compares outcomes to existing benchmarks to assess relative quality. Johnstone et al. note, however, that the application and implementation of Six Sigma is appropriate only in certain settings — such as procedures that have a large number of iterations, that involve double checking by humans, and that have potential for computer oversight.

Another issue with this approach, as this author’s own Six Sigma calculations reveal, is that the result is very sensitive to events that can happen by chance — an error might occur in a process where it usually does not — if there are not enough observations, or repetitions, of a particular process, but the result is not so sensitive when there are many repetitions (see Johnstone et al. 2003).

Six Sigma has been implemented successfully by a number of health care providers in the United States, as reported in several studies, and is becoming increasingly popular there. In Canada, however, hospitals undertake little in the way of continuous data monitoring, which is a requirement for Six Sigma. Moreover, because hospital comparisons do not exist in Canada, high error rates are tolerated and there is little accountability.

---

7 See, for example, Bahensky, Roe, and Bolton 2005; Sunyog 2004; and Loree, Maihack, and Powell 2003.
Consumers need to be able to compare the quality of different products and services — whether it is the gas mileage of different models of cars or the quality of health care provided by different hospitals — if they are to make an informed choice. However, when the data that are used to compile relative rankings are poor, when outcomes are measured improperly, when the indicator is basically irrelevant to those who rely on the rankings, and when results are more likely to cause a gaming of numbers than a serious improvement in quality, then performance comparisons can actually do more harm than good.

The one element that the various models of comparative hospital performance seem to lack, or to convey clearly, is relevance to patients. Instead, these measures are oriented to providers. ISO standardization, for example, while thorough, is business oriented and does not focus on outcomes. Hospital league tables focus on certain outcomes, but the lack of standardization makes such rankings difficult to interpret. The main concern of patients, in contrast, is how likely they are to get better, and whether this is related to the care they receive at a particular hospital.

Much more exploration remains to be undertaken to find appropriate indicators, to ensure proper data collection, and to make hospital report cards or league tables more relevant and easier to understand for those who use them. Reaching these goals, however, will help to build a solid foundation for establishing quality assurance and comparability among health care institutions, to the benefit of all stakeholders.


A Finger on the Pulse


Selected Publications from the AIMS Library

Other AIMS work on health care:

Definitely NOT the Romanow Report, by Brian Lee Crowley, Brian Ferguson, David Zitner, and Brett J. Skinner

Payment Is Powerful: Overcoming Canada's Shortage of GPs by Increasing Family Practice Compensation, by Ida Rayson

Operating in the Dark: The Gathering Crisis in Canada's Public Health Care System, by Brian Lee Crowley, Dr. David Zitner, and Nancy Faraday-Smith

Expenditure on Medical Care in Canada, by Brian Ferguson

The Non-Sustainability of Health Care Financing under the Medicare Model, by Brett J. Skinner

Medicare and User Fees: Unsafe at Any Price? by Carl Irvine and Dr. David Gratzer

Alice in Borderland: Why Canadians Cannot Afford to Be Complacent about US Drug Re-importation, by Brian Ferguson

This Won't Hurt a Bit: Why the Vaccine Crisis Shows We Shouldn't Believe What Health 'Planners' Tell Us about How to Reform Drug Policy, by Brian Ferguson and Julia Witt


Issues in the Demand for Medical Care: Can Consumers and Doctors Be Trusted to Make the Right Choices? by Brian Ferguson

How Should We Decide What to Cover under Medicare? by Julia Witt

Profits and the Hospital Sector: What Does the Literature Really Say? by Brian Ferguson

Improving Canadian Health Care: Better Ways to Finance Medicare, by Brett J. Skinner

Better Medicine, by David Gratzer

Public Health, State Secret, by Dr. David Zitner and Brian Lee Crowley

Road to Growth: How Lagging Economies Become Prosperous, by Fred McMahon

Looking the Gift Horse in the Mouth: The Impact of Federal Transfers on Atlantic Canada, by Fred McMahon (photocopies only)

Commentary Series

Locking Up the Pork Barrel: Reasoned Economic Development Takes a Back Seat to Politics at ACOA, by Brian Lee Crowley and Bruce Winchester

Following the Money Trail: Figuring Out Just How Large Subsidies to Business Are in Atlantic Canada, by David Murrell

First, Do No Harm: What Role for ACOA in Atlantic Canada? by Brian Lee Crowley

Research Reports

Debtor's Prison II: Shortening the Sentence, by Roland T. Martin

Debtor's Prison: Public Sector Debt and What It Means for You, by Tom Riley

The Atlantica Power Market: A Plan for Joint Action, by Gordon L. Weil

New Brunswick's Plastics Industry: Rhetoric versus Reality, by Isaiah A. Litvak

Fencing the Last Frontier: The Case for Property Rights in Canadian Aquaculture, by Robin Neill

Grading Our Future: Atlantic Canada's High Schools' Accountability and Performance in Context, by Rick Audas and Charles Cirtwill

Rags to Riches: How “The Regions” Can and Should Be Leading Canada’s Productivity Push, by Brian Lee Crowley

Taxing Incentives: How Equalization Distorts Tax Policy in Recipient Provinces, by Kenneth J. Boessenkool

Fiscal Equalization Revisited, by Professor James M. Buchanan, Nobel Laureate

Testing & Accountability: The Keys to Educational Excellence in Atlantic Canada, by Charles Cirtwill, Rod Clifton, and John D'Orsay

Books

Retreat from Growth: Atlantic Canada and the Negative-Sum Economy, by Fred McMahon

These publications are available at AIMS, 2000 Barrington St., Suite 1006, Halifax NS B3J 3K1
Telephone: (902) 429-1143 Facsimile: (902) 425-1393 E-mail: aims@aims.ca
They can also be found on our web site: www.aims.ca