

## APPROPRIATENESS IN HEALTH CARE DELIVERY: DEFINITIONS, MEASUREMENT AND POLICY IMPLICATIONS

John N. Lavis, MD, MSc; Geoffrey M. Anderson, MD, PhD

### Abstract • Résumé

A major focus of the current health care debate is the notion that a substantial proportion of the health care delivered in Canada is inappropriate. There are two types of appropriateness: appropriateness of a service and appropriateness of the setting in which care is provided (i.e., inpatient v. outpatient or home care). Measuring both types objectively requires the comparison of observed patterns of care with explicit criteria for appropriate care. The few studies of appropriateness conducted in Canada have shown that inappropriate services are provided and inappropriate settings are used. Reducing inappropriate health care delivery could involve active strategies for the implementation of guidelines and better cooperation and coordination within the health care system. However, lower rates of health care delivery or even of inappropriate health care delivery will not necessarily translate into higher quality care or lower costs overall.

Un des grands sujets du débat en cours sur les soins de santé est le concept selon lequel une partie importante des soins de santé fournis au Canada ne seraient pas pertinents. Il y a deux types de pertinence. La pertinence d'un service et celle du contexte où les soins sont fournis (c.-à-d. service interne c. service externe ou soins à domicile). Pour mesurer objectivement les deux types de pertinence, il faut comparer les tendances observées des soins à des critères explicites sur les soins pertinents. Les rares études sur la pertinence effectuées au Canada ont démontré qu'on fournit des services non pertinents et qu'on fournit des services dans des contextes non pertinents. La réduction des soins de santé non pertinents pourrait passer par des stratégies actives de mise en oeuvre de lignes directrices et par une amélioration de la collaboration et de la coordination à l'intérieur du système de soins de santé. Une baisse des taux de prestation de soins de santé ou même de prestation de soins de santé non pertinents ne se traduira toutefois pas nécessairement par une hausse globale de la qualité ou par une baisse globale des coûts.

The quality and cost of health care delivery have come under increasing scrutiny in recent years. A major focus of the current health care debate is the notion that a substantial proportion of the health care delivered in Canada is inappropriate.<sup>1,2</sup> As conventional wisdom has it, reducing inappropriate care would not only improve the quality of care but also save money. If true, reducing inappropriate care is an appealing approach to addressing the fiscal realities faced by many provincial governments while preserving the health care system that has served this country so well.

The appropriateness of the health care delivered in Canada will clearly continue to interest health care providers, policymakers and the public. To help inform future discussions, this article addresses four basic questions: What is meant by "appropriate" and "inappropri-

ate" care? What general approaches and specific methods have been used to measure the levels of inappropriate care? What do we know about the levels of inappropriate health care delivery in Canada? And, what lessons can be drawn for future efforts to measure and reduce inappropriate care?

### THE MEANING OF APPROPRIATENESS AND INAPPROPRIATENESS

Appropriateness is often treated as a single concept. However, there are two distinct types of appropriateness: appropriateness of a service and appropriateness of the setting in which care is provided. The differences between the two parallel the differences between two other concepts in health care: effectiveness and cost-

*Dr. Lavis is from the Division of Health Policy Research and Education, Harvard University, Boston, Mass., and Dr. Anderson is from the Institute for Clinical Evaluative Sciences in Ontario, North York, Ont., and the Department of Health Administration, University of Toronto, Toronto, Ont.*

*Reprint requests to: Dr. Geoffrey M. Anderson, Institute for Clinical Evaluative Sciences in Ontario, G-106, 2075 Bayview Ave., North York ON M2N 3M3*

effectiveness. Clinical research determines whether the health of a patient is expected to improve after a service or procedure is provided (effectiveness). Economic evaluation, building on the results of clinical research, addresses cost-effectiveness: identification of the least expensive mix of services required to improve the patient's health.

The appropriateness of a service is the effectiveness of that service for a particular type of patient. Appropriateness is determined by whether symptoms, physical findings and results of diagnostic tests indicate that the particular patient is expected to benefit from the service. This type of appropriateness involves the choice of service that should benefit the patient. When appropriateness of a service is discussed, it is assumed that the service is provided in a high-quality way. Typically, costs are disregarded, and no attempt is made to determine whether the benefits are worth the costs, given the alternative uses (within or outside the health care system) for the money spent on the service.

We use the following definitions. An appropriate service is one that is expected to do more good than harm for a patient with a given indication or set of indications. An inappropriate service is one that is not expected to benefit the patient or, in the more extreme case, may harm the patient. An equivocal service is neither clearly appropriate nor clearly inappropriate.

The appropriateness of the setting in which care is provided is related to cost-effectiveness. This type of appropriateness is determined by whether the patient's clinical characteristics, and the services required for his or her care, match the setting in which the care is provided. Setting is a proxy measure of the resources used to provide care. Just as effective care can be provided in a way that is not cost-effective, appropriate services can be provided in inappropriate settings. When appropriateness of setting is considered, it is assumed that the services are appropriate and are provided in a technically correct way.

Coronary artery bypass grafting (CABG) can serve as an example of how the effectiveness and appropriateness of a service are related. Clinical trials have shown that CABG is effective in improving survival rates among patients with certain indications.<sup>3</sup> For example, the results of these trials show that CABG is effective (and thus appropriate) if it is performed on a patient with severe angina, positive results of noninvasive tests and 90% narrowing of the left main coronary artery. However, the same trials may show that CABG is ineffective (and thus inappropriate) if it is performed on a patient with no symptoms, negative results of noninvasive tests and only 30% narrowing of the circumflex coronary artery.

To illustrate the relation between the cost-effectiveness of care and the appropriateness of the setting, we will use

the example of endoscopic examination. Providing an endoscopic examination in an outpatient setting is less costly for the health care system (although it may be more costly for the patient) than admitting the patient to hospital for this purpose. However, a patient undergoing a gastric endoscopic examination may require inpatient care for other clinical reasons such as bleeding or a serious co-existing illness. Such a patient is too sick to be examined as an outpatient. In contrast, if the patient has symptoms of a peptic ulcer but is in no distress and has no signs of bleeding, hospital admission for an endoscopic examination may be an inappropriate use of resources. This patient could safely undergo the procedure on an outpatient basis. Hence, an endoscopic examination may be the appropriate service for both types of patient, but hospital inpatient care may be an inappropriate setting in which to provide the service to the second type of patient.

Different conclusions can be drawn from the two types of inappropriate health care delivery. A service that is inappropriate for a specific type of patient should not be provided in any setting. The service is not expected to benefit the patient and, therefore, is not needed. An inappropriate setting means that care could be provided in an alternative setting (usually on an outpatient basis or at home) at a lower cost. However, "could" is the operative word: the services the patient needs may not be available in an alternative setting. For example, a patient may not be sick enough to warrant hospital admission but may be too sick to be sent home without adequate home care. If home care services are unavailable, then hospital admission may theoretically be inappropriate, but it is the better available option. Hence, identifying care settings as inappropriate depends on the availability of alternative, less expensive settings.

## APPROACHES TO MEASUREMENT

Measuring appropriateness objectively hinges on the comparison of observed patterns of care with criteria for appropriate care. To assess the appropriateness of a service, detailed clinical information in the medical record is reviewed to determine the indications for the service and any relevant risk factors or coexisting illnesses. To assess the appropriateness of setting, the medical record is reviewed to determine the severity of illness and the intensity of services needed to care for the patient during the period studied. For both types of appropriateness the results of the medical-record review are compared with a list of criteria to determine whether the care was appropriate, equivocal or inappropriate.

Assessments of the appropriateness of a service are based on criteria specific to the diagnosis or procedure. These criteria are typically arrived at through a critical

appraisal of the research literature, followed by a process to achieve consensus among a group of experts. The consensus process is needed to interpret the research literature and establish the appropriateness of a service for each possible or common combination of indication, risk factor and coexisting illness.

For assessments of appropriateness of setting, the criteria are typically independent of diagnosis; they are applicable to most categories of patients. These criteria are developed through an expert-consensus process. There is little research literature on the most appropriate setting for care of patients with a wide range of clinical presentations; the existing literature is usually specific to a single clinical condition.

#### SPECIFIC METHODS TO MEASURE APPROPRIATENESS

Although these general approaches are straightforward and broadly applicable, only a small number of well-documented methods are commonly used to measure appropriateness. The best known of these was developed by researchers at RAND, an independent US research organization.<sup>4</sup> Criteria for the appropriateness of seven services have been applied in published studies; the services are coronary angiography, CABG, percutaneous transluminal coronary angioplasty, carotid endarterectomy, hysterectomy, endoscopy and, most recently, placement of tympanostomy tubes.<sup>5,6</sup> To develop criteria, each member of a panel of experts ranks all possible indications for a given service from 1 (inappropriate) to 9 (appropriate) on the basis of a critical review of the research literature. Experts then meet to discuss the results of this ranking exercise and the research literature, then rank the indications a second time. Through this process, a service is classified as appropriate, equivocal or inappropriate on the basis of the indications, risk factors and coexisting illnesses documented in a patient's medical record.<sup>4</sup>

The validity of this method is, and will remain, uncertain, since there is no ideal standard against which to validate such methods. The definition of inappropriate service provision varies depending on the nationality of the experts and the decision rule used to define agreement.<sup>2</sup> In the absence of an agreed standard for measuring appropriateness, the sensitivity and specificity of measurement tools cannot be determined. However, we do know that no measure of appropriateness is perfect. The use of these measures will sometimes lead to classifying services as appropriate when they are inappropriate and vice versa. The extent of bias is uncertain.<sup>7</sup>

The most commonly used tools for assessing the appropriateness of setting are the Appropriateness Evaluation Protocol (AEP)<sup>8,9</sup> and the Intensity-Severity-Discharge-Appropriateness (ISD-A) review system (In-

terQual Inc., Westborough, Mass., 1978). Both include criteria for assessing the appropriateness of a hospital admission or a day of care. These criteria were originally developed by consensus process, are easily updated and can be modified to suit the local environment. A hospital admission or day of care is classified as appropriate or inappropriate on the basis of these criteria and of an assessment, from the patient's medical record, of the severity of illness and the intensity of service needed for the care of that patient. (It is more accurate to classify hospital admissions and days of care as requiring or not requiring acute care, since the determination that the setting is "inappropriate" depends on the availability of alternative, less expensive care settings.) Used prospectively, some of these tools can also attribute hospital admissions and days of care when acute care is not required to the patient, physician, hospital or environment. The validity of these two measurement tools has been established through comparisons with assessments by panels of physicians in which implicit techniques were used.<sup>10,11</sup>

#### LIMITATIONS OF EXISTING METHODS

There are obvious limitations to these two types of assessment of appropriateness. The comprehensiveness of the RAND method, which provides an appropriateness rating for every conceivable combination of indication, risk factor and coexisting illness, has meant that criteria have been developed for only seven services (all of which are hospital-based procedures provided by specialists). These criteria have been used in Canada only for research purposes.

The AEP and ISD-A criteria for appropriateness of setting are applicable only to acute care provided in hospitals. There is no systematic approach to identifying, for example, patients in long-term care facilities who could receive home care (or vice versa) or patients in home care programs who could use self-care (or vice versa). As well, these measurement tools do not assess the appropriateness of resource use in a setting. The level and mix of health care providers is assumed to be fixed. For example, the possibility of substituting a licensed practical nurse for a registered nurse to care for certain patients is not considered.

Limitations to existing methods can be attributed to the environments in which they were developed and in which they evolved. Both of these methods are often used prospectively in the United States to determine whether a physician or hospital is paid for providing a service or inpatient care. Because health care insurers and other third-party payers in the United States are mainly concerned about "big-ticket" items, such as specialist-delivered procedures and hospital-based care, appropriateness criteria have been developed in these ar-

cas. However, there is no reason why appropriateness criteria that meet the specific needs of the Canadian health care system could not be developed for any service or type of health care setting. In the Canadian context, assessing the appropriateness of high-volume, "small-ticket" items may have much more impact than assessment of big-ticket items on the overall quality (and possibly even cost) of care.

Existing methods rely exclusively on the review of medical records of patients who have received a service or who have been admitted to acute care institutions. As typically applied, these methods define only appropriate or inappropriate use of services or hospitals. However, measurement tools such as the RAND criteria or ISD-A review system could theoretically identify patients who need a particular service or access to a particular health care setting but do not receive it. They have never been used in this way, although measuring underuse of appropriate services or settings is as important as measuring inappropriate overuse.

## LEVELS OF INAPPROPRIATE HEALTH CARE DELIVERY IN CANADA

To illustrate the application of existing methods and their usefulness in answering relevant policy questions, we reviewed and critically appraised studies of appropriateness conducted in Canada and relevant studies conducted in other jurisdictions.<sup>2</sup>

### INAPPROPRIATE SERVICE PROVISION

Three Canadian studies have measured the rate of inappropriate service provision for selected procedures. The results of these studies are presented in Table 1. The first study determined the rate of inappropriate hysterectomies performed in selected hospitals in Saskatchewan

in the early 1970s.<sup>12</sup> The second study, in which the RAND method was used with criteria developed by US physicians, compared rates of inappropriate coronary angiography and CABG in a Winnipeg hospital with those in three US hospitals in the early 1980s.<sup>13</sup> It showed that the proportion of inappropriate cases was lower in the Winnipeg hospital than in the US hospitals. The most recent study performed in Canada also used the RAND method but with criteria developed by both Canadian and US physicians.<sup>14</sup> The study examined a random selection of all cases of coronary angiography and CABG performed in New York, Ontario and British Columbia and showed similar rates of inappropriate procedures in the two countries.

The rate of inappropriate service provision in the United States during the past decade has been measured in many studies. Rates for procedures that have been studied three or more times are summarized in Table 2. Rates of inappropriate coronary angiography<sup>15-18</sup> and CABG<sup>17,19,20</sup> shown in New York in 1990 were lower than those shown among other states from 1979 to 1982.

The proportion of cases defined as inappropriate was substantially higher when the experts involved in the ranking process were from the United Kingdom.<sup>17</sup> This suggests that, even when presented with the same research evidence and asked to ignore cost considerations in the ranking process, experts in different countries have different definitions of inappropriate service provision.

These results suggest that rates of inappropriate service provision vary according to the year of the study, the location of the study and the nationality of the experts who develop the criteria. Rates of inappropriate provision have also been shown to vary according to the decision rule used to define agreement.<sup>2</sup> Further research is needed to determine whether rates vary depending on the characteristics, such as age and social class, of the population being studied.

Table 1: Proportion of services deemed inappropriate for selected procedures in Canada and the United States

Procedure	Study	Study year	Canadian study location	% of services deemed inappropriate			
				In Canada		In the United States	
				According to Canadian criteria	According to US criteria	According to Canadian criteria	According to US criteria
Hysterectomy	Dyck et al <sup>12</sup>	1971-73	Saskatchewan	24	NA*	NA	NA
Coronary angiography	Roos et al <sup>13</sup>	1981	Winnipeg	NA	6-9	NA	15-18
	McGlynn et al <sup>14</sup>	1989-90	Ontario and British Columbia	9	5	10	4
Coronary artery bypass grafting	Roos et al	1981	Winnipeg	NA	1-3	NA	6-23
	McGlynn et al	1989-90	Ontario and British Columbia	4	3	6	2

\*NA = not applicable.

Although absolute rates of inappropriate service provision in one time and place may not be generalizable, the approach can be used to answer relevant policy questions. For example, the RAND method has been used to explore the relation between small-area variations in the rates of specific services and rates of inappropriate provision of these services. Although the caveat concerning the possibility of biased estimates should be kept in mind, high rates of services in some jurisdictions have not been shown to be correlated with high rates of inappropriate service provision. This result helps to answer the policy question concerning whether high rates of services can be used as a marker of inappropriateness. According to evidence from the United States, the answer is probably No,<sup>15,27</sup> however, this is still subject to debate.<sup>28-30</sup>

## INAPPROPRIATE SETTING

Two studies have systematically measured the rate of inappropriate acute care hospital use (i.e., use of acute care hospitals when that setting for care was not justified) by adults in Canada. Results are summarized in Table 3. The first examined admissions to selected hospitals in British Columbia.<sup>31</sup> The second, in Saskatchewan, was Canada's first province-wide study of inappropriate use of acute care.<sup>32</sup> In both studies, researchers used the ISD-A measurement tool. Studies conducted in the United States have shown rates of hospital admissions when acute care is not required ranging from 7% to 43%, and rates of days of care when acute care is not required ranging from 20% to 48%.<sup>2</sup>

Other studies have measured the rate of hospital use

**Table 2: Proportion of services deemed inappropriate for selected procedures in the United States**

Procedure	Study	Study year	% of service deemed inappropriate	
			According to US criteria	According to UK criteria
Coronary angiography	Chassin et al and Brook et al <sup>15-17</sup>	1979-81	17-27	42-60
	Bernstein et al <sup>18</sup>	1990	4	NA
Coronary artery bypass grafting	Brook et al and Winslow et al <sup>17,19</sup>	1979-82	13-14	35
	Leape et al <sup>20</sup>	1990	2	NA
Carotid endarterectomy	Chassin et al, Merrick et al, Winslow et al and Leape et al <sup>16,21-23</sup>	1981	13-32	NA
Upper gastrointestinal endoscopy	Chassin et al and Kahn et al <sup>16,24</sup>	1981	11-19	NA
	Kahn et al <sup>25,26</sup>	1982-93	24	NA

**Table 3: Rates of hospital use when acute care was not required (inappropriate setting) among adults and children in Canada**

Study population	Study	Study year	Study location	Rate of inappropriate acute care hospital admissions or inappropriate days of care in hospital, %	
				Hospital admissions	Days of care
Adults	Anderson et al <sup>31</sup>	1987	Victoria	24	13-33*
	Health Services Utilization and Research Commission (HSURC) <sup>32</sup>	1991-92	Saskatchewan	38-48†	48-64†
Children	Kasian et al <sup>33</sup>	1988-89	Saskatoon	NA	16
	Gloor et al <sup>34</sup>	1988	London, Ont.	NA	24
	Smith et al <sup>35</sup>	1990	Vancouver	29	22
	HSURC	1992	Saskatchewan	44-56	27-48
	Davis‡	1992	Toronto	NA	13

\*Range based on different definitions of inappropriate days of care.

†Range among three different categories of hospitals; lower limit is result from regional hospitals and upper limit is result from small community hospitals.

‡W.M. Davis, Hospital for Sick Children: personal communication, 1994.

among children in Canada when acute care is not required; all except a study in Saskatchewan were conducted with the use of the pediatric AEP measurement tool. Two of these studies examined admissions,<sup>32,35</sup> whereas five examined days of care<sup>32-35</sup> (W.M. Davis, Hospital for Sick Children: personal communication, 1994). In the United States, a study showed that in 11% of admissions among children acute care was not required<sup>9</sup> and that in 13% to 21% of days of care for children acute care was not required.<sup>9,36</sup>

Rates of hospital use when acute care was not required also vary depending on the year of the study and the location of the study population. Hence, absolute rates of inappropriate use based on one time and place may not be generalizable. To establish the local rate of inappropriate use of acute care, existing methods need to be applied locally.

As in the case of inappropriate service provision, the approach can also be used to answer policy questions concerning care settings. For example, the AEP measurement tool has been used to explore the effects of user charges and of care provided through prepaid group practices (such as health maintenance organizations in the United States) on inappropriate hospital care. User charges have been shown to reduce the use of hospital care but not the inappropriate proportion of that use.<sup>37</sup> Physicians working in prepaid group practices, in contrast to those working under fee-for-service arrangements, face a financial disincentive to provide "excess" hospital-based care. One study showed that physicians working in prepaid group practices had lower rates of hospital admissions but not lower proportions of hospital admissions when acute care was not required.<sup>38</sup> These results suggest that neither user charges nor care provided through prepaid group practices can be relied upon to reduce hospital use when acute care is not required.

## LESSONS FOR THE FUTURE

Appropriateness in health care delivery is clearly an attractive concept. It allows us, theoretically, to progress from describing how much care is being provided to analysing how much of that care is expected to benefit patients or whether that care could be provided less expensively. These types of analysis are important steps in our efforts to improve the quality of care while controlling costs. Yet the steps taken to date in Canada have been very tentative.

### MEASURING INAPPROPRIATE HEALTH CARE DELIVERY

In only one study conducted in Canada during the last two decades have researchers used clearly defined,

explicit criteria, developed by Canadian physicians, to examine the appropriateness of a service in a comprehensive sample of hospitals.<sup>14</sup> Furthermore, after more than a decade of efforts involving the RAND method, criteria have been developed to analyse the appropriateness of only seven procedures. Attempts to measure systematically the appropriateness of services provided by Canadian physicians will require a major effort to develop evidence-based criteria. Such criteria should be arrived at through an explicit process combining critical appraisal of the literature with expert consensus. This process should be coupled with the development of systems for the routine collection of data to enable researchers to assess the appropriateness of services in representative samples of cases. These two steps would allow the more widespread use of this approach.

Existing methods of assessing the appropriateness of setting apply only to acute care hospitals. Similar methods could be developed to assess the appropriateness of other health care settings such as long-term care facilities or home care programs. This would permit the application of this general approach to all settings in the health care system and would promote accountability throughout the system. Recent research suggests that a substantial proportion of the care provided in acute care hospitals in Canada could be provided, perhaps less expensively, in other settings. This is only one step on the road to establishing the most appropriate setting for each patient. More needs to be learned about the availability and appropriateness of settings other than the hospital. All of the studies conducted in Canada to date have been able to identify only inappropriate overuse. However, in a system designed to provide access to care for all who need it, it seems prudent to begin to look at the potential underuse of appropriate health care as well.

### REDUCING INAPPROPRIATE HEALTH CARE DELIVERY

Actions to address inappropriate health care delivery will probably not wait for these refinements in measurement. Inappropriate care is being provided in Canada, and ways to reduce such care will be implemented. However, efforts to reduce inappropriate health care delivery should take into account the unique organization of the Canadian health care system.

Efforts to improve quality of care and contain costs in the United States have been compared to the use of reins, whereas efforts in Canada have been compared to the use of fences.<sup>39</sup> Third-party payers in the United States attempt to exert influence at each turn in medical decision making, much like a rider would control a horse. In contrast, provincial health care systems in Canada have instead adopted approaches, such as global budgets for hospitals and utilization caps for physician



services, that act as fences within which health care providers must operate.

This analogy to reins and fences can also be applied to efforts to reduce inappropriate health care delivery. For example, efforts to reduce inappropriate delivery in the United States have included measuring appropriateness prospectively to determine the eligibility of a service or admission for reimbursement on a patient-by-patient basis. Care that does not meet the criteria is not paid for. If such efforts were undertaken in Canada, it would mean a fundamental transformation of the relationship between providers and governments.

In contrast, future efforts to reduce inappropriate delivery in Canada could involve a traditional "fences" approach. Rather than making case-by-case judgements concerning whether care will be reimbursed, providers and governments could make a commitment to develop clearer lines of responsibility for quality assurance and utilization management. Clinical practice guidelines<sup>40</sup> could provide the basis for defining appropriate services. However, guidelines are a necessary but not sufficient condition for improving the quality of care; therefore, active strategies to implement these guidelines will also be needed.<sup>41</sup> Guidelines can be implemented successfully through an understanding of the local barriers to change and through comprehensive strategies to overcome these barriers.<sup>42</sup> We need a system that will support such interventions.

Reducing care provided in an inappropriate setting must begin with the realization that such change requires cooperation and coordination among institutions, organizations and providers. Studies of inappropriate care settings in Canada have measured the levels of inappropriate hospital use, not the reasons for inappropriate use. Without an understanding of the factors that lead to inappropriate use, it will be impossible to address this issue.

#### A CAVEAT: LESS IS NOT NECESSARILY BETTER

Reductions in health care spending have led to, or may necessitate, many initiatives to reduce health care utilization. If such efforts are not targeted specifically at reducing inappropriate care, appropriate care could be reduced at the same time as inappropriate care is eliminated, resulting in lower overall quality of care. Lower rates of use do not necessarily mean lower rates of inappropriate care. Quality concerns do not end with ensuring that delivered services are appropriate for the patient and are provided in the appropriate setting. In a system designed to provide universal access to necessary services, the analysis of appropriateness of care should extend to an examination of those who did *not* receive an appropriate service or did *not* receive care in the required setting. There is evidence of overuse of services in Canada, but we know very little

about underuse and nothing about underuse precipitated by untargeted reductions in health care utilization.

The cost implications of initiatives to reduce inappropriate care are also complex. If such initiatives are based on a broad view of quality of care, which incorporates the notions of overuse and underuse, they may not save money. For example, careful examination of care provided outside of acute care settings may result in a net flow of patients from other settings to acute care settings. Moreover, the type of targeted efforts needed to reduce inappropriate care will require new administrative structures and substantial financial investment.

## CONCLUSION

Many in the Canadian health care system are facing demands for increased accountability. A central feature of accountability is ensuring that appropriate services are provided in appropriate settings. To reach this goal we must first establish criteria for appropriate care through a process that is open and explicit, combining critical appraisal of the literature with expert consensus. We must then assess performance systematically against these criteria. This assessment should be methodologically sound, but it must also involve local health care providers. Change is easier when the need for it is clear and acceptable to all. The commitment and cooperation of everyone in the system are needed for change to take place. To meet our shared goal of accountability for the quality and costs of health care delivery, we must also recognize the unique Canadian relationship among health care providers, policymakers and the public.

---

This article is based on a report commissioned by the Queen's-University of Ottawa Economic Projects as part of the Cost-Effectiveness of the Canadian Health Care System Project. We thank Jonathan Lomas for helpful comments on the report and on an earlier version of this article.

Dr. Anderson was supported by the National Health Research and Development Program (grant no. 6606-5073-48), the Institute for Clinical Evaluative Sciences in Ontario (ICES) and the University of Toronto. Dr. Lavis received support for this project from ICES.

## References

1. Working Group on Utilization Management: *When Less Is Better: Using Canada's Hospitals Efficiently*, Conference of Federal/Provincial/Territorial Deputy Ministers of Health, Ottawa, 1994
2. Lavis JN, Anderson GM: *Inappropriate Hospital Use in Canada: Definition, Measurement, Determinants and Policy Implications* [University of Ottawa Graduate School working paper 93-08], Cost-Effectiveness of the Canadian Health Care System, Queen's-University of Ottawa Economic Projects, Ottawa, 1993
3. Yusuf S, Zucker D, Perduzzi P et al: Effect of coronary artery bypass graft surgery on survival. *Lancet* 1994; 344:

4. Brook RH, Chassin MR, Fink A et al: A method for the detailed assessment of the appropriateness of medical technologies. *Int J Technol Assess Health Care* 1986; 2: 53-63
5. Park RE, Fink A, Brook RH et al: Physician ratings of appropriate indications for six medical and surgical procedures. *Am J Public Health* 1986; 76: 766-772
6. Kleinman LC, Kosecoff J, Dubois RW et al: The medical appropriateness of tympanostomy tubes proposed for children younger than 16 years in the United States. *JAMA* 1994; 271: 1250-1255
7. Phelps CE: The methodologic foundations of studies of the appropriateness of medical care. *N Engl J Med* 1993; 329: 1241-1245
8. Gertman PM, Restuccia JD: The Appropriateness Evaluation Protocol: a technique for assessing unnecessary days of hospital care. *Med Care* 1981; 19: 855-871
9. Kreger BE, Restuccia JD: Assessing the need to hospitalize children: Paediatric Appropriateness Evaluation Protocol. *Pediatrics* 1989; 84: 242-247
10. Strumwasser I, Paranjpe NV, Ronis DL et al: Reliability and validity of utilization review criteria: Appropriateness Evaluation Protocol, Standardized Medreview Instrument and Intensity-Severity-Discharge Criteria. *Med Care* 1990; 28: 95-111
11. Kemper KJ, Fink HD, McCarthy PL: The reliability and validity of the Paediatric Appropriateness Evaluation Protocol. *Qual Rev Bull* 1989; 15: 77-80
12. Dyck FJ, Murphy FA, Murphy JK et al: Effect of surveillance on the number of hysterectomies in the province of Saskatchewan. *N Engl J Med* 1977; 296: 1326-1328
13. Roos LL, Bond R, Naylor CD et al: Coronary angiography and bypass surgery in Manitoba and the United States: a first comparison. *Can J Cardiol* 1994; 10: 49-56
14. McGlynn EA, Naylor CD, Anderson GM et al: Comparison of the appropriateness of coronary angiography and coronary artery bypass graft surgery between Canada and New York State. *JAMA* 1994; 272: 934-940
15. Chassin MR, Kosecoff J, Park RE et al: Does inappropriate use explain geographic variations in the use of health care services? A study of three procedures. *JAMA* 1987; 258: 2533-2537
16. Chassin MR, Kosecoff J, Solomon DH et al: How coronary angiography is used: clinical determinants of appropriateness. *JAMA* 1987; 258: 2543-2547
17. Brook RH, Kosecoff JB, Park RE et al: Diagnosis and treatment of coronary disease: comparison of doctors' attitudes in the USA and the UK. *Lancet* 1988; 1: 750-753
18. Bernstein SJ, Hilborne LH, Leape LL et al: The appropriateness of use of coronary angiography in New York State. *JAMA* 1993; 269: 766-769
19. Winslow CM, Kosecoff JB, Chassin M et al: The appropriateness of performing coronary artery bypass surgery. *JAMA* 1988; 260: 505-509
20. Leape LL, Hilborne LH, Park RE et al: The appropriateness of use of coronary artery bypass surgery in New York State. *JAMA* 1993; 269: 753-760
21. Merrick NJ, Brook RH, Fink A et al: Use of carotid endarterectomy in five California Veterans Administration Medical Centers. *JAMA* 1986; 256: 2531-2535
22. Winslow CM, Solomon DH, Chassin MR et al: The appropriateness of carotid endarterectomy. *N Engl J Med* 1988; 318: 721-727
23. Leape LL, Park RE, Kahan JP et al: Group judgments of appropriateness: the effect of panel composition. *Qual Assur Health Care* 1992; 4: 151-159
24. Kahn KL, Kosecoff J, Chassin MR et al: The use and misuse of upper gastrointestinal endoscopy. *Ann Intern Med* 1988; 109: 664-670
25. Kahn KL, Kosecoff J, Chassin MR et al: Measuring the clinical appropriateness of the use of a procedure. Can we do it? *Med Care* 1988; 26: 415-422
26. Kahn KL, Park RE, Vennes J et al: Assigning appropriateness ratings for diagnostic upper gastrointestinal endoscopy using two different approaches. *Med Care* 1992; 30: 1016-1028
27. Leape LL, Park RE, Solomon DH et al: Does inappropriate use explain small-area variations in the use of health care services? *JAMA* 1990; 263: 669-672
28. Davidson G: Does inappropriate use explain small-area variations in the use of health care services? A critique. *Health Serv Res* 1993; 28: 389-400
29. Park RE: Does inappropriate use explain small-area variations in the use of health care services? A reply. *Health Serv Res* 1993; 28: 401-410
30. Cain KC, Diehr P: Does inappropriate use explain small-area variations in the use of health care services? A commentary. *Health Serv Res* 1993; 28: 411-418
31. Anderson GM, Sheps S, Cardiff K: *Evaluation of VI-CARE: a Utilization Management Program of the Greater Victoria Hospital Society* [working paper 93:1D], Health Policy Research Unit, University of British Columbia, Vancouver, 1993
32. *Barriers to Community Care: Summary Report*, Hospital Services Utilization Review Commission, Saskatoon, 1994
33. Kasian GF, Zinkiew K, Senthilselvan A: Inappropriate hospital bed days at a Canadian paediatric tertiary care centre. *Ann R Coll Physicians Surg Can* 1992; 25: 275-278
34. Gloor JE, KISSOON N, JOUBERT GI: Appropriateness of hospitalization in a Canadian paediatric hospital. *Pediatrics* 1993; 91: 70-74
35. Smith HE, Sheps S, Matheson DS: Assessing the utilization of in-patient facilities in a Canadian pediatric hospital. *Pediatrics* 1993; 92: 587-593
36. Kemper KJ: Medically inappropriate hospital use in a pediatric population. *N Engl J Med* 1988; 318: 1033-1037
37. Siu AL, Sonnenberg FA, Manning WG et al: Inappropriate use of hospitals in a randomized trial of health insurance plans. *N Engl J Med* 1986; 315: 1259-1266
38. Siu AL, Leibowitz A, Brook RH et al: Use of the hospital in a randomized trial of prepaid care. *JAMA* 1988; 259: 1343-1346
39. Grumbach K, Bodenheimer T: Reins or fences: a physician's view of cost containment. *Health Aff* 1990; 9 (4): 120-126
40. *Guidelines for Canadian Clinical Practice Guidelines*, Canadian Medical Association, Ottawa, 1994
41. Anderson GM: Implementing practice guidelines. *Can Med Assoc J* 1993; 148: 753-755
42. Davis DD, Thomson MA, Oxman AD et al: Evidence for the effectiveness of CME. *JAMA* 1992; 268: 1111-1117