Commission on the
Future of Health Care
in Canada

Commission sur
l'avenir des soins de santé
au Canada

DISCUSSION PAPER NO.35

Influence of Physician Payment Methods on the Efficiency of the Health Care System

by

Michel Grignon, Valérie Paris and Dominique Polton

with the co-operation of

Agnès Couffinhal and Bertrand Pierrard

CREDES, Paris, France

November 2002

Catalogue No. CP32-79/35-2002E-IN

ISBN 0-662-32963-5

Although the views expressed in the papers are those of the authors, each paper was subjected to an independent peer-review process. The Commission would like to thank the Institute of Health Services and Policy Research (IHSPR), of the Canadian Institute of Health Research, for their oversight and administration of the peer-review process. The work of the authors, the reviewers and IHSPR will serve to make these papers an important contribution to the Commission's work and its legacy.

Contents

Abstract

The study concerns the impact of physician payment methods on the costs of medical care.

There are three main methods of remunerating health care professionals: fee-for-service, capitation and salary. In theory, fee-for-service can lead the professional to generate an excessive volume of care, while lump-sum methods can jeopardize quality and equality of access. The magnitude of these respective risks basically depends on the nature and extent of the physician's discretionary power over the patient. If the physician simply influences demand, risks are limited through the control exercised by the patient; if the physician induces demand, the risks are greater.

There is no satisfactory empirical test of the nature of the physician's discretionary power. However, the majority of experimental studies comparing volume and quality of care according to payment method come to the conclusion that fee-for-service has a significant impact on the volume of service generated by primary care physicians (general practitioners and pediatricians). Conversely, capitation increases referrals and prescriptions, which can impose costs on other health care expenditure budgets. Finally, natural experiments in Norway show that the method of remuneration has less impact than the health characteristics of the patient.

In the real world (outside natural experiments), payment methods are rarely pure:

- First approach: adjust fee-for-service through financial incentives. In the United Kingdom, delegation of a pharmaceutical budget to general practitioners has made it possible to cut down on drug consumption. However, such financial incentives raise fears of conflict of interests between the physician's income and the quality of care he is required to provide. Other incentives are based on performance targets to be achieved, for example screening or immunization. The studies are contradictory, but most indicate that these incentives have a positive impact.
- Second approach: link the level of remuneration to a collective outcome. This type of approach seems ineffective if the group is too large (e.g. all physicians in a country), but payment by capitation of physician groups seems a promising avenue for reconciling cost containment and efficient organization of care.
- Third approach: non-financial incentives (protocols, assessment of practices). Few empirical studies have been conducted on this type of incentives, but initial results indicate a weak impact unless the "messages" are delivered by opinion leaders.

Executive Summary

A large body of economic literature attempts to model the delivery and consumption of health care as a market characterized by several informational asymmetries on which profit is one possible motive of a physician's practice.

Within this literature, we are interested in works dealing with the impact of physicians' modes of payment on the volume of services provided or prescribed to the patient, the quality of these services, and inequality of access to care.

We can distinguish three main modes of physician compensation: fee-for-service, capitation (lump sum payment per patient managed over a given period), and salary. Of course, this typology is simplistic, as most national health care systems combine these three modes of payment. In addition, the fee-for-service payment mode reflects, to some extent, a synthesis of medical care activity.

The following theoretical postulates apply: under a fee-for-service arrangement, the price received by the physician takes into account the fixed as well as the variable costs of his practice, which means that at the margin, an additional medical service will always increase the physician's monetary gain. Fee-for-service thus creates favourable conditions for the provision of more medical services than what the patient wants. If the physician has some control over the patient's demand, fee-for-service will increase the volume of medical services provided. By contrast, under a capitation system, the physician has no reason to produce more medical acts than necessary. On the contrary, the concern is that physicians will limit their effort by restricting patient access to care, especially the intangible resources required for the production of health care, such as the time devoted to the patient, the ability and willingness to listen, and the mental effort spent on the case – what is often called the quality of care. Finally, all lump sum payment arrangements raise the risk of case selection: because physicians receive the same amount for cases requiring varying levels of effort, they will prefer to take the simplest cases.

Also on a theoretical level, the robustness of these postulates will more or less depend on the influence the physician may exert on the consumption of care by patients: a high level of influence magnifies the expected negative effects (increase in the volume of care provided under fee-for-service, inferior quality under capitation), while a certain level of control on the patient's part will tend to dampen these effects. Two models of the physician-patient interaction account for this level of influence : imperfect competition and inducement. In the imperfect competition model, the patient is able to judge his physician's actions, but cannot sanction as severely behaviours that he perceives negatively than under conditions of perfect competition: the imperfect competition comes from the patient's bond with his regular physician, whom he knows personally. In the inducement model, the patient is under stronger influence from his physician and cannot control *ex post* the physician's behaviour, who will be in a position to convince the patient that he is receiving exactly the care he is seeking; any competition thus disappears.

The problem is that no satisfactory empirical test exists to evaluate these two theoretical models. We must then resort to observing physician-patient interactions in a real setting, under different payment methods. In order to make inferences from these observations on the impact of payment methods, we must ensure that the comparison reflects "all other things held constant." Two natural experiments (random assignment of a clinic's physicians among various modes of payment) confirm that fee-for-service increases the volume of care provided, but also that the level of payment seems to play a larger role than the method of payment. Results from "before-after" comparisons (sudden and exogenous change in payment method) show that the impact of the payment method strongly depends on the physician's environment, especially peer pressure. Another result is that payment by capitation increases the volume of drug prescriptions.

In the real world, methods of payment are seldom observed in pure form: most payers are seeking an optimal combination, based on the notion that the influence exerted by the physician depends on the type of care provided. A key innovation has been to introduce, in traditional methods of payment, corrective features at the margin targeting specific health outcomes, generally measured on a large population base (vaccination rate), or financial results. We refer to these practices as "performance payment." In the United Kingdom, for example, the portion of income that is independent of the level of services is 65 %, while fee-for-service payments represent 25 %, and performance payments 10 %.

Incentives linked to financial results are effective, particularly in decreasing hospitalisation rates, but they raise a conflict of interests that tends to lower the patient's level of confidence toward the physician and to alter the nature of the physician-patient interaction.

Today, payers prefer incentives linked to health outcomes, which is more compatible with medical ethics. In the United Kingdom, target payments seem to have a positive impact on vaccination or screening rates (simple observation "before-after"). Results from randomized studies are however contradictory: no impact on cancer detection among Medicaid clients, but a positive impact on the vaccination of Medicare clients or on quality indicators such as cesarean rates, length of hospital stays or certification rates for private-sector insured patients. Studies examining both the impact on costs and on quality are difficult to find.

Another innovation is to balance individual incentives with group incentives: under a fee-for-service payment regime, the unit value decreases as the collective volume increases. Applied to a an excessively large group of providers, this system raises a "free-rider" problem, where each individual provider increases his volume to compensate the reduction in price. In this case, controls over individual activity levels, or discounted rates are then necessary. Another solution is to reduce the size of the group targeted by incentives, on the model of the British Primary Care Groups, or the U.S. Medical Groups. The difficulty is to strike a proper balance between physicians' accountability for the group, and the financial and technical strength of the group, which suggests a minimum group size.

Finally, most group incentives rest on non-financial measures, such as the diffusion of best practices protocols, or peer-reviewed individual practices, even selective contracting based on commonly defined practices. However, the impact of these non-financial incentives has not yet been evaluated.

List of Abbreviations and Meaning

HMO:	Health Maintenance Organization Health plan with a limited panel of physicians (patient is not managed if he consults outside the panel)
MEDICAID:	U.S. public health insurance for the indigent
MEDICARE:	U.S. public health insurance for disabled persons and persons over 65
PCG/T:	Primary Care Group or Trust Group of general practitioners in the United Kingdom which is assigned an overall budget and manages all care for a given population
PPP:	Purchasing-power parity Unit for comparing expenditures between countries, with variability of exchange rates factored in
PPO:	Preferred Provider Organization Health plan with panel of recommended physicians (patient is only partially managed if he consults outside the panel)

General Question

> *How do payment mechanisms and incentives effect the utilization of health care services and their costs? To what extent can changes in policy, practice, and incentives change the cost structures and trajectories in health care?*

Introduction

The efficiency of the health care system affects the public economy and economic policy because health spending is always at least mutualized, if not socialized. The question of efficiency has long been confined to controlling the rising costs of health care, but it is now being raised in a more complex fashion: how to reconcile reasonable spending with the requirement for quality and fairness.

This question has generated a great deal of economic literature, modelling the characteristics of the medical care market and the behaviours of producers and consumers of care in order to suggest optimum contracts to the regulator for satisfying these three contradictory objectives (expenditure containment, quality, fairness). This theoretical literature, focussed on latent behaviours within the payer/patient/physician triad, has given rise to unresolved empirical questions, particularly as regard the extent of the physician's influence over the patient.

Our interest here is in the segment of this literature that addresses the impact of the method of physician payment on the volume of services the physician provides or prescribes for the patient, the quality of those services, and inequalities of access to care. The method of paying the physician is a true strategic choice, and there is no "natural" mode of payment. Because the physician is subject to an obligation of means and not of outcome (for reasons of uncertainty), he is not in fact a classic producer who sells an outcome or end-product and takes the risk that this end-product may be worth less than its production cost. Besides, we are talking about the remuneration of physicians, which is what makes them factors of production, and not the retail price of medical services. And yet, the physician is not a mere factor of production, because his degree of autonomy in organizing the elementary actions that contribute to the production of his services is a component of the quality of the services provided.

After a brief presentation of the different methods of physician remuneration, Part 1 will summarize the economic literature on their impact on health care costs. We explain the theoretical issues behind the various payment mechanisms and examine the extent to which the empirical studies conducted in various national contexts confirm these theoretical postulates.

Economists hypothesize that, while the physician's profit is not the sole motive behind a medical decision, it is an element of it. This hypothesis is shared and supported by medical sociology, which shows that when the consequences of a medical decision for a patient's well-being are not clear, the trade-offs for the physician in terms of leisure, reputation or income take on greater importance. It is then no longer certain that the patient is receiving the exact amount of care that his treatment requires.

Economic analysis implicitly raises questions like : Given the motivations of physicians, what impact does the method of payment have on medical expenditure? Such an attitude does not mean that motives other than profit (professional standards, respect of peers, or the patient's interest) have no importance, but rather that we can hypothesize that payment methods have no impact on them. In reality, this hypothesis is certainly erroneous, and experiences of changing physicians' modes of payment can show us why the mechanisms adopted can depart significantly from economic recommendations, particularly because the payment method can interfere with these other motivations of the physician (Part 2).

The Various Mechanisms of Physician Payment

It is customary to distinguish three main methods of physician remuneration:

- Fee-for-service (FFS);
- Capitation, which is a lump-sum payment per patient managed for a given period of time; and
- Salary.

Different countries use one or another of these three forms of compensation for general practitioners (GPs). Specialists generally receive a fee-for-service or are on salary.

This traditional typology raises certain questions.

First of all, the overall remuneration of a physician can combine a number of methods of payment: these mixed or blended payments are the practice in some health care systems, as we shall see in Part 2.

Next, while it is convenient to oppose these three mechanisms, they can also be considered on a rising scale of inclusiveness, in which intermediate situations can also be imagined:

- At one extreme, the fee-for-service is the most specific and least inclusive form of payment for a medical service. Yet, it still constitutes a fixed price: when a consultation is paid at a flat rate, as in France for example, it makes no difference whether it is long or short, whether a clinical examination is performed or a prescription is simply renewed. And one notes that the degrees of inclusiveness of the nomenclatures used in different countries for pricing services vary, and that a comparison of those nomenclatures could be one way of describing different types of FFS.
- Capitation adds another level of inclusiveness, since the physician is paid the same amount for a patient regardless of the quantity of services provided. The physician's income varies only according to the number of patients taken.
- Salary is the all-inclusive extreme form of payment, where income does not change regardless of the physician's level of activity – in terms of number of patients in care or quantity of services provided per patient.

On this scale, one can imagine intermediate situations, for example, between fee-for-service and comprehensive capitation: a lump-sum payment per episode of care, or a lump-sum payment for monitoring a chronic pathology over a given period of time. A few experiments with these types of mechanisms will be described in Part 2.

In summary, even though the three "archetypal" methods of payment are traditionally opposed – a practice we shall adopt in our review of the economic literature below – there is no reason why intermediate methods cannot be devised.

Finally, as regard the forms of payment of health professionals, we are witnessing the emergence of mechanisms that do not fit any of the three models mentioned above, but are similar to results-sharing mechanisms, whether those results are financial in nature (cost containment) or target the quality of the medical service.

This type of mechanism, of which we provide examples in Part 2, might be considered a fourth mode of compensation, as an alternative to the three previous ones. That is, in fact, the approach taken by some authors who are now distinguishing between four methods, often labelling the fourth "target payments". In practice, however, this type of results-sharing is only complementary and peripheral to the main method of payment. It is thus difficult to consider it as an alternative – it is, rather, an adjunct to the three main methods of payment. In future, it could possibly become a greater factor in physicians' remuneration.

Part 1:
The Economic Literature

We present here a review of what the economic literature has to offer on the impact of the method of payment of physicians' services on the total health care budget.

The Facts and the Simple Theoretical Postulates

Based on aggregated data, a country's total health expenditure is affected by supply variables in general and the method of physician payment in particular. In 19 OECD countries, Gerdtham et al. (1992) find that, all other things being equal, "physician fee-for-service" (i.e. a variable indicating that this is the majority payment method in the country) increases health spending by 11%. A more recent study employing panel data (for 22 OECD countries between 1970 and 1991) no longer finds that FFS per se has any direct effect, but rather an indirect one: medical density increases expenditure when physicians are paid on a fee-per-service basis (Gerdtham et al. 1998). However, these relations observed on aggregate data are difficult to interpret, for it is possible that total national spending and specific institutional arrangements such as FFS are both positively correlated to a third variable, such as the fact that physicians practice in isolation.

The economic literature thus attempts to describe the micro-economic process by which the method of physician payment might have an influence on medical spending; to formalize this process, a payment method is characterized by its degree of inclusiveness. In the rest of this first part, we will focus on the opposition between systems of inclusive or lump-sum payment (capitation and salary) and fee-for-service schemes.

First, we can identify how payment methods are expected to impact on health care costs by isolating the other objectives of the health system (quality and fairness). To do so, let us briefly describe the physician's activities: to care for patients, the physician uses certain resources. Resources such as an office are necessary whether there are patients or not; others are used only if the are clients. In reality, the physician does not set his price himself: he receives it from a payer. The resources used in production are of two types: quantifiable material resources and non-material resources that are difficult to quantify. An example of the former would be the consumables of the medical practice; examples of the latter would be the time the physician devotes to each patient, his competence (a function of his continuing training, the effort he makes to keep up to date), his ability to listen to the patient, and his mental effort. In economics, the level of non-material resources is usually referred to as the "physician's effort". An important difference between material and non-material resources is that the payer can take the former into account in the payment scheme, whereas doing so for the latter is more difficult. It is difficult for a payer to compensate a resource when there is no proof that the resource was actually used.

With in hand this simplifying model, what then can be said about the methods of payment?

In fee-for-service, the price takes into account fixed costs and variable costs (if not, the physician refuses to work) and so is higher than the average cost of the variable resources alone: thus caring for an extra patient brings the physician a financial profit. If, in addition, the

physician has some power over the patient's demand, FFS creates favourable conditions for a larger quantity (and hence expenditure) of services than the patient may wish to receive, or than may have been medically necessary. Under this model – the physician has some power over the patient – if the payer reduces the unit price, the physician will respond by increasing quantity[1] to maintain his profit; this explains why a reduction in the price of care leads to an increase in the supply of services.

Under a capitation formula, on the other hand, the physician is not reimbursed proportionally to the resources committed, and so has no interest in multiplying his services, since he would then lose money or waste efforts. He prefers to treat the patient using the least possible amount of resources to achieve a given level of quality.

The Cost-Quality-Fairness Trade-off: Status of the Problem

If the question were limited to cost containment, one might wonder why all regulators have not opted for some form of lump-sum payment. Obviously, there are historical reasons that support fee-for-service in spite of the evident superiority of the lump-sum payment as an instrument of budget control (see, for example, Emery, Auld and Lu 1999). However, in the culturally homogenous United States, one notes that different payers have adopted different arrangements: some prefer a form of capitation, while others prefer FFS; the latter are surviving, and even seem to be recapturing market share lately.

So there are certainly reasons of cost-effectiveness that also explain why the lump-sum payment method is not universally adopted:

- An initial reason is related to the quality of the medical service provided: while lump sum payments offer a structural means of resolving the problem of increased volumes of services delivered by physicians, it raises the risk that the physician will provide too few services (consultations, examinations) and too few non-material resources (shorter consultations, less attention) to the patient. Of course, the regulator can impose a minimum level of services per patient, but it will be difficult to control the level of non-material resources. If the physician is paid to manage a patient and is not sanctioned in any way,[2] he has an incentive to collect the lump sum payment and not provide the corresponding services and effort. The economic literature tries to shed light on this point and to define the circumstances where the lump sum payment method can jeopardize medical quality. More generally, lump sum payments can be seen as a delegation to the physician of responsibility for the optimum organization of services that are factored into health care, whereas with fee-for-service payment, this role is shared between physician and regulator. Where there are multiple payers, FFS raises an additional problem of inconsistency among pricing regimes, but to our knowledge this problem has been studied only for hospital services (Tai-Seale, Rice and Stearns 1998).
- This quality problem can also have consequences for spending: the physician is not only a producer of services, he is also a prescriber. Capitation is generally accompanied by a referral responsibility for the primary care physician (PCP): he makes decisions concerning quantities of analyses, medications and specialized or hospital consultations. If the physician reduces the intensity of his effort because he is paid in a lump sum fashion, even as he

remains concerned for his patient's health, he can very easily generate a high volume of health care costs by increasing his prescription rate.

- Finally there is the issue of universal access to care. The lump sum formula always entails a financial risk for the physician: if doctors are averse to uncertainty, they will refuse to take patients who incur highly variable *ex ante* costs. If doctors have better selection skills than payers, they will make money by treating individuals paying a high lump sum and refusing to provide care to others (e.g. see Mulligan 2002).

Ultimately, the question is then: what form of payment can best reconcile the objectives of expenditure control, quality of medical service and universality of access to care? We will first look at the theoretical models that explain the physician's discretionary power over the patient's consumption, models which conflict on the potential ability of the physician to "induce" patient demand (Evans 1974). This ability has some important consequences for our study, for it increases the risk of lump-sum payments leading to a reduction in the medical services received by patients. This discussion concludes with an analysis of the physician's behaviour as a prescriber of medical services that he does not provide himself.

The Cost-Quality-Fairness Trade-off: Discretionary Power of the Physician

Theoretical approach

We saw earlier that the physician will profit from fee-for-service if he is able to impose on the patient more services than the latter wants or needs. How is this made possible?

There seems to be broad agreement that, in interacting with his patients, a physician does indeed have a superiority that allows him to influence the consumption of his services.

One reason for the physician's superiority is that he has a knowledge advantage over his patient. There is abundant theoretical literature debating the fact that superior knowledge is sufficient for an expert to impose an agency relationship, that is, a relationship in which the lay person puts himself entirely in the expert's hands. This literature concludes that in a competitive context among experts, if the lay person can get a fairly accurate idea of the link between the expert's effort and his results, then superiority of knowledge is not enough to establish an agency relationship. The expert is selling so-called "experience" goods, the quality of which will ultimately be determined by the buyer. On the other hand, if some fundamental uncertainty prevents the buyer from forming a notion of how the expert's effort impacts upon the result he obtains, then an agency relationship can exist. The expert is then selling "confidence" goods, as the lay buyer has to have absolute trust in him.

What about medical goods? Rice (1993) advances the radical idea that there simply is no such thing as informed and independent demand for care on the part of the patient. He bases his argument on the fact that, according to the findings of Rand researchers, the additional care obtained by the best-insured patients is not, in the view of independent experts, really appropriate given their health status.

The idea that there is no independent patient demand for care is no doubt excessive (we all make health decisions independently of our physicians), but on the other hand, it is true that part of the medical services provided by the physician is not observable. This difficulty in assessing the medical services provided is undoubtedly one of the reasons why physicians in developed countries are not generally authorized to own pharmacies or sell medications: if they could increase their earnings in this way, they might abuse their expert's position by persuading patients that it is in their interest to consume drugs.

One can also make a few *a priori* predictions. For example, it is certainly more difficult to observe the quality of a specialist's services whom we rarely see: to resolve the asymmetry of information, the regulator will prefer to have the quality of this specialist's services observed by another physician through referral[3] by a primary care physician. However, empirical proof of a degree of inducement on the part of physicians is very difficult to establish: even if it is determined that different treatments were provided by two doctors to two identical patients, and that a connection in found between this variation in practice and the economic interest of the physician, it is still necessary to ascertain whether the variation is not perceived by the patient as a deterioration of services received.

In fact, the agency relationship assumes not only that the physician imposes his views on the patient, but also that the patient does not realize it (he trusts the expert). As McGuire (2000) demonstrates, however, one can easily imagine a physician imposing his views even if the patient realizes that the physician is not doing exactly what the patient wants, or what the physician would do for himself. In technical terms, an agency relationship is not necessary for the physician to hold some discretionary power over the patient: it is enough for the physician to be in a situation of imperfect competition with patients for the latter to accept from a service delivery that is not totally satisfactory.

One good reason explaining this imperfect competition between physician and patient is the latter's attachment to certain characteristics of his doctor, which results in the patient not seeking out another doctor. The patient's attachment for "his" doctor (family doctor, the usual source of care) is abundantly documented, and is explained by the fact that long-term familiarity allows the physician to specialize in the patient. Obviously, this attachment of the patient to his doctor applies more to the GP (or primary care physician) or chronic care specialist.

As a result of this attachment, physicians are not perfect substitutes for each other: each doctor more or less has a free hand with his patient and can impose different choices upon him than those the patient would spontaneously select if he were able to replace his physician with no adverse effects. In economic terms, having this freedom translates into "having market power" and behaving like a monopoly. The physician is precisely in a situation of monopolistic competition with his patient: like a monopoly, the physician can take optimum advantage of his patient's willingness to pay by playing on prices and quantities (whereas in a situation of perfect competition, there is only one price, that of the marketplace); but for all that, if the physician goes too far (strays too much from his patient's expectations or needs), the patient can leave him for another doctor.

Attachment and agency can even be said to complement each other in the way they operate: when the patient sees the doctor only rarely, he has little attachment to him, but also has more chance of being in an agency relationship and having to bestow blind trust; on the other hand, when the patient is in a long-term relationship with his doctor, it is easier for him to observe the quality of services received, but more difficult to act in a competitive manner and oppose services that leave him dissatisfied.

What impact might the nature of this discretionary power of the physician have on the problem that concerns us here? Schematically, the closer one gets to a situation of monopolistic competition, the more the physician has to take into account the risk of the patient's departure: with fee-for-service, that threat limits the physician's power to provide too many services the patient does not want; with capitation, it limits the physician's power to excessively reduce his volume of services, and above all the effort he devotes to his patient. If an individual's attachment to a given doctor is a constant variable, the physician arbitrates between the length of his client roster and the benefit he derives from each client listed on it. The physician is then said to influence the patient's consumption.

On the other hand, the closer we get to an agency relationship, the less likely the patient is to pose a threat of departure for the physician, who can provide too many services if they are paid on a per-service basis, or not enough effort if he is receiving a lump-sum payment. In this case, the physician is said to induce demand.

A large body of literature attempts to prove empirically, in an indirect way, that the physician has absolute discretionary power, since the patient is not in control. A classic test is to measure the response of physicians to a variation in medical density in a FFS context. Remember that, in the case of inducement, the physician has no competition: if medical density increases, the physician will have on average a lower number of patients (where inducement is involved, the quality of the physician has little impact on his attractiveness, for he is not in competition) and so will react to preserve his income by increasing the volume of services used by his patients. But in a context of monopolistic competition, higher medical density raises the risk of the patient leaving the physician (a break being less costly to the patient), so the physician will not create a wider gap between the consumption wanted by the patient and the services provided.

Apart from the empirical issues (the truly exogenous nature of the increase in density, the difficulty of observing major variations), this test is not perfect; it is based on the arguable hypothesis that, in monopolistic competition, an increase in medical density facilitates the departure of unhappy patients; in fact, if a patient's attachment is due to personal characteristics of the physician that are not affected by medical density, one can find an increase in services following an increase in medical density in a context of influence. Pauly and Satterthwaite (1981) raise a similar objection: a higher density of physicians increases the capacity of each physician to influence demand via a process of medicalization of public expectations.

Overall, it does not seem possible to decide on the degree of inducement on the basis of measurements of physician activity. Direct surveys of patients' level of satisfaction and their motives for choosing or leaving a physician would doubtless be of great interest. But it must be acknowledged that there has been no measurement of the degree of inducement based on such

direct statements by patients. An invaluable indication is provided by the fact that, in the United States, the sickest persons insured by Medicare prefer to remain in FFS plans and retain the freedom to access the physician of their choice instead of opting for the more extensive coverage proposed by HMOs. This is reason to suspect that the physician's latitude to persuade the patient that what he needs is precisely what will maximize the physician's income may increase with the patient's degree of morbidity.

Empirical approaches

Faced with varying medical practices, it is impossible to know beforehand whether a given practice is excessive or whether another is restrictive. Practices can be compared after the fact to some technical standard, such as a recommendation, or to a statistical norm, such as an average practice, but there is nothing to validate such comparisons.

Hence we can only measure the impact of the method of physician payment on the volume of services consumed by the patient, quality and universality of access; but nothing can be said on the proper mode of payment based of such studies. To measure this impact, researchers look for comparisons where all other things are held equal except the payment method. Scott and Hall (1995) have found seven empirical studies published in English on this theme, noting that this small number is linked to the fact that these are "opportunistic" studies which make use of data compiled for other purposes. Gosden et al. (2001) list another similar study.

For comparisons to be meaningful we must first be sure that the contexts are comparable: comparing the impact of different compensation systems in areas that are far removed from each other geographically or socially does not make much sense. Therefore, one must find homogeneous contexts with sufficient variability of payment methods. A second problem with this type of comparison is that it is difficult to ensure that all things apart from the payment method are in fact equal: the fact that a patient chooses a physician who is paid on a per-service basis or by capitation can be correlated with some unobservable characteristics that are themselves related to the use of the health care system. Similarly, physicians can specialize by type of contract (Glied 1998), again for reasons not unrelated to their work methods and the volume of services they provide. Finally, it is difficult to be sure that the level of payment is the same from one method to the next, and that one is not measuring the combined effect of a difference in payment method and level.

Hickson et al. (1987) conducted an experiment controlling for the first two problems: 15 pediatricians, members of the same clinic, were randomly allocated a mode of payment (salary or fee-for-service). The study compares the volume of preventive visits to physicians on salary and those paid on a per-service basis: with controlled observable characteristics of physicians and patients, the pediatricians paid on a per-service basis delivered significantly more visits.

This additional volume also reflects an ambiguous variation in quality: patients treated by physicians paid per service say they receive better follow-up, but the volume of preventive visits suggested by these physicians departs more from the recommendations of the American Academy of Pediatrics than the volume observed for salaried physicians. According to

Gosden et al. (2001), this average difference between salaried physicians and those paid on a per-service basis involves two different phenomena: visits in excess of the recommended number are fewer for patients of salaried physicians than for those of FFS physicians, but the recommended visits missed by the former are more numerous than those missed by the latter.

Another study with physicians randomly assigned to payment methods is that of Davidson et al. (1992), on 80 physicians treating children under the U.S. Medicaid program. The study monitors three groups: a control group, paid on a per-service basis; a fee-for-service group receiving higher rates; and a group paid by capitation. The volume of primary care sessions increased for the two tests groups, underscoring the impact of the level of remuneration regardless of payment method. However it increases more in the FFS group (+0.8-0.9) than in the capitation group (+0.5-0.6).

The price to be paid for these methodological gains (random assignment of physicians to payment methods) is obviously the sample weakness and results that cannot be easily generalized.

We therefore turn to some "before/after" comparisons, in which a payer changes the payment method of a single group of physicians and patients. In this case, we can assume a real exogenous change in payment method and observe its effect free of any bias. Furthermore, in implementing reforms, efforts are made to maintain physicians' income at a constant a level, so the effect of payment method can be measured separately from the level effect.

An initial study of this type was done by Stearns et al. (1992). It reached the conclusion that capitation has a weak impact on volume of services provided: in a health plan for Wisconsin government employees, PCPs moved from fee-for-service to individual capitation. At the same time, the unit price of specialists' services were greatly reduced. Paradoxically, there was an increase in the number of ambulatory visits, but it was visits to specialists that increased, and there is reason to believe that specialists increased their self-referral to compensate for the drop in service rates. The volume of visits to GPs remained unchanged, so the transition to capitation had no effect. It is possible however that the experiment was disrupted by another factor: at the same time, the GPs were allowed to share in the profits from reduced hospital expenditures. It is thus possible that they were motivated by this profit-sharing and stepped up their effort in primary care medicine in order to limit hospitalizations. The study in fact shows a perceptible drop in the hospitalization rate (though not in the length of stay). This study is not completely free of bias, because certain primary care physicians opted out of the plan, or refused plan patients (or certain services). Some patients may also have left the plan, for in addition to the change in physician remuneration, there was a transition to an exclusive HMO approach (an insured person is not managed if she consults outside the panel), whereas the plan previously functioned as a PPO (reduced management of an insured person who consults outside the panel).

An experiment offering better control of these voluntary separations was conducted in Denmark (Krasnik 1990). General practitioners in the city of Copenhagen moved from pure capitation to mixed capitation and FFS, while regional/county physicians remained under capitation. The author was thus able to compare changes in volumes in the "test" group

(the Copenhagen GPs) and the "control" group. There was a significant increase in the volume of services provided by the physicians in the test group with the transition to fee-for-service.

This effect is not universal, however, and the Norwegian experience seems less conclusive on the impact of fee-for-service. Using records of contacts by patient and by physician[4] for Norway, Kristiansen and Holtedahl (1993) show that FFS increases the volume of home visits, all other things equal (although the patients' state of health was unknown), but that the effect of the payment method is weak compared with the influence of a patient's individual characteristics: the physician remains more motivated by health concerns than by his income.[5] This precedence of the clinical over the economic in Norway seems to be confirmed by Grytten and Sorensen (2001), who show that physicians paid on a per-service basis do not perform more examinations than salaried physicians, even when one controls for factors that lead physicians to choose a given type of contract. The authors claim that this primacy of clinical concerns is mainly due to strong peer control in Norway, which prevents the physician from taking advantage of his position of superiority over the patient.

In addition to the damage caused in terms of average quality, inducement can also be damaging to the fairness of the system. Vulnerable segments of populations (children, seniors, the mentally ill, the poor, linguistic minorities) have more difficulty observing and tracking the quality of the physician's service, so it is tempting for the physician to use this ability to induce demand primarily on members of these groups. It is noteworthy that HMOs ration clinical quality for these groups, while leaving average quality unchanged (Miller and Luft 1997).

The Cost-Quality-Fairness Trade-off: The Physician's Capacity to Increase Prescriptions

Referrals and prescriptions are important components of the physician's practice and of total health care costs. In France, 33% of visits to a physician are the result of a recommendation from another physician. In the United Kingdom, 5% of PCP visits lead to a referral (Coulter, 1992), and there is a very large variation in physicians' practices which cannot be explained by patient traits (summarized by Scott 2000).

A simple but sound theory of physician behaviour is that payment via capitation or salary encourages the physician to transfer his own work to others and increase the volume of prescriptions so as to cut down on unremunerated effort (time spent with the patient, listening, auscultation).

This prediction has wide empirical confirmation. The study by Krasnik et al. (1990) shows that when GPs changed to FFS, their rates of referral and prescription dropped substantially. This study uses a control sample whose payment method remained unchanged over the same period.

International comparisons suggest that less time is spent in consultation when payment is by capitation: according to Sandier (1990), while the French GP, paid per service, devoted an average of 15 minutes (in 1979) to each consultation, the English GP, paid by capitation, was content with 8 minutes (in 1987). However, in a homogenous context and using individual data,

there is no confirmation that capitation or salary reduces consultation time: Kristiansen and Mooney (1993), using Norwegian data, even find that GPs paid per service spend less time in consultation than those on salary, although the difference is not significant.

One way to prevent capitation from raising the cost of prescribed care is to make the prescriber responsible for a patient's total care budget. This is a form of vertical integration of care, assigned to the primary care physician; we will outline it below.

Part 2:
In Search of the Optimum Method of Payment: Lessons from a Variety of Experiences

The question of incentives associated with payment methods has received growing attention since health care systems have been faced with major issues of balancing cost, quality and fairness. Through general reforms or local experiments, various solutions have been tried in an attempt to reconcile these potentially contradictory objectives by means of the methods of compensating professionals. These solutions will be reviewed in this part, distinguishing between a number of approaches.

One approach tries to combine various incentives within individual physician remuneration, using blended payment systems that include lump sum, capitation and fee-for-service. Specific financial incentives may be incorporated into these, the most innovative being mechanisms for sharing in financial or health outcomes (performance-based payment).

A second approach deals with different financial incentive systems at the individual and collective levels of the medical profession.

The conflicts of interests resulting from this second approach are such that it lends itself more to smaller, rather than very large groups of professionals; this has led to a third approach, that of making physicians' groups responsible for reconciling financial and quality objectives.

A final approach to the best cost-quality ratio is to combine financial incentives with incentives of other kinds, such as information provided to the physician, formulation of practice benchmarks, and different tools for their dissemination. The collegial dimension of the medical profession can facilitate change in practices, and this affords a means of potentially gaining the interest of the groups or networks mentioned above.

Incentive Combinations within Individual Physician Remuneration

Before turning to mixed payment systems, it should be noted that no method of payment has emerged as an overall alternative to the standard three: fee-for-service, capitation and salary.

Episode-of-care payment, transposed from hospital payment to diagnostic-related groupings (DRG), was considered by Medicare in the United States in the mid-1980s, but seems to have had few concrete applications in ambulatory medicine. It can be envisaged for technical care in acute episodes, e.g. a surgeon receiving comprehensive remuneration for pre-op, surgery and post-op (Robinson 2001), and it has been studied in cardiology, e.g. for coronary bypasses (Showstack, Garnick and Rosenfeld 1987). But extending this mode of payment to all medical activities raises numerous technical problems.

In terms of practical applications, therefore, innovation and inventiveness have mainly developed around the manner of organizing different elements of payment and using marginal incentives for a minor percentage of remuneration.

One interesting example of mixed payment is that provided to British general practitioners, whose compensation includes (Redbook 2002):

- A basic allowance to cover a fraction of operating costs, in the form of a sliding lump sum per patient based on the number of patients. An additional allowance offsets extra costs of practising in a rural area. Top-ups are also paid based on the physician's experience ("seniority" fees).
- A capitation per patient on the GP's roster, with three levels according to age (under 65, 65 to 74, and 75 and over); patients in disadvantaged areas also give rise to an additional capitation reflecting their greater needs.
- FFS for night visits, vaccinations, emergency consultations for colleagues' patients, minor surgery, as well as more comprehensive specific fees for monitoring contraception, child health and maternity medical services.
- Target payments for cervical cancer screening (payment depends on the rate of women who have had a smear in the previous five years) and for immunization of two-year-old and pre-school children.

The portion independent of level of service makes up 65% (50% for capitation and 15% for basic allowances), FFS accounts for 25% and target payments for 10%. In addition, U.K. physicians working in poorly served rural areas can dispense medications to their patients and receive specific fees for this service.

Have incentive payments had the results expected? In fact, it appears that activities paid on a per-service basis have increased most of the time, though not always (for example the proportion of physicians agreeing to monitor pregnancies to term decreased from 1960 to 1980). In any case, it is not possible to apportion within this increased activity the role of financial incentives and that of other factors related to the general evolution of professional practices and patient demand (Whynes and Baines 1998).

While in the National Health Service (NHS) certain productivity incentives have been added to what were originally lump-sum payments, we are now seeing the introduction of packages or lump sums into FFS systems, for example:

- In Quebec, for hospital specialists, mixed remuneration has taken the form of introducing a lump-sum portion, and in return paying for specific services at a fraction of their rate and eliminating billing for related services.
- In France, GPs who volunteer to be "referring physicians" receive an annual lump sum of 46 Euros for general patient monitoring, in addition to billing per service (although this option has been chosen by only 10% of GPs and 1% of patients).

The most innovative forms of blended payments incorporate elements of sharing in the outcomes (target payment). Those outcomes can relate to health matters (as in the British example above), but they can also be financial, i.e. associated with cost control. Such incentives are used by "managed care" organizations in the United States to reduce specialist referrals and use of hospitals. They are generally based on a formula that gives the physician an interest in the organization's financial performance. What assessment can be made of such outcome-sharing experiments?

Interest in Financial Outcomes

The 1991 reform in the United Kingdom ("fundholding") gave British general practitioners a financial interest in medication and programmed surgery expenditures for patients on their roster: the volunteer GP was assigned a budget at the start of a period for all his patients and could retain any year-end surpluses. Theoretically, such surpluses are not income, since the GP has to reinvest them in his practice, but in reality it is difficult to separate the physician's personal income from the revenue of his practice. Evaluation of this experiment has shown that physicians did indeed realize savings on their medication budgets, but not much on programmed surgery, one of the reasons for this being the quasi-monopoly held by British hospitals (Coulter 1995).

A study by Debrock and Arnould (1992) measured the effect of financial incentives to the physician by controlling for existing group incentives. Their conclusions for the 35 Illinois managed care organizations studied between 1985 and 1987 were that individual financial incentives reduced the hospital admission rate by 16% and cut the number of consultations per client by half.

While giving physicians an interest in limiting expenditures seems an effective tactic, the conflict of interests it can raise between professional ethics and the financial interest of the practitioner is negatively perceived by the profession, and above all by the general public:

- A survey of 766 U.S. physicians, 58% of whom were subject to group or individual incentives, showed that 57% of practitioners felt pressure to limit referrals and 17% thought that this could negatively affect the quality of care management (Grumbach, Osmond, Vranizan 1998, quoted by Armour, Pitts, Maclean et al. 2001). This survey also gives an idea of the magnitude of such bonuses: according to the physicians surveyed, the bonus accounts for about 7% of their median income.
- Although not particularly sizeable, these financial incentives are raising concerns among patients, who fear that they may limit the professional autonomy of their physician. There are few empirical indications of the extent of this concern, apart from many legal actions concerning financial incentives of physicians. In 1996, U.S. legislation set certain limits on contracts between managed care organizations and physicians and made their publication mandatory. According to Hall, Dugan, Balkrishnan et al. (2002), the publication of incentives has the effect of slightly increasing patient confidence if the physician is paid by capitation, but has no effect if he is paid per service.

Many authors, insurers and HMOs now think that performance incentives must be used to improve quality by focussing on health as opposed to financial objectives. According to Rosser and Kasperski (2002), Ontario physicians say they are in favour of a transition to capitation provided that it is mitigated by fees for certain services such as being on call, and by premiums for attaining prevention objectives.

Interest in Health Outcomes

This does not directly concern health objectives – a seductive but not very practicable idea – but rather the management of the population concerned. Prevention, both primary (immunization) and secondary (screening), is its most frequent domain of application.

The example from the U.K. cited above seems to show that these "target payments" have a positive impact: according to the British Department of Health (quoted by Scott and Hall 1995), the proportion of GPs reaching the objective set rose in five months, for vaccination, from 55% to 70% (objective: a vaccination rate of 90% or 70%) and for smears, from 53% to 67% (objective: for 80% of women from 25 to 64 to have a smear within five years). Of course, this simple observation of before/after results cannot control for other possible explanations and categorically claim that the increase is due to financial incentive. That is the point made by Armour et al. (2001) regarding another experiment in Northern Ireland, involving childhood immunization: there again, while there was a substantial increase in the vaccination rate during this period, the pure effect of this measure is not distinguished from other concomitant effects.

A bonus experiment to improve cancer screening, set up by a managed care organization for people on Medicaid, has been evaluated (Hillman, Ripley, Goldfard 1998). A six-month bonus ranging from 10 to 20% of capitation for all female clients was paid to the six medical practices achieving the highest screening rate. Fifty-two practices were randomly assigned to an intervention group (with incentives) or a control group. The study revealed no significant difference in screening rates between the two groups.

However, positive conclusions were drawn from an incentive experiment for flu vaccination for people over age 65 involving 54 New York practices, which was evaluated via a controlled test (Kouides, Benett and Lewis 1998). In addition to remuneration for the vaccine ($8), the physicians received $0.80 per vaccine beyond a vaccination rate of 70% for persons 65 years of age and over and $1.60 for a rate over 80%. The immunization rate in the group receiving the bonus (69%) was six points higher than that of the control group, and 52% of the practices receiving the bonus had an immunization rate of 70%, versus 44% in the control group. An analysis controlling for a number of environmental factors suggests to the authors that the introduction of bonuses served to boost the immunization rate of the elderly by 7%.

A recent trend seems to be to tie the bonus to various quality indicators. In one experiment done by Aetna (Hanchak 1997), groups of contracting physicians received bonuses based on their score on a quality-of-care index incorporating a number of aspects, including patient satisfaction, appropriateness of treatment, and efficiency. The authors concluded that there was a positive impact on all aspects of quality (2% decline in the cesarean rate, 25% decline in average

length of hospital stay, 85% increase in biopsies, and a certification rate approaching 95% at the end of the program) except patient satisfaction, which does not seem to have improved.

Armour, Pitts, Maclean et al. (2001) conclude from a review of the literature that the impacts of incentives are contradictory, and do not allow for any easy final verdict. They also note that studies rarely deal with volume effects and quality effects at the same time, and in this they see a recognition of some kind of trade-off between cost containment and quality.

Different Incentive Combinations at the Individual and Group Levels

A second approach is to attempt, not to counterbalance different incentives within individual remuneration, but to offset individual incentives with collective incentives.

This is illustrated, for example, in the limiting overall envelope systems associated with FFS. Germany opted for such a system in the early 1980s, and other countries or regions have adapted the same principle, using varying methods suited to their own context.

German physicians are paid per service but are bound by collective moderation: fees-for-service for all physicians in a Land are capped, and the amount is distributed among practitioners by the physicians' unions proportionally to the activity of each, expressed in points on a list of services. In other words, the unit value of the point is not known in advance, but is estimated after the fact by dividing the total envelope by the accumulated number of points of all physicians (Busse 2000).

The direct remuneration of practitioners through the funds in the form of a cumulative payment for a given period (in Germany, per quarter) affects the operational nature of such a system. However, certain systems in which these conditions are not met have also introduced expenditure caps (Quebec, France, the United States for Medicare): the difference is that the unit rate cannot be automatically adjusted in real time, but must be decided – and negotiated – for the following year, based on actual outcomes. Obviously, this method is likely to be less effective in terms of strict adherence to the budget envelope.

In theory, this type of mechanism leads to a conflict, for each physician, between the collective interest in moderating volumes (therefore achieving a higher point value and a level of income that is unchanged overall with a lower output of work) and the individual interest which continues to be to maximize one's own volume of activity at a given unit price. Hence, it encourages "opportunistic" behaviours, with each physician thinking that he will be able to profit personally from collective moderation if he raises his volume of activity, something which ultimately leads to inflated volumes, contrary to the target objective.

Clearly, this is related to the dilution of responsibility within a very large collective. Each physician on his own has no chance of affecting the unit value through virtuous behaviour; on the contrary, he has every interest in being this "opportunistic" element in the great anonymous collective. We shall return to the question of the size of physicians' groups later.

The theoretical postulates of systems of this type – ongoing inflation of volumes and a correlative decline in unit prices – seem indeed to be confirmed by the observations in Germany (even though one can debate the proportion of the increased volumes that is to be attributed to opportunistic physician behaviour or to other factors: demand, technological progress, etc.).

So a complementary mechanism is needed to offset this incentive to individual behaviours that are too production-oriented. That mechanism can include:

- Control: this is the case in Germany, where individual activity is examined by the fund physicians. Incidentally, one might think that this control will be more effective as it is organized by the profession itself; but its feasibility depends on the profession's availability to co-manage financial constraints with the regulators, as demonstrated *a contrario* by the French example.
- Incentive: as in Quebec, with sliding-scale rates beyond a threshold of service volume; this has the effect of curtailing intensive activities.

The Physicians' Group as Locus for Integrating Clinical and Economic Objectives

In different countries we seem to be witnessing convergence toward a more collective practice of medicine, in various forms – networks, single-discipline or multidisciplinary groups.

Since the mid-1990s, this trend has emerged in such different national contexts as the British NHS and American "managed care" (Bocognano, Couffinhal, Grignon et al. 1998). In England, the 1997 reform instituting the Primary Care Groups or Trusts (PCG/Ts) made the first spontaneous associations of fundholder GPs a general fact of life; in the more competitive local contexts of the United States, physician groups have developed very rapidly to contract with health insurers seeking to build networks, and in extreme cases even to incorporate the insurer's functions themselves.

In both cases, the approach is to make these groups into highly structured networks with their own management infrastructure.

These groups are sometimes given financial responsibility and delegated management. For example, the British PCTs are remunerated via comprehensive payments, on condition that they distribute this funding among their various constituent medical practices. Following up on the "fundholder" experiment, budgets allocated include not only physicians' compensation but prescription and hospitalization budgets.

The PCTs are thus at once responsible for meeting an overall budget, making decisions between different components of that budget, and distributing fees and prescription budgets among their members.

To this end, they have to develop an expertise that allows them to optimize their utilization of resources. For example, in the area of drug prescription, all PCTs monitor their members' prescriptions by means of trend charts provided by the national level, almost all of them have a pharmaceutical consultant, 70% use common prescription reference systems, and 49% have produced their own formulary (Wilkin, Gillam and Coleman, 2001).

This example illustrates that the interest of physicians' groups lies in the ability to link financial accountability to clinical governance in a collective that is:

- managed by professionals, and so has both the technical competence to make the best decisions and a quality code of ethics;
- large enough to acquire a number of expert and assessment tools, but at the same time small enough to make each professional feel involved.

The enthusiasm for physicians' groups, considered to be in the best position to implement efficient management of care processes, and thus to reconcile quality of care with economic use of resources, has also been very strong in U.S. managed care programs. In the most competitive areas, the second half of the nineties saw very fast development of such medical groups and their management and bargaining skills, both with insurers upstream and hospitals downstream.

Within these networks, specific specialty groups can be assigned overall budgets, with the task of distributing them among the individual practitioners they manage. Practitioners can be paid per service, but the group sets up procedures (e.g. peer review, verification of the usefulness of care) which are more operational within a single specialty than in multidisciplinary settings (Robinson 2001). Within this framework, capitation mechanisms called "contact capitation" can be implemented – a fairly innovative approach for some specialists: once a patient is referred by a GP to a group physician, the latter receives a capitation for a period of 12 months, during which time he manages the patient as he sees fit.

This principle of compensation is more adapted to specialties whose activity is relatively homogenous (cardiology, gastro-enterology, etc.). It can be adjusted to account for the gravity of cases, e.g. by balancing patient numbers for certain diagnoses or services, or by allowing FFS for certain procedures (Franck 1999). This payment method has been used in numerous HMOs, and managed care specialists recently predicted that it has a promising future (Carlson 1997). The Swedish experience seems to confirm that group incentives can have a strong impact, even if individual income is not affected by the change in mode of payment: some hospitals moved to one payment per pathology, while others stayed with the fixed annual envelope, but in both cases the physicians were salaried hospital staff. The result is that hospital physicians paid per pathology significantly reduced lengths of stay, even though total cost in these hospitals did not decrease (Forsberg, Axelsson and Arnetz 2001).

The experiences described above entail very formalized networks with substantial delegation of responsibility. But the idea of promoting a more collective practice of medicine in the form of networks or groups is also developing in other systems, in a less structured fashion (France, Germany, some Canadian provinces with experimentations of family physician groups). Here the objective is rather to improve the coordination and continuity of care, something that an isolated

FFS exercise does not spontaneously encourage. But the underlying idea is the same – that between the isolated physician and the profession as a whole, the locally established group of physicians, sharing skills and a common culture, can be a tool for changing practices in the desired direction.

However, recent developments in the United States seem to reflect upon the excesses of this trend. California, where the medical group reigns supreme, has experienced many bankruptcies in recent years, and is now seeing its groups unravel, with individual contracting picking up momentum. The decline of the groups is happening in tandem with the decline of capitation as a method of payment (Managed Care Outlook 2001): the percentage of groups with at least one capitation contract, which rose from 1998 to 1999 (66% to 72%), fell in 2000 (55%).

The factors behind the fragility of the groups and their lack of staying power (diseconomies of scale, difficulty of incorporating insurance perspectives, demand for free choice of patients, etc.) aptly illustrate that the reconciliation of cost containment and quality objectives through the virtue of collective "clinical governance" has its limits, and that the context of extreme financial accountability is perhaps not the most conducive to making the best use of these collective organizations which, in spite of everything, seem to be the way forward in many systems.

Non-financial incentives

If physicians' groups can coordinate cost containment and quality of care, it does not simply rests on the alchemy of group and individual financial incentives: it is because the latter are associated with other incentive mechanisms.

These are of several types. The first is the **development and distribution of clinical protocols and recommendations,** found today in most public or private systems. These can be national or local, and may provide guidance or be mandatory (e.g. medication formularies). There is abundant literature on the comparative efficacy of different methods of disseminating references (training, interviews, computerized reminders) for their implementation by physicians, following the recommendations of the Cochrane Collaboration on Effective Professional Practice. It is beyond our purview here to reproduce and comment on this literature, but we can highlight its main lessons, as summarized by Durieux (1999): no studies have found that distribution of standards via print material or ongoing medical training has a significant impact; however, distribution by opinion leaders or, better yet, by visits from colleagues, does have a significant impact on physicians' practices.

Consistent with this finding, the **evaluation of individual practice**, particularly in terms of these reference sources, or compared with colleagues (release of information to the physician, peer review, etc.) is a tool for changing individual behaviours.

The care network or physicians' group is a preferred forum for applying these incentives, for the reasons given above. This is especially true since these networks select their members based on their support for the group's values – clinical, economic and cultural. A further step can be taken in multiple-insurer systems with **selective contracting** (where the payer is able to manage only a limited number of professionals), which is obviously a powerful non-financial incentive.

Conclusion: What are the Lessons for the Canadian Context?

Canada ranked fifth among OECD countries in 1998 for health spending as a percentage of GDP, and the interim report of the Commission on the Future of Health Care in Canada notes that the problem lies more with the unstable evolution of these expenditures than with their absolute level (OECD 2001; Commission 2002).

Also in 1998, per capita ambulatory care expenditures put Canada in seventh place, with a level ($PPP 333) representing a third of that in the United States and 80% of the German level, but 40% higher than the French level. The price trend for physician services in Canada over the 1985-1996 decade matches that of other OECD countries (excluding the United States), and the increase in volumes, which was quite high at the outset compared to other countries, seems to have been curbed from 1995 onward.

So Canada does not appear to have a serious problem of controlling health care costs.[6] However, the Commission's interim report underscores Canadians' great dissatisfaction with the accessibility and quality of their health care system.

To resolve these problems, must fee-for-service necessarily be called into question? We have seen that payment by capitation can pose a risk for quality of care and accessibility, particularly since the physician need not fear competition, should there be a shortage of doctors, for example. On the other hand, by modulating the physician's compensation according to the type of population served, capitation can be a means of addressing specific problems such as the uneven geographic distribution of physicians.

The current debate on the future of the Canadian health care system certainly provides an opportunity for discussing payment methods. Especially since the latest survey by the Canadian Medical Association reveals that, although 62% of Canadian doctors derive almost all of their income from fee-for-service, only 37% prefer that method of payment to other alternatives (Martin 2000).

Specific questions

What is the impact of physician payment mechanisms on both utilization of physician services, and overall health care utilization and costs?

How does fee-for-service payment for physicians influence utilization and costs compared to other mechanisms, and where are the observed differences?

There is no real alternative to the canonical payment systems of fee-for-service, capitation and salary. Recent innovations are more in the nature of marginal adaptations. Current knowledge of the impact of payment methods on health care costs is summarized, according to their degree of inclusiveness: fee-for-service allows the physician to increase the volume of medical services provided to the patient, and so limits the scope of modes of regulating by unit prices.

Conversely, lump-sum forms of payment (capitation and salary) limit volumes, but can adversely affect the quality of care provided, limit access to care for vulnerable populations and have unexpected consequences on the level of spending. The extent of the impact on quality and access is critically dependent on the patient's ability to observe the physician's effort, an ability upon which it is difficult to make any empirical statement. The unexpected effect on costs comes in the form of prescriptions (a lump sum method of payment encourages the primary care physician to refer and prescribe more) and long-term effects (if the physician reduces service in the short term, the system can pay for the consequences of this in the long term).

The concrete responses are designed to combine mechanisms so that weaknesses offset each other. Payments with a health outcome (intermediate objectives) serve to increase the capacity of payment by capitation to account for the diverse nature of the physician's work, and to combine public health objectives on the one hand with access and quality objectives on the other. Delegation of responsibility to the medical profession (the regulator sets an envelope which is then managed in a decentralized fashion) can come up against opportunistic behaviours if the decentralization is too weak, and conflicts of interests if it is too strong (physician groups that are too small). After being described as the solution of the future, delegation to physician groups now seems to be marking time, to the benefit of non-financial incentives focussed on patient and doctor information.

What is the evidence on the impact of changes in payment mechanisms or incentives on utilization and costs? Where have the most dramatic changes taken place, and how?

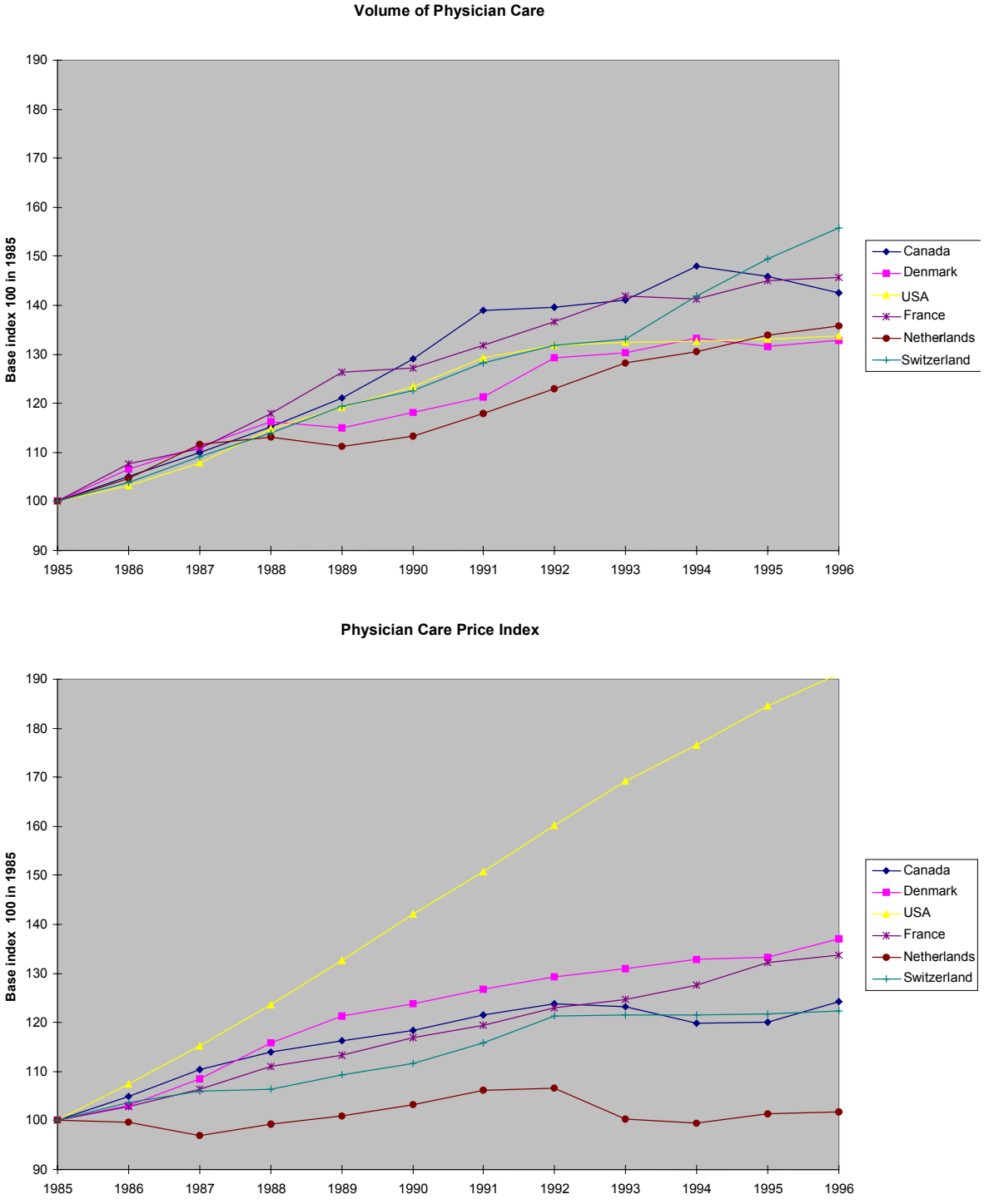
Quantitative data are rare, opportunistic in nature (coming from experiences carried out for other purposes) and always biased. Noteworthy exogenous changes have been made in the United Kingdom, with the introduction of fundholding and PCTs, in Denmark (Copenhagen experiment) and in Germany (envelope and floating points). In the United States, there are numerous reduced-sample experiments (a clinic or a plan) which are difficult to generalize. The reform of Medicare rates has also been a learning ground, but the main problem is that it mixes level and method of remuneration.

Empirical quantitative data suggest that the transition from capitation to FFS increases volumes and reduces prescriptions, although there is no real proof of a reciprocal effect. The quality assessment is more ambiguous, even though patients seem dissatisfied with the lump sum payment method. The target payment experiments seem moderately conclusive in terms of the impact of incentives on the desired outcome. In addition to quantitative evaluations, we have qualitative assessments centered on stakeholders' opinions. Tracking the evolution of physicians' groups in the United States and their internal arrangements is yielding numerous results, which however are difficult to generalize to single-payer contexts.

What non-monetary incentives affect utilization and costs? How do they operate?

Two main innovations are suggested: To assist physicians by providing them with validated and synthesized information about clinical protocols and efficiency; in particular, replacing promotion from the medical products industry; and help physicians and patients get some idea of variations in practices. This second type of incentives is aimed at strengthening patients' control of the physician's effort, which can either increase variability (adapting practices to demand) or decrease it (standardization). This is not simply a matter of finger pointing: physicians may also find it advantageous to compare their practice to that of their colleagues. The impact of these non-monetary incentives is strongly context-dependent, particularly the level of organization of the profession and the ability of physicians to show respect for standards of good practice among their peers.

Figure 1:
Evolution of Prices and Volumes* of Physician Services, 1985-1996



* Volumes correspond to ambulatory care expenditures deflated by the corresponding price index.
Source: *OECD Health Data* 2001.

Notes

1 Increasing quantity can mean changing the structure of services offered by increasing the proportion of better-paid services. Taking the example of Quebec between 1977 and 1983, Rochaix (1993) shows that freezing fees has led to a substantial increase in the proportion of elaborate consultations.

2 A salaried physician could be sanctioned by his employer. With capitation, he will be sanctioned by the patient if inferior quality is noticed and if professionals are in competition to attract clients.

3 Referral is the action whereby one professional refers a patient to another professional.

4 These data allow one to control bias in comparative studies by capturing the effects linked to each individual; however because the data are complicated to obtain, samples are less often representative.

5 Another study on the same set of data appears to show a strong and significant impact of fee-for-service on the volume of urine tests (Kristiansen and Hjortdahl 1992). But Scott and Hall (1995) point out that this effect is measured without correcting for the bias related to the non-independence of the observations (many contacts per patient are observed).

6 This verdict is based on data available at the time this article was written (in 1996 or 1998, depending on the series). Canadian experts might possibly come to another conclusion with more recent data on trends in expenditures, prices and volumes generated.

References

Armour, Brian S., et al. 2001. The effect of explicit financial incentives on physician behaviour. *Archives of Internal Medicine*, 161(10): 1261-2166.

Bocognano, Agnès, et al. *Mise en concurrence des assurances dans le domaine de la santé : théorie et bilan des expériences étrangères***.** CREDES-INSEE, 1998.

Brunelle, Yvon, and Howard Bergman. 1999. Capitation, efficience et rationnement. *Ruptures, revue transdisciplinaire en santé*, 6(2): 270-288.

Busse, Rheinardt. Health [sic-TR] Systems in Transition: Germany, European Observatory on Health Care Systems, 2000.

Carlson, Bob. "Capitation: State of the art". 1997. Available httpwww.managedcaremag.com/archives.

Commission on the Future of Health Care in Canada. Shape the Future of Health Care – Interim Report, February 2002.

Coulter, Angela. 1995. Evaluating general practice fundholding in the United-Kingdom. European Journal of Public Health*, 5: 233-239.*

Davidson, Stephen, M. Manheim, S. Werner, et al. Prepayment with office-based physicians in publicly funded programs: results from the children's Medicaid program. *Pediatrics*, 89: 761-67.

Debrock, Larry, and Richard, J. Arnould. 1992. Utilization control in HMOs. *Quarterly Review of Economic Finance*, 32: 31-53.

Durieux, Pierre. Actions de diffusion, d'information et de formation médicale continue, in *Comment influer sur les pratiques professionnelles des médecins?* Institut d'Études des Politiques de Santé, 1999.

Emery, J.C. Herbert, Chris Auld, and Mingshan Lu. *Paying for physician services in Canada: the institutional, historical and policy context*, Working paper 99-6. Institute of Health Economics, 1999.

Escarce, José. 1996. Using physician practice cost functions in payment policy: the problem of endogeneity bias, *Inquiry,* 33: 66-78.

Epp, Michael, et al. 2000. The impact of direct and extra billing for medical services: evidence from a natural experiment in British Columbia, *Social Science and Medicine,* 51: 691-702.

Evans, Robert G. 1974. Supplier induced demand : some empirical evidence and implications, in M. Perlman, ed., *The Economics of Health and Medical Care.* McMillan, pp. 162-73.

Evans, Robert G. 1998. New bottles, old wine: right and wrong in physician supply, *Canadian Medical Association Journal*, 158: 757-9.

Ferrall, Christopher, Allan W. Gregory and William G. Tholl. 1998. Endogenous work hours and practice patterns of Canadian physicians, *Canadian Journal of Economics.*

Forsberg, Ewa, Runo Axelsson, and Bengt Arnetz. 2001. Financial incentives in health care. The impact of performance-based reimbursement. *Health Policy*, 58(3): 243-262.

Franck, Clifford R. 1999. Can the pie be divided more fairly? *Contact capitation for specialists*, 5(2). Available http://www.cost-quality.com/5,2art3.html.

Gerdtham, Ulf-G, Jes Sogaard, Fredrik Andersson and Bengt Jönsson. 1992. Econometric analysis of health expenditure: a cross-section study of the OECD countries. *Journal of Health Economics*, 11: 63-84.

Gerdtham, Ulf-G, Bengt Jönsson, M. MacFarlan and H. Oxley. 1998. The determinants of health expenditure in OECD countries, in *Health, the Medical Profession, and Regulation*, under the direction of Peter Zweifel, Kluwer.

Gillett, James, Brian Hutchison, and Stephen Birch. 2001. Capitation and primary care in Canada: financial incentives and the evolution of health service organizations. *International Journal of Health Services,* 31(3): 583-603.

Glied, Sherry. 1998. Payment heterogeneity, physician practice, and access to care. *American Economic Review*, 88(2): 127-31.

Gosden, Toby, Frode Forland, Ivar Sombo Kristiansen, et al. 2001. Impact of payment method on behaviour of primary care physicians: a systematic review. *Journal of Health Services Research and Policy*, 10(1): 44-55.

Grumbach, Kevin, Dennis Osmond, Karen Vranizan, et al. 1998. Primary care physicians' experience of financial incentives in managed care systems. *New England Journal of Medicine*, 339: 1516-21.

Grytten, Jostein and Rune Sorensen. 2001. Type of contract and supplier-induced demand for primary physicians in Norway, *Journal of Health Economics*, 20: 379-393.

Hall, Mark, A., Elizabeth Dugan, Rajesh Balkrishnan, et al. 2002. How disclosing HMO physician incentives affects trust. *Health Affairs*, 21(2): 19-27.

Hanchak, NA. 1997. A performance-based compensation model for obstetricians/gynecologists. *Clinical Obstetrics and Gynecology*, 40: 437-45.

HCFA, Study on Pharmaceutical Benefit Management, PricewaterhouseCoopers LLP.

Health Canada. 2001. *Canada Health Act*, *2000-200*, [*sic*-TR] Annual Report.

Hemenway, David. 1995. Financial incentives for childhood immunization, *Journal of Policy Analysis and Management,* 14: 133-139.

Hickson, B., A. Altemeier, and M. Perrin. 1987. Physician reimbursement by salary of fee-for service: effect on physician practice behaviour in a randomized prospective study. *Pediatrics*, 80: 344-50.

Hillman, Alan L., Kimberly Ripley, Neil Goldfard, et al. 1998. Physician financial incentive and feedback: failure to increase cancer screening in Medicaid managed care. *American Journal of Public Health*, 88: 1699-1701.

Hutchinson, James M. and Robert N. Foley. 1999. Method of physician remuneration and rates of antibiotic prescription. *Canadian Medical Association Journal*, 160: 1013-117.

Jelovac, Izabela. 2001. Physicians' payment contracts, treatment decisions and diagnosis accuracy. *Health Economics*, 10: 9-25.

Josseran, L., G. King, and J. Chaperon. 2000. Les Primary Care physicians américains. *Journal d'économie médicale*, 18(6): 375-387.

Kouides, R.W., N.M. Benett, B. Lewis, J.D. Cappuccio, W.H. Barker, and F.M. LaForce. 1998. Performance-based physician reimbursement and influenza immunization rates in the elderly. *American Journal of Preventive Medicine*, 14: 89-95.

Kralewski, John E., et al. 2000. The effects of medical group practice and physician payment methods on costs of care. *Health Services Research*, 35(3): 591-613.

Krasnik, Allan, P. Groenwegen, Poul A. Pedersen. 1990. Changing remuneration system: effects on activity in general practice. *British Medical Journal*, 300: 1698-1701.

Kristiansen, Iwar S., and K. Holthedahl. 1993. The effect of the remuneration system on the general practitionner's choice between surgery consultations and home visits. *Journal of Epidemiology and Community Health,* 47: 481-4.

Kristiansen, Iwar S., and Gail Mooney. 1993. The general practitioner's use of time: is it influenced by the remuneration system? *Social Science and Medicine*, 37: 393-9.

Kristiansen, Iwar S., and Per Hjortdahl. 1992. The general practitioner and laboratory utilization: why does it vary? *Family Practice,* 9: 22-7.

Managed Care Outlook. 2001. "Interest in capitation back off". Available http://www.managedcaremag.com/archives

Mc Guire, Thomas G. 2000. Physician agency in Handbook of Health Economics, *under the direction of Anthony J. Culyer and Joseph P. Newhouse, 1: 463-536.*

Martin, Shelley. 2000. CMA survey shows fee-for-service not dead yet. Canadian Medical Association Journal*, 163(5): 601.*

Miller, Robert H. and Harold, S. Luft. 1997. Does managed care lead to better or worse quality of care? *Health Affairs*, 16(5): 7-25.

Mulligan, Pamela. 2002. La capitation: une direction mal avisée dans la réforme des soins de première ligne. Le Médecin de famille canadien*, 48: 244-246.*

Namovicz-Peat, S., J. Swann and K. Tureski. HMO & PBM strategies for pharmacy benefits: recent results, current practices, future plans. AIS - Atlantic Information Services, 2001.

Philips, Karen, and William R. Swan. 1996. Health care system in transition: Canada. Preliminary version. World Health Organization - Regional Office for Europe – Copenhagen.

Red Book 2002. Available http://www.nhs.uk/redbook/2.htm.

Rice, Thomas. 1993a. A model is only as good as its assumptions: a reply to Peele. *Journal of Health Economics*, 12: 209-211.

Rice, Thomas. 1993b. Demand curves, economists, and desert islands: a response to Feldman and Dowd. *Journal of Health Economics*, 12: 201-204.

Robinson, James C. 2001. Theory and practice in the design of physician payment incentives. *The Milbank Quarterly*, 79(2): 148-177.

Robinson, James C. 2001. Physician organization in California: crisis and opportunity. *Health Affairs*, 20(4): 81-96.

Rochaix, Lise. 1993. Financial incentives for physicians: the Quebec experience. *Health Economics*, 2: 163-176.

Rosser, Walter, and Jan Kasperski. 2002. En faveur d'une rémunération mixte. *Le Médecin de famille canadien*, 48: 247-249.

Sandier, Simone. 1990. Le paiement des médecins dans quelques pays de l'OCDE, in *Les systèmes de santé, À la recherche d'efficacité*, Paris: OECD, 1990.

Scott, Anthony, and Jane Hall. 1995. Evaluating the effects of GP remuneration: problems and prospects. *Health Policy*, 31: 183-195.

Scott, Anthony. 2000. Economics of general practice, in *Handbook of Health Economics,* under the direction of Anthony J. Culyer and Joseph P. Newhouse, 1: 1175-1200.

Showstack, Jonathan A., Deborah W. Garnick, Kenneth E. Rosenfeld, et al. 1987. Episode-of-care physician payment: A study of coronary artery bypass graft surgery. *Inquiry*, 24: 376-383.

Stearns, Sally C., Barbara L. Wolfe and David A. Kindig. 1992. Physician responses to fee-for-service and capitation payment, *Inquiry*, 29: 416-429.

Tai-Seale, Ming, Thomas H. Rice, and Sally C. Stearns. 1998. Volume responses to Medicare payment reductions with multiple payers: a test of the McGuire-Pauly model. *Health Economics*, 7: 199-219.

Whynes, David and Darrin L. Baines. 1998. Income-based incentives in UK general practice. *Health Policy*, 43: 15-31.

Wilkin, David, Steve Gillam, and Anna Coleman. The national tracker survey of primary care groups and trusts, 2000/2001.