



Program Evaluation and Economic
Consulting
1780 Vernon Street, Halifax, NS

Quality Based Incentive Payments: International Experiences, Lessons for Canada

Dominika Wranik, Ph.D.
Martine Durier-Copp, Ph.D.

**Report Submitted to
The Government of Manitoba
Minister of Health**

Table of Contents

1. Introduction and Context	3
1.1. Increase in chronic disease	3
1.2. Realization of inappropriate care use	4
1.3. Shift in focus toward value for money	4
1.4. Supporting policy documents	6
2. Theory of physician reimbursement	8
2.1. Classification of payment methods	8
2.2. Physician level payments	10
2.3. Institution level payments	15
3. Quality Based Incentive Payments	17
3.1. Goals	17
3.2. Types of incentives	18
3.3. Quality measurement	21
3.4. Examples	25
4. Experiences to date	30
4.1. Theorized effects	30
4.2. Empirical findings	32
5. Lessons Learned and Recommendations	34
5.1. Quality measures	34
5.2. Data collection	35
5.3. Program structure	35
5.4. Program implementation	36
6. Conclusions	37

1. Introduction and Context

Methods of physician remuneration are considered an essential component of effective health care policy and management. As such, they have been the subject of debate among economists, clinicians, and governments for nearly four decades. There has been a great deal of experimentation in physician payment systems over the past two decades in particular as governments grapple with increasing costs for services and mounting demand. Economic theories serve to provide the foundation for the kind of incentive structures created by various physician remuneration methods. Based on these theories, governments can design the type of incentive structure which can best respond to the most pressing concerns of the health care system and obtain the desired behaviours on the part of clinicians.

In general terms, four major models of physician remuneration exist, namely, payments based on items of service, payments based on patient numbers, payments based on time, and payments based on quality of care or performance. The latter is subject to investigation in this report, and is termed Quality Based Incentive Payment (QBIP) for this purpose.

Models of QBIP are relatively new in the health care sector, although payment for performance has long been advocated in management theory and in many other industries, where employers cannot easily observe effort. Modern public administration has seen new emphasis placed on “results”, effectiveness, and value for money, and this trend has been reflected in the management of public sector domains, such as health and education. In the health care system, quality based management frameworks have principally been observed in three jurisdictions, notably the UK, Australia and the US, which have served as models for other countries. QBIP models are used in the United Kingdom, under the name Performance Based Reimbursement (PBR), in the United States, under the name Pay for Performance (P4P), and in Australia, under the name Practice Incentive Program (PIP). In all jurisdictions, similar developments led to an increase in popularity of QBIP programs: increases in chronic disease prevalence and a growing aging population, evidence of care being used inappropriately, and the shift in policy focus toward value for money considerations. These three factors led to a reassessment of health care management by governments, including the incentive structure. A series of measures were introduced to better address policy concerns.

1.1. Increase in chronic disease

The World Health Organization declared chronic diseases an epidemic with large economic impact in terms of health expenditures and forgone production.¹ At the global scale, the goal is to reduce the projected trend in chronic disease death rates by 2% each year, until 2015.

In the U.S., 60% of patients report chronic health problems, which amounts to over 59 million individuals. The leading chronic diseases are diabetes, mental health issues, arthritis, and asthma. Over a quarter of these individuals experience three or more chronic diseases simultaneously.

Chronic disease is a large burden in Canada, as well. The Canadian Diabetes Association reports that over two million Canadians have diabetes.² One hundred and fifty three thousand Canadians were diagnosed with cancer in 2006.³ 8.1% of the Canadian population have been diagnosed as having asthma in 2001.⁴ Estimates of prevalence of chronic disease in Manitoba indicate that 20-30% of the population 19 years of age and older suffers from arthritis, 11-17% suffer from asthma, 6-8% suffer from Coronary Heart Disease, 6-8% suffer from Diabetes, 22-23% suffer from hypertension, and 3-4% suffer from stroke. (Lix et al., 2006)

¹ WHO: Preventing Chronic Disease: A Vital Investment

² <http://www.diabetes.ca>

³ http://www.cancer.ca/vgn/images/portal/cit_86751114/31/21/935505792cw_2006stats_en.pdf

⁴ <http://www.asthma.ca/corp/newsroom/pdf/asthmastats.pdf>

1.2. Realization of inappropriate care use

In 1999, IOM published a report, providing evidence that medical errors were a leading cause of death in the United States, ahead of breast cancer, AIDS, and motor vehicle accidents. (Shea, 2006).

In the late 1990s it was recognized that quality of care was being compromised through underuse, overuse, or misuse of health care (Chassin *et al.*, 1998). Underuse is to mean a situation, when services are not provided, despite having potential positive outcomes for the patient. For example, hypertension is not treated in roughly 50% of patients, mammograms are not given to 65% of women under FFS, and 45% under managed care, and depression is not treated in 46% and 58% of patients under FFS and managed care, respectively. Overuse refers to a situation, when services are provided, despite their potential for harm being greater than their potential for benefit. For example, it is estimated that in 1992 21% of antibiotic prescriptions are given inappropriately, 17% of coronary angioplasties were inappropriate, and 32% of carotid endarterectomies were inappropriate. Misuse refers to a patient receiving appropriate services, then experiencing preventable complications. (Chassin *et al.*, 1998) Each one of these is motivated by a different reimbursement method – capitation and salaries support underuse, fee for service supports overuse, and no method focuses on diverting misuse. P4P methods are aimed to mitigate the occurrence of all three situations. (Baker *et al.*, 2003).

Establishing programs linking pay with performance and results was sparked by the realization that none of the other payment methods (FFS, salary, capitation, blends, and all other modalities described earlier) gave incentives to provide high quality of care, and that care quality is what matters most to patients and their well being. The problems with overuse, underuse, and misuse of health care partially fuelled by the physician being faced with a multitasking work requirement. There is a multitude of responsibilities and tasks a physician must perform, and when time and resources are limited, economic theory predicts that she will focus on tasks that are rewarded explicitly. Payments targeted at previously not remunerated tasks represents an attempt to protect the priority of those tasks (Rosenthal *et al.*, 2004). Discussions of the National Roundtable on Health Care Quality in the U.S. resulted in three main conclusions. First, quality of health care can be precisely defined and measured; second, problems with health care quality are extensive both in terms of delivery systems and financing mechanisms; and third, selected health plans, hospitals, integrated delivery systems, have made efforts to improve care quality (Chassin *et al.*, 1998). Recently, the Medicare Payment Advisory Commission (MedPAC) recommended that Congress add incentives for quality into payment systems to hospitals, physicians, home health agencies, facilities treating dialysis patients. These are referred to as Medicare Advantage Plans.

1.3. Shift in focus toward value for money

Many jurisdictions, including Australia, Canada, the United Kingdom, and the United States, have been struggling with rising costs of health care over the last three decades. In absolute terms, when measured in \$US per capita, health care costs have increased in all four countries, from roughly \$500 per capita in the mid 70s, to roughly \$2500 in 2002, with the U.S. facing health expenditures in excess of \$5500 per capita in 2002 (a ten fold increase over 25 years). When health care costs are considered as a percentage of GDP, these have risen less dramatically in most countries, with the U.S. again being an outlier. Canadian health care costs as a percentage of GDP have increased from 7% in the 70s, to around 8% in the 80s, and around 9% in the 90s. Costs in the U.K. have been lower through all years, with 5% in the 70s, 6% in the 80s and 7% in the 90s. In the U.S., health care costs as a percentage of GDP have increased from around 7% in the 70s, to 15% in 2002. (OECD Health Data 2005). In response to these trends and growing demand for health care services, Canada, Australia, the U.K. and the U.S. have responded with cost control measures, including experiments with physician remuneration.

Health care costs can be contained from either the demand side, by creating incentives for patients to reduce consumption, from the supply side, by creating incentives for providers to reduce production, or via regulation. (Cutler, 2002)

Demand controls are not the focus of this report, hence their description is brief. Patients can be motivated to consume less health care through various cost sharing arrangements, all of which can conflict with the principle of equality of access. Cost sharing arrangements are minimal in countries like Canada, Italy, Germany or the U.K., and higher in France, Japan, or the U.S. (Cutler, 2002) Demand side incentives do not get the same level of attention in the literature as the supply side, likely because they are easier to understand, and theoretical predictions about their effects are accepted. There is little controversy that higher co-payments deter patients from seeking care, dispute lies with ethical arguments surrounding co-payments, and the effects of co-payment on socio-economic groups and on health outcomes. Since need for care is not positively related to ability to pay, countries can choose to provide care based on need or based on ability to pay. The question is largely ideological, the theoretical and empirical evidence on the structure of incentives is agreed upon.

Supply side incentives are more controversial in the literature for several reasons. First is the question of whether physicians are motivated by financial incentives at all. The proposition appears obvious to some, and preposterous to others. Second, there is a discussion and debate about which financial arrangements create what kinds of incentives, and what outcomes are desirable. For example, cost containment is supported by a different set of provider incentives than is care quality, physician productivity, or responsiveness to community needs. No agreement exists as to which of the conflicting goals should be focused upon, nor upon the effects of various financial incentives on providers. This report discusses the incentives created, their effects on health care delivery, and in particular focuses on payment methods in support of health care quality.

Health systems have undergone reform in the name of cost containment in most of the Western world since the late 70s. Beginning in the late 70s and early 80s, global spending limits were introduced in Canada, U.K., France, Germany, and Japan (Cutler, 2002). Public health systems could limit expenditures via regulation or policy measures to reallocate or limit certain types of expenditures, while predominantly private systems like the U.S. had to rely more heavily on managing health care at the micro level, including physician remuneration (Kesteloot, 1999).

Facing the greatest increases in health care costs, the U.S. has experimented heavily with various provider reimbursement methods. The U.K. has also been a world leader in the introduction of alternative methods of physician remuneration. Trends in U.S. and U.K. physician payment policies can be seen as precursors to physician payment reforms in other Western economies. Historical developments in physician remuneration have been categorized into three chronological phases, spearheaded by the U.S. and often followed by other jurisdictions. During the early 1990s, concern with rising health care costs sparked the rapid spreading of managed care organizations in the United States (Korn, 2004). Managed care was characterized by capitation payments of physicians, which were also introduced in other Western health systems around the same time, such as for instance the development of fundholding in the U.K. Health Maintenance Organizations (HMOs), global capitation, and selective provider contracting encouraged the development of integrated care, characterized by hospital ownership over primary care (PC) practices, Physician Hospital Organizations (PHOs), hospital sponsored management service organizations (MSOs), and hospital sponsored IPAs. As a result of capitation payment, PC providers were needed by hospitals to recruit new patients for practice (Lake, 2003).

Canadian cost control efforts in the first half of the 1990s took place at both the macro level and micro levels. At the macro level, spending caps on physician fees were implemented in many provinces. Some of these were “hard” in the sense that any overruns had to be repaid by the profession in full, and others were “soft”, whereby the profession shared the responsibility for overruns with Ministries. In addition, the supply of physicians was restricted via adjustments to medical school enrolments, and restrictive policies on foreign medical graduates. At the micro level, physician incomes were controlled through billing number restrictions, maximum amounts to be billed, introduction of salaries and capitation, and monitoring efforts (Barer *et al.*, 1996).

The second phase was characterized by continued support for managed care, as evidenced for instance by the Medicare Prescription Drug, Improvement, and Modernization Act (2003) assigning subsidies to Managed Care organizations. The enthusiasm for capitation, integration, and selective networks, however, began to decline by the late 1990s (Korn, 2004). In the U.K., fundholding was abolished, as well.

It is reported that managed care has failed to achieve its original growth expectations (Lake, 2003). The introduction of managed care in the U.S. was not accompanied by a decline in health care costs, nor by slower growth rates. Instead, during the 1990s, health care costs in the U.S. increased from \$2700 per capita, to \$5600 per capita, and from 11% to 15% of GDP (OECD Health Data 2005). Today's health policy makers are preoccupied with stimulating physician performance, and promoting quality of health care – a need that was identified almost a decade ago (Chassin *et al.*, 1998). In economic terms, this represents a shift in emphasis from the concept of cost control alone, toward inclusion of effectiveness, which examines results achieved. The achievement of maximum effectiveness at minimum cost is known as efficiency. At the level of institutions, such as hospitals, and physician organizations in the U.S., performance based reimbursement is becoming increasingly more popular. At the level of the individual physician in a small practice setting, it is more difficult to base remuneration on performance, due to data and measurement difficulties (Cunningham, 2004, Korn, 2004, Super, 2006). Nonetheless, “the new mantra of physician remuneration is pay for performance in small practice settings” – an approach, which supports the notion that doctors have the primary responsibility for decisions regarding health care (Korn, 2004). The trend is accommodated by the formation of new arrangements between Integrated Delivery Systems (IDS) and physicians, where emphasis is shifting away from owning or contracting with PCPs, and toward establishing solid relationships with specialists (Lake, 2003).

To sum up, each decade was characterized by changes in remuneration structures. During the 1980s, focus was given to creation of diagnostic related groups (DRGs) and tying of physician payments to them. During the 1990s, capitation was heavily experimented with, as were alternative funding plans. This decade is characterized by focus on quality, and tying of physician payment to quality and performance.

1.4. Supporting policy documents

In the late 1990s, the National Roundtable on Health Care Quality in the U.S. launched a discussion on the urgent need to improve health care quality.

In 2001, the Institute of Medicine (IOM) in the U.S. published a report titled “Crossing the Quality Chasm: A New Health System for the 21st century” presented a broad strategy to increase safety and accountability in health care.

A decade after the National Roundtable on Health Care Quality discussions, the rationale for these recommendations is similar to issues raised by the Roundtable. First, evidence exists of unsafe care being provided, and the quality of care is low for many beneficiaries. Second,

resource-based payment systems are essentially fee for service in terms of incentive structure - quantity is rewarded. Third, the focus at system level shifted from cost reduction to low cost and good value for money, P4P programs support this shift at the micro-level (MedPac, 2003, Milgate *et al.*, 2006).

2005 Deficit Reduction Act in the United States directs the secretary of Health and Human Services to develop a plan for value based hospital payments in Medicare, to be implemented in 2009. As a result, Centers for Medicare and Medicaid Services (CMS) are working on an across the comprehensive strategy to adapt P4P programs for health plans, physicians, and institutions.

PricewaterhouseCoopers has identified Pay for Performance as a key trend in the health care market between 2002 and 2007. The trend is described as a grassroots movement, spearheaded by individual payers, who question a structure, where worst and best performers receive the same payments. (Shea, 2006)

The Department of Health in the United Kingdom has published several policy documents in support of its new Payment by Results system.

In 2002, the Department of Health in the U.K. published report on a similar subject, titled "Reforming NHS Financial Flows: Introducing Payment by Results". This document describes how funds are to flow through the NHS. (U.K. DOH, 2002) This publication is supported by several more specific documents, such as for instance "Care Planning in Diabetes", a report from the Joint Department of Health and Diabetes Care. This document sets out the policy context for care planning, including principles and processes, use of evidence, workforce issues, and quality assurance processes.

In the U.K., the Payment by Results experiments have been introduced through the National Health System's creation of the Foundation Trusts. The experiments started gradually in 2003 and early results have been interesting, though the NHS cautions that much remains to be enacted and more time is needed before one can make any realistic assessment of the PbR methods.⁵

"Standards for Better Health", as introduced by the U.K. Department of Health in April 2005, set out areas to be included in quality measurement and the quality incentives package. The areas include safety, clinical and cost effectiveness, governance, patient focus, accessible and responsive care, care environment and amenities, and public health. (DoH, 2005)

The Health and Social Care Standards and Planning Network for 2005-2008 lays out principles to be followed in planning, commissioning and providing services. With respect to remuneration of physicians, it is suggested that incentives be designed to support increased quality of patient care, and to provide better scope for physicians to address local priorities. National priority areas, also to be supported, include health and well being of the population, care for long term conditions, access to services, and good patient experiences. The new system sets as a target the achievement of efficiency savings of 2.7% per year for the 2005-2008 period. (DoH, 2005)

2. Theory of physician reimbursement

⁵ For more, see

http://www.dh.gov.uk/Consultations/ClosedConsultations/ClosedConsultationsArticle/fs/en?CONTENT_ID=4069462&chk=nppgJP

This section provides a theoretical framework to analyze remuneration methods, a classification of these methods in health care, a description of their structure, and their theorized effects on various aspects of health care delivery. The framework of agency theory is used, which recognizes that the world operates under conditions of imperfect information, risk and uncertainty. With reference to the health care market, since the payer or principal cannot observe physician behaviour, diligent contract design aims to align physician behaviour with organizational or government goals. Financial rewards are only one, albeit an important one, among a variety of mechanisms for eliciting the desired behavior. Non financial incentives, especially as embedded in QBIP contracts, are also discussed.

Four types of goals in clinical practice need to be accounted for in the process of contract design.

a) Physician productivity and patient services: Physicians should be encouraged to work long hours, perform many procedures and be attentive to the needs and preferences of each individual patient.

b) Risk Acceptance: Physicians should receive extra praise and compensation for treating the sickest patients and should not be rewarded financially for skimming the healthy and avoiding the ill.

c) Efficiency and appropriate scope of practice: Physicians should be rewarded for providing the appropriate level of care in the appropriate setting and for maintaining an appropriate scope of practice.

d) Cooperation and evidence-based medicine: Physicians should be encouraged to cooperate with other clinicians, to adopt evidence-based best practices, and to narrow the unjustifiable variations in how similar patients with similar conditions are treated across the health care delivery system. (Robinson, 2001).

Compensation methods, such as fee for service (FFS), capitation, salary, or blends thereof each support some, but not all of the above goal areas. Piece-rate pay (FFS) is theorized to favor the first two features (with capitation performing rather poorly), whereas capitation is seen as promoting the last two features (and FFS showing limitations in favoring these features). Choice of appropriate remuneration method largely depends on policy goals, population needs, or identified shortcomings in service delivery.

It should also be recognized that an important physician incentive is the organizational structure, in which she/he participates. First, ownership of practice assets creates an extremely high-powered set of performance incentives. Second, large amounts of bureaucracy might create a sense of loss of control over decision making for the physician. Third, particular organizational structures might be more conducive to particular life styles. For instance, the 24 hours/7 days a week commitment that is often implied in capitation schemes may not be suitable for all physicians.

2.1. Classification of Payment Methods – Payments not linked to Quality

Payment methods used to reimburse providers can be categorized along two dimensions: the relationship between payment and activities performed; and, the relationship between payment and resources used (Jegers *et al.*, 2002).

The degree to which payment depends on the level of activities performed by the physician determines the extent to which a payment method is fixed or variable. In a purely fixed system, payment does not depend on the level of activity. The physician is not remunerated

for production of additional units, as the lump sum payment is determined *ex ante* and not related to production. The opposite case is a variable system, wherein payment is linked to the level of activity performed. Incremental income of the physician is equal to the amount received for the production of an additional unit of care. These two pure methods identify the end points of a continuum of possibilities, where a system becomes increasing more fixed the broader is the unit of reimbursement. The unit of reimbursement can be the item of service, the day, the patient, the case, or the period of time. Each type of provider payment, as outlined below, has a fixed and variable component. The more variable a system is, the easier it is for the physician to control his earnings (Jegers *et al.*, 2002).

The distinction between prospective and retrospective methods is based on the linkage between resources used in care provision, and payment received. A retrospective system reimburses the provider's own costs *ex post*. The amount of payment is based on real costs of care production. On the other hand, prospective systems are those wherein provider payment rates or budgets are determined *ex ante*. There is no link between payment nor the individual costs incurred by the provider. In most cases, providers are allowed to keep financial surpluses, and are responsible for financial deficits. In a situation where rates of payment are determined *ex ante*, a large mismatch between payment and costs provides more incentive to under-provide care and decrease care quality. To minimize this risk, rates are often set according to patient and physician characteristics.

Given the framework above, this paper will classify its analysis of the effects of the remuneration methods according to the following categories (BCMA, 2005, Carrin, 2003, Chaix-Couturier *et al.*, 2000, Langenbrunner, 2004, CMA, 2005, Robinson, 2004)

1. Payments at the individual physician level, used to reward PCPs and specialists:
 - Fee for service (RCU, Resource Based Relative Value Scale (RBRVS))
 - Capitation (population based, rostering, fundholding)
 - Time based payments (salaries)
 - Blended payments (B1, B2, and B3, as described later)
 - Subsidies for facilities, equipment, staff
2. Payments at the individual physician level, used mainly to reward specialists:
 - Specialty budget with capitation (contact/referral)
 - Financial support for teaching
 - Funding for research
 - Administrative stipends
3. Payments at the institutional level :
 - Per diem rates (Diagnostic Related Groupings (DRG))
 - Case rates per episode of illness
 - Budgets (line item, global)
4. Performance based incentives, financial and non-financial:
 - The new approach to incentive structuring that is being tried at all levels.
 - Bonuses for selected activities
 - Rewards for achieving targets, withholds for failure to achieve targets
 - Quality grants
 - Additional payment for chronic care, care management, and investment in IT systems.
 - Employer health plan incentive agreements.
 - Patient benefit design and tiered networks.
 - Administrative relief.
 - Public report cards, community recognition, and honour rolls.

- Scoring of providers.

Using Jegers *et al.* (2002) classification, these payment methods can be organized as follows.

	Retrospective	Prospective
Variable	Global budgets based on previous period costs and activities. Bonuses and target payments	Fee for service Capitation Specialty budgets (referral based) Per diem rates Case rates per episode
Fixed	Not possible	Salary Research Funding Financial Support for Teaching Administrative Stipends Quality Grants

If incentives are poorly structured, or are not properly aligned at the individual and institutional level, a multitude of negative consequences could arise. Quality of care might be affected through a decrease in the range of services, the improper use of services, or a decrease in the range of services provided. Physicians might face ethical conflicts: for instance, they might have to choose between more care for one patient, or care for more patients at a lower intensity, or choosing between cost concerns and care provision. Conflicts of interest between patient and provider might also be fuelled. Furthermore, poor incentives might decrease time for teaching and research in academic settings. And finally, if several payers and institutional structures exist, physicians might find themselves with conflicting guidelines from various sources (Chaix-Couturier *et al.*, 2000).

Each payment method is described in terms of its structure, including modalities, its theorized incentives, its use in the world, and its effects as shown empirically. In addition to payment modalities, specialists might also receive other forms of financial incentives, such as shareholding in diagnostic facilities or “kickbacks” from pharmaceutical or medical supplies industry (Chaix-Couturier *et al.*, 2000). Furthermore, specialists also face non-financial restrictions, such as drug formularies, threat of medical malpractice suit, clinical practice guidelines/ protocols, selective contracting, physician supply management, direct controls over infrastructure, technology, quantities of medical personnel, and monitoring of behaviour, among others (Carrin, 2003, Chaix-Couturier *et al.*, 2000, CMA, 2005).

2.2. Physician level payments

Fee for Service

The fee for service system is most common in Canada and internationally. It is a retrospective system that allows physicians to bill for each item of service they have provided to a patient. In most systems, fees are determined *ex ante*, often through negotiations between the payer and the profession. Rarely are fees open-ended, although they are allowed in some instances in the U.S. and Canada. Generally, fee schedules are negotiated, as is the case in Germany, France, Belgium, Canada, and the U.S. Systems like Japan and China rely on a regulatory body to set the fee schedule (Langenbrunner, 2004). The FFS system is often used in social health insurance-based schemes, such as

those in Canada, Australia, the U.K, and in mixed systems, like that of the U.S. (Medicare component) (Carrin, 2003).

The method by which fees are estimated has changed over time in the U.S., and shortly thereafter in other jurisdictions. Traditionally, fees were determined using what is referred to as the usual, customary, reasonable method (UCR) (Mitchell, 1987). Fees were set based on historical values, with little empirical or theoretical underpinnings, and no reflection of production costs. This method was replaced with a more systematic approach around the late 1980s that is still in use today. The current FFS system is estimated using Current Procedural Terminology (CPT), which assigns a unique identifier to each clinical task, and Resource Based Relative Value Scale (RBRVS), which permits the indexing of CPT and conversion into a monetary unit (Robinson, 2001). The RBRVS has been assessed as fair, in that it approximates a competitive market price – it is estimated based on resources used, time required to perform a procedure, type of setting, level of difficulty, and the level of provider training (Dorwart, 1988).

Fee for service as an incentive structure has been extensively evaluated in the theoretical and empirical literature. The following arguments in favour of FFS remuneration have been brought forth in the literature (CMA, 2005, Langenbrunner, 2004, Robinson, 1999)

- Encourages service provision or physician productivity
- Motivates long hours
- Encourages efficient time management
- Motivates performing of difficult procedures
- Motivates caring for the chronically ill
- Cost efficiency of health technologies can be encouraged, if true costs are known
- Reward is fair, if prices reflect relative value
- Detailed record keeping provides a rich data source
- Familiarity with the system by payers and providers yields high acceptance

High physician productivity can be viewed positively, or negatively, depending on the necessity of services provided. The point is, FFS does not provide any incentive to restrict services for any reason, including increased risk, which is associated with difficult procedures, and caring for the chronically ill. In a FFS environment, the payer assumes all risk resulting from uncertainty in health care. If in any given time period, the number of severely ill patients rises and more costly procedures are performed than anticipated, the payer assumes the financial consequences, not the provider.

Many shortcomings of the FFS method have also been cited in literature (Carrin, 2003, CMA, 2005, Dorwart, 1988, Langenbrunner, 2004, Robinson, 1999)

- Requires detailed record keeping of procedures performed, and is disliked by payers due to high administrative costs, and difficulties in budgeting.
- FFS can result in over-provision of care, provision of inappropriate services, or supplier induced demand
- Physicians have no incentive to perform un-remunerated tasks, such as patient education, counseling, returning patient phone calls or correspondence.
- Motivates fraudulent upcoding of visits and procedures.
- Motivates “ping pong” referrals between specialists.
- Is not fair, when fees are not reflective of relative values.
- Rising health care costs, supported by FFS
- Fragmentation and failures of coordination in health care, supported by FFS

In addition, it is suggested that FFS performs well when desired behaviour is standardized and easily measured, but does not work well when multiple tasks are required, some of which cannot be measured, and not when team cooperation is important (Robinson, 1999).

In general, empirical literature confirms that FFS is associated with a higher quantity of services provided than alternative payment methods. In a systematic review of the literature, Chaix-Couturier *et al.* (2000) demonstrate that FFS does have demand-inducing effects in certain situations. For instance, one study showed that a ceiling on fees resulted in increased quantity of services provided. Another study showed that gynecologists are more likely to provide elective procedures under FFS than under capitation, though no difference was found in emergency procedures (Chaix-Couturier *et al.*, 2000). Also, empirical studies discovered an increase in quantity of services paid under FFS in Denmark after switching from capitation to FFS. U.S. FFS practices were shown to have higher hospitalization rates than HMOs, which compensate their physician on a capitation basis (Carrin, 2003). Evidence refuting the idea of induced demand is also available (Grytten, et al., 2000, Sorensen, et al. 2001).

Capitation

Capitation refers to the practice of paying physicians per capita, i.e. per patient. Physicians receive one payment per person per fixed time period, such as a year, or a month. The fixed fee can be flat, or adjusted for various risk factors, which at the least include sex and age structure of the population. The fee is determined prospectively (Dorwart, 1988, Langenbrunner, 2004). The number of patients can be determined either on the basis of the population in a defined geographical area around the physician, or it can be determined based on rostering. Rostering is the practice of patients' formally enrolling with a physician, who then receives a per capita fee for each patient on the roster. The former method creates an economic incentive structure identical to that of a salary payment, discussed in the next section. Capitation based on a roster creates the incentive for physicians to attract as many patients as possible to their office - preferably perfectly healthy patients, who would never require actual care.

Literature presents the following arguments in favour of capitation (CMA, 2005, Hadley et al., 1999, Horowitz, 1999, Langenbrunner, 2004, Robison, 2001)

- Performs better than FFS in terms of curbing the incentive to over-provide care.
- Performs better than salary in terms of incentive to see all patients.
- Encourages good communication with patients and fosters long term relationships, especially when patients are rostered. Encourages personalized care.
- Motivates early treatment and prevention.
- Encourages cooperation between providers, since delegation of task does not result in revenue loss.
- The approach is marketable to specialists and payers, because it bears resemblance to FFS, by linking a large amount to a specific patient.
- The system is fairly easy to understand and is workable.
- Encourages good relationships and communications between specialists and PCPs, who are in control of referrals.

The capitation method also carries several shortcomings. (CMA, 2005, Horowitz, 1999, Langenbrunner, 2004, Robinson, 2001).

- Creates the incentive to increase practice volume, and to decrease intensity of care.
- Encourages selective admission of patients, as the method transfers the risk created by uncertainty in health care to providers, who might be inclined to refer or not accept risky patients.
- Patients with chronic illness might lose, as their care is costly.

- There is no incentive to limit referrals, from the perspective of PCPs, since the system does not promote risk sharing between specialists and PCPs.
- Excessive referrals unnecessarily increase overall health care costs.
- Unless correctly adjusted for illness severity, those specialists, who treat the sickest are left underpaid.
- Physicians found less likely to perform discretionary care, and experience higher level of discomfort with clinical decision making (Shen et al., 2004)

Empirical evidence is mixed. On the one hand, capitation in its numerous variants⁶ is shown to have resulted in no changes to primary physician workloads, no changes in referrals to specialists, and no changes in hospital utilization. (Chaix-Coutourier et al., 2000, Ogden et al., 1990, Rice, 1997) Other studies indicate positive effects of capitation. The methods seems to have resulted in decreases in elective surgery, decreases in referrals to private clinics and referrals in general, and decreased health care spending. (Carrin, 2003, Chaix-Coutourier et al., 2000, Hillman et al. as cited in Leger, 2000, Rice, 1997). Lastly, negative impacts of capitation include increased specialists costs, decreases in service provision, and increases in morbidity. (Rice, 1997, Stearns et al., 1992, Siu et al., 1988, Luft, 1981, Manning et al., 1986, as cited in Leger, 2000)

Salary

A salary is payment per unit of time, not related to the costs of production, activities performed, nor time spent with patients. Accordingly, it does not stimulate physicians to change behaviour affecting any of those areas. From a theoretical perspective, to a salaried physician, it makes financial sense to see as few patients as possible, and give each of them as little care as possible. Clearly, financial incentives are not the only motivating factors to a physician, and salaried physicians certainly do provide care to patients.

Salaries paid to health care providers are seen to have the following advantages. Longer visits with patients are encouraged, since there is no financial opportunity cost to spending time. This creates better relationships with patients. Providers are also more likely to engage in group consultations, design appropriate treatment plans, have no incentive to over-provide, and refer patients when necessary than they are in a FFS setting. From the perspective of physicians, salaried positions have the distinct advantage of predictability in terms of income, and often also in terms of work hours, and do not require additional administrative work associated with both FFS billings and capitation rostering (Carrin, 2003, CMA, 2005, Langenbrunner, 2004).

Salaries, however, encourage low productivity and on the job leisure. Physicians are more likely to cream skim⁷ and to refer unnecessarily than they are in FFS settings. They are also more likely to not accept patients than in both FFS and capitation setting. In addition, salaried physicians are typically in an employee relationship, a status that is generally less preferred than the independent entrepreneur status of FFS physicians (CMA, 2005, Robinson, 2001).

Empirical studies showed that salaried physicians referred less frequently, and had a lower level of activity than FFS counterparts. FFS physicians provided more visits and to more patients than salaried physicians, and offered better care continuity (Chaix-Coutourier *et al.*, 2000).

Blends – capitation with fee for service carve outs

⁶ HMO and other organizational structures in the U.S.; fundholding experiments in the U.K.

⁷ Cream skimming refers to the practice of accepting relatively healthy patients, while turning away relatively ill patients, so as to avoid associated risk.

Blended models of capitation are prototypes or variants of Quality Based Incentive Payments. Experiences with these models across Canada should be given due consideration. Given that FFS and capitation create opposing incentives, and support different goals of health care provision, a blending of the two methods is believed to minimize undesirable effects, and stimulate desirable ones. As Robinson (2001) describes, in today's health care environment, which is characterized by uncertainty and asymmetry of information, the following behaviours are required of physicians in the given health care environment:

- Physician productivity and patient service
- Risk acceptance
- Efficiency and appropriate scope of practice
- Cooperation and use of evidence based medicine.

The first two aspects are well supported by a FFS system, but the last two are discouraged in a FFS system. By itself, capitation leads to under provision of care, but when combined with FFS it provides an antidote to supplier-induced demand, which constitutes inappropriate scope of practice. Fee for service does not reward collaborative discussion, nor adaptation of protocol, nor changes in practice patterns, such as adoption of telemedicine, health education classes, delegation of tasks, or e-mail usage, among others. Capitation could draw attention to epidemiological patterns, and does not discourage collaboration, when such is advocated to physicians (Robinson, 2001).

Blended payment methods are recommended to achieve all of the above four objectives, rather than only two at a time. Blended payments include capitation with FFS carve outs, specialty department capitation with individual physician contact capitation, or FFS, and case rates for episodes of illness at the institutional level (described in a later section).

In Canada, blended payments are common in primary care remuneration, in academic health science centres, and in disadvantaged communities. The goal of blends is to preserve desirable incentives of individual payment modalities. Care is improved only if methods are blended effectively. For example: fee for service blended with capitation encourages seeing more patients and delivering services; bonuses encourage desirable target activities; time based payments encourage time consuming forms of care (CMA, 2005). Blended payment methods used in Canada can be categorized into three types (North South Group, 2005). The first, B1, is a blend of capitation and fee for service payments, as used at the primary care level in British Columbia and Ontario. Capitation with fee for service carve outs consists of two components: (i) a flat monthly payment per enrolled patient, adjusted for age and sex, limited by stop-loss provisions; and, (ii) supplemented with fee for special carved out services (often preventive, such as vaccinations and mammography, or services requiring costly supplies). The second, B2, combines a fixed payment component with fee for service billings. This type is used at the primary care level in Ontario and Québec. The third category, B3, includes all other blends, such as fee for service billings combined with bonuses, or sessional fees. These are used in Manitoba, Prince Edward Island, Nova Scotia, and Ontario.

Subsidies and grants

While not very large in terms of revenues of physicians, subsidies and grants play an important role in assisting any initiative that benefits the patient population. These might include financial support toward the purchase of equipment, facilities, or staff. In general, these financial incentives have little direct influence on patient-doctor relationship, except through changes in quality of care. They are often accompanied by conditions which might increase undue demands on physicians (CMA, 2005).

2.3. Institution level payments

Physician institutions, be they hospitals or other organizations, can receive their funding in four possible formats, payment per diem, payment per case, line item budget, or global budget. Each is described below, and each can be apportioned to individual physicians working within these institutions through any of the physician payment methods described above.

Per diem rates

Per diem rates are the equivalent of fee for service remuneration of physicians. Hospitals are paid a fixed amount for each day a patient spends in the hospital. There are various methods of estimating the value of the fixed amount. Historically, current year's payments were determined based on the previous year's statistics (Langenbrunner, 2004). More commonly today, per diem rates are risk adjusted to better differentiate among various types of illnesses and care requirements. The diagnostic related group (DRG) calculation is a method of risk adjustment for hospital patients. The institution receives a fixed sum for treatment of illness defined by the DRG. (Dorwart, 1988)

Within the DRG, the payment is predetermined and does not vary with intensity of care provided on any given day. Payment per day creates the incentive to increase the number of hospital days used by patients. If payments are based on DRGs, it makes financial sense to increase the number of hospital days for the most lucrative DRG category. Increasing the number of hospital days can be achieved through admitting more patients, or through increasing the length of stay for existing patients. The latter is much easier to justify to a payer (Carrin, 2003, Dorwart, 1988, Langenbrunner, 2004). In addition, the institution has no financial motivator to provide high intensity care, and in fact would provide no care, if profit were the only goal.

Case rates per episode of illness

Case payment per episode of illness allocates a fixed amount per case (Langenbrunner, 2004). Often the amount is based on the DRG system (Carrin, 2003). The amount is to cover the costs of care from admission of a patient to discharge. Cost overruns are the responsibility of the hospital, while cost savings can be retained by the hospital.

Because the payment per episode is divorced from actual costs of care, it is financially attractive to a hospital to admit patients with the most severe cases, and provide them with little to no care.

Case rates per episode are used in outpatient care in China and Lebanon, and for inpatient care with DRG adjustment in Austria, Brazil, Ireland, Portugal, Sweden, Spain, and U.S. Medicare, to name a few (Carrin, 2003, Langenbrunner, 2004).

While traditionally used within institutions, the case rates per episode approach has also been used at the physician level in surgery, and is slowly moving to other specialties. Surgeons are paid a set fee for pre-operative preparation, the procedure, and post-operative monitoring. The case is based on the diagnosis, and payment is made in accordance with patient condition and treatment required. This allocates the risk of the patient population being sicker than the anticipated average to the insurer, while leaving the physician with the risk related to utilization and cost control.

This method bundles the physician's direct expenditures, without turning her into a budget holder paying others, as would a fundholding arrangement. It is easiest to implement on a specialty specific basis, for instance in cardiology, where a "case" would include the evaluation and management services, physician fees while the patient is in hospital or ER, in office tests, and some procedures; but would exclude the institutional components of

hospital or ER costs, laboratory expenses, use of in-office injectibles, and non-office testing, such as radiology (Robinson, 2001)

Line item budgets

Line item budgets are allocated to institutions with predetermined specified amounts for salaries, equipment, maintenance, etc. These support cost containment and offer predictability (Langenbrunner, 2004). If the budget is based on previous years' costs, then the incentive to contain costs is diluted.

Global budgets

Global budgets are one line budgets that are meant to cover all of the institution's health care costs. Budgets can be soft, where the purchaser assumes the risk of cost overruns, or hard where the risk of cost overruns is shifted to the provider. Global budgets can be put together on the basis of inputs (as in Canada), historical spending and activities, or volume of services provided and types of cases (as in France, UK, Germany) (Carrin, 2003, Langenbrunner, 2004).

Global budgets give incentive to cream skim patients, and admit only those requiring low cost care. Global budgets also give the incentive to under-provide care, much like salaries at the provider level.

Physician Organizations

Aside from hospitals, physicians can also create organizations by forming physician groups. This is prevalent in the U.S. in both the primary and secondary care sector. Specialists are more likely to organize into groups by similar specialty. Physician organizations in the U.S. serve as intermediaries between physicians and payers, which can be private, for example Health Maintenance Organizations (HMOs), or public, for example Medicare and Medicaid.

Independent Provider Associations (IPAs) are the most common type of physician organization. IPAs negotiate with HMOs or other insurers. IPAs bring doctors together for purposes of managed care contracting, credentialing, claims payment, and medical management. Physicians retain independence over ownership and practice operation. Younger IPAs do as little as negotiate discounted fee for service rates with insurers, while mature IPAs contract with GPs and specialists. They are intermediaries between the insurer and the physician; payment by insurer to the IPA can, and often does, take different structure than payment from the IPA to the physician. The former often includes a pre-paid amount for a spectrum of services (likely on a capitation basis), while the latter has the option to pay providers fee for service, PC capitation, specialty department capitation, or various blended forms. The IPA dilutes insurance risk of capitation, and aims to motivate appropriate scope of practice through blended payment methods, and to focus physicians' attention on desired services (Robinson, 1999).

The Canadian Medical Association cautions that when physicians have to share a common pool of money, several issues might arise. Payers generally like shared pools, as they provide a predictable budget, encourage collective action, and give the opportunity to redistribute income among participants. Physicians dislike having to assume risk and having to determine how to share fairly. On the other hand, physicians appreciate having predictability in income. Physicians benefit, if the need for care falls below that anticipated by the payer, but lose, if the costs exceed estimations. (CMA, 2005)

The pilot sites for the Manitoba Physician Integrated Network experiment with Quality Based Incentive Payments are examples of physician organizations, where primary care providers and specialists share a common pool of money. Each clinic has a unique formula for distribution of clinic revenues between individual physicians.

Quality Based Incentive Payments are the latest addition to the list of remuneration methods of both individual physicians and hospitals. QBIPs are the focus of this report, and as such, are discussed in greater detail in the next few sections.

3. Quality Based Incentive Payments (QBIP)

QBIP programs have emerged across jurisdictions in a practical pragmatic fashion, rather than being founded on extensive theoretical debate. The description of QBIPs in this report is structured around real existing programs. The focus is on three jurisdictions: the United Kingdom and its Performance Based Reimbursement (PbR), the United States and its Pay for Performance (P4P) programs, as well as Australia and its Practice Incentive Program (PIP). Programs are delineated in terms of their goals, their incentive structures, their quality measures, and experiences and lessons learned. The description is general, insofar as there are commonalities between programs. Differences between programs are delineated when applicable.

3.1. Goals

Currently in the U.S., there are over 100 different P4P efforts, including Medicare/ Medicaid's 10 P4P demonstration sites: 53 P4P programs are targeted at physicians, 34 are targeted at hospitals, 18 at health plans, and 13 at consumers. Overlap exists, in that several P4P programs target two or three groups. In addition, some P4P programs are implemented in more than one state. (CMS website, 2007, Herndon, 2005, Leapfrog Group website, 2007). P4P contracts specify that health care providers (hospitals or doctors) must satisfy externally imposed quality metrics in order to be selected by the payer to enter payment contracts, or to be paid additional amounts over non-qualifying providers (O'Hare, 2005). Opponents to this approach argue that P4P could lead to doctors gaming the system, making access to care more difficult for non-compliant patients, and turning doctors' goals from healing to scoring (Herndon, 2005).

The participants in P4P programs include purchaser groups, health plans, and healthcare providers. Each of these groups has different goals when implementing or accepting P4P programs. Purchasers (in the U.S. they are often employers), support P4P programs, as they hope to see the following improvements to quality: reduction in clinical practice variation, reduction in errors, reduction in acute exacerbations, and an increase in the transparency of provider performance for consumers. Health Plans, while sharing the purchasers' goals, also consider objectives of increased market share, and administrative simplicity and cost. Providers are least involved in the design of the program, and mostly act as recipients/ participants, who fear losing patient volume, if they do not join P4P programs. Providers hope for the ability to do the right things for their patients, without losing market share, and perhaps achievement of a competitive edge in the marketplace (Baker *et al.*, 2003). Public health insurance systems, such as that of Australia, or the U.K., implemented one national QBIP. The following is a categorization of goals, as appear common to most forms of QBIP programs. (U.K. Department of Health, 2007, Payment by Results, Hanchack, 1997, Herndon, 2005, Pink, 2006, Vounasis, 2006).

The aim is *"To incorporate incremental payments that reward physicians for delivering high quality care in areas that are not compensated via traditional payment methods. Desirable components of care include non office visit based care, care by mid level non-physician providers, investment in office systems and electronic medical records."* (Landon, 2006)

Goals with respect to patient care

- To improve health care quality
- To improve patient health outcomes, and patient safety
- To support patient choice and diversity
- To improve quality through direct explicit financial rewards and support

Goals with respect to physician

- To provide feedback to specialists, and increase overall performance
- To motivate and optimize individual performance
- To link payment to activity (adjusted for case mix)

Goals with respect to health care system

- To decrease costs,
- To increase likelihood of achieving organizational targets
- To provide transparent, rules-based system for paying trusts
- To ensure fair and consistent basis for hospital funding
- To encourage activity for sustainable waiting time reductions

Goals with respect to insurer

- To decrease costs per amount of quality, and hence increase in health system efficiency
- To increase emphasis on management of care
- To reform physician payment systems
- To reward efficiency

3.2. Types of incentives

Linking compensation to performance is not a new concept in industrial organization and other economic literature. Principal agent theory, which, as mentioned, serves as a framework for discussion of provider incentives in the health care market, prescribes that payment be linked to outcome, whenever effort levels of the employee are not observable. While straightforward in conceptual form, operationalization of this idea is fraught with challenges.

Two crucial decisions need to be made. First, what should be the mechanism that ties reimbursement to performance? This question is addressed in this section of the report. Second, how can performance be measured? This question is subject of the next section.

QBIP programs are added to existing sources of physician income, such as FFS billings, capitation payments, or salaries. GMS and PMS practices in the United Kingdom are reimbursed on a capitation basis. An individual physician's office in the U.S. is usually involved with several health plans concurrently, and has varied sources of income. P4P programs in the U.S. are employer-supported, and give additional payment for achievement of quality. Below is a description of the most common types of financial and non-financial incentives. Incentive payments are paid quarterly, semi-annually, or annually (Baker, *et al.*, 2003). The literature notes that most QBIP programs put no emphasis on quality improvement, nor make any operational distinction between quality improvement and meeting a fixed target (Rosenthal *et al.*, 2004).

Since QBIP programs are relatively new, a widely used typology for this kind of payment system has not been established in the literature. Conrad *et al.* (2004) observe four types of incentive structures emerging in the marketplace:

- Bonus payments for superior quality,
- Tiered co-payments (lower patient cost sharing for use of higher quality providers)
- Higher reimbursement rates for high quality doctors,
- Quality "infrastructure" grants.

These are specific examples, which can be placed into a set of broader categories.

The following classification might be useful, based on Baker *et al.*, 2003, CMA, 2005, Golden, 2006, MedPac, 2003, Pink, 2006, Wranik, 2007) . QBIP programs include two types of incentives,

financial (extrinsic), and non-financial (intrinsic). Both types can be aimed at changes in physician performance, or enablers toward that change. Enablers include practice capacity development, as well as influences on patient behaviour. The following matrix categorizes QBIP incentives.

	Physician Performance	Enablers
Financial Rewards (Extrinsic)	<ul style="list-style-type: none"> • Bonuses • Withholds • Performance fee schedule • Higher rates for higher quality providers • Additional Payments 	<ul style="list-style-type: none"> • IT infrastructure grants • Quality grants
Non-financial Rewards (Intrinsic)	<ul style="list-style-type: none"> • Performance profiling and publicizing • Practice sanctions • Selective contracting with health care plans or providers 	<ul style="list-style-type: none"> • Technical assistance • Data collection on quality • Educating patients on quality issues • Patient incentives to choose high quality providers (ex. tiered co-payments)

Extrinsic Rewards for Physician Performance

Bonuses – Annual payments by health plans that are based on achievement of minimum targets of specified measures. These range from 5 to 20 percent of total physician reimbursement (Baker *et al.*, 2003). Target payments are often offered for services that are clinically desirable, but not financially attractive, unless performed in bulk.

Target payments and bonuses are generally viewed positively by patients, as they tend to support activities shown to be desirable. They might, however, detract from other regular physician activities, especially when coupled with regular non-performance based payment, or can force the physician to perform disliked tasks. There is also potential to manipulate clinical practice decisions made by physicians, by giving incentives to focus on specific activities that would not be prioritized based on clinical decision making alone.

Withholds – Failure to achieve minimum targets of specified measures is penalized by the health plan withholding a percentage of reimbursement. This option is generally not liked by physicians, as it puts their compensation at risk; therefore participation in these programs is not voluntary. Withholds correspond to competitive QBIP systems, where funds to reward best performers are generated via penalties garnished from worst performers.

Adjustable fee schedules for specialists – Based on specific measures in the previous period, the fee schedule is adjusted accordingly. Fees for desirable elective procedures are increased, to encourage provision thereof.

Additional payments – Reimbursement for tasks that are beneficial to patients, but not otherwise remunerated via the existing payment mechanism. Examples of such tasks include care planning and management of chronically ill patients, patient education, reporting results from patient registries, and other tasks. In a fee for service system, additional payments are equivalent to adding new billing codes, and as such, do not constitute a change to the payment method.

Shared savings – refers to a calculation of savings from quality improvements, and a sharing of them with those who invested resources to improving quality. The calculation of savings is based

on cost effectiveness, or cost utility estimates, which are notoriously difficult to develop for preventive treatment, or care for chronically ill patients.

Extrinsic Rewards for Use of Enablers

Quality grants – Funding for a specific quality improvement project to a physician organization (IPA, MSO, IDN) awarded by the health plan or purchaser. Quality grants are allotted based on pre-determined criteria, and awarded via an application process.

IT Infrastructure grants – Similar to general quality grants, but earmarked specifically for investment in software that allows tracking of required quality indicators. Investment in IT should, and often is, a prerequisite for acceptance into a QBIP plan. Without reliable data tracking mechanisms, a QBIP system cannot be implemented.

Intrinsic Rewards for Physician Performance

Administrative relief – Exemptions from various administrative requirements for high performing providers. This option is similar to flexible oversight, which identifies potentially less burdensome regulatory requirements, if an organization demonstrates a high level of performance or effort.

Public report cards, community recognition, and honour rolls – Recognition of best demonstrated practices among high performing providers. Physicians are rewarded with prestige and positive public exposure. Disclosure of quality information on individual providers improves care, because providers want to maximize their performance, and because more patients are likely to opt for better quality care.

Scoring of providers – Scoring of providers is achieved via two methods or a combination thereof. Threshold scoring is when the performance of a provider is measured against a yardstick in absolute terms. Ranking of providers compares the performance of a provider relative to other providers, often done with reference to a statistical distribution.

Additional intrinsic incentives, that are not unique to QBIPs, include peer review, second opinion, utilization management, medico legal sanctions, and continuing medical education.

Intrinsic Rewards for Use of Enablers

Employer health plan incentive agreements – Large employers or coalitions influence the insurance contract between insurer and provider, thereby indirectly affecting the provider's incentive structure and practice behaviour. These exist in privately funded systems, such as the U.S.

Patient benefit design and tiered networks – Patients are channeled by the insurer to high performing providers, giving providers additional incentive to perform according to insurer expectations. Examples of patient plan design include requiring lower cost-sharing amounts for enrollees for plan premia, or lower co-payments for choosing high quality providers.

Competitive and Non-competitive models

Competitive models, those that reward relative performance, are most common in 37 P4P programs described in Rosenthal *et al.* (2004). Competitive models often offer tiered incentive payments. Non-competitive models reward providers for meeting a specified threshold, or for implementing processes and systems.

3.3. Quality measurement

The selection of indicators of quality, and their measurement, is arguably the most difficult task in the design of a QBIP program. At the same time, the task is a crucial determinant of effectiveness of the QBIP payment system. Even the most intricate design of incentives fails, when payment is tied to inappropriate indicators. This section addresses the question of defining quality, operationalizing the concept, and selecting indicators. Examples of commonly chosen indicators are provided.

The Institute of Medicine (IOM) defines quality of care as:

“The degree to which health services for individuals and populations increase likelihood of desired health outcomes and are consistent with current professional knowledge.” (IOM, 2006)

In this view, valid measures of quality include processes or outcomes that are affected by medical or health care. Important to note is the focus on achievement of desired health outcomes. The implicit assumption is that population health has been assessed, population health needs are known, and gaps in service provision have been determined, PRIOR to implementation of the QBIP program.

Operationalization of quality measures

The broad definition of quality measures needs to be operationalized before they can be applied in practice. Operationalization refers to the process of indicator selection and specification of measurements. The table below shows the desirable characteristics good indicators should possess, as well as the dimensions, along which those characteristics have to be met. (Baker et al., 2003, MedPac, 2003, Milgate et al., 2006, Wranik, 2007).

		Dimensions	
		<i>Technical</i>	<i>Logistical</i>
Characteristics	Credibility	Validity	Evidence based
		Reliability	Broadly understood
			Accepted by providers
	Feasibility	In provider’s control	Data easily generated
		Responsiveness	Data easily collected
		Attainable	

The characteristics listed are not unusual in the selection of indicators in all areas of empirical social research. Description is provided with focus on QBIP in health care. Validity refers to the extent, to which a selected indicator actually reflects the underlying concept. Validity can be discussed using various methods. Concept, or face validity is established, when the indicator logically matches the concept it is to reflect. Empirical validity is established, when a selected indicator empirically behaves as one would expect the underlying concept to behave. Given the IOM definition of quality in health care, a measure of use of evidence based treatments carries face validity, while a measure of ambiance in the hospital does not. Similarly, patient satisfaction carries empirical validity as a measure of quality, if it has a positive empirical relationship to other quality indicators.

Reliability refers to the degree to which the value of an indicator is independent of environment and time. Assessment of patient status by health professionals may not be a reliable measure, if performed by different health professionals at different times. BMI is a reliable measure, insofar

as the height and weight of a patient will have the same value, regardless of researcher, or location.

In the logistical sense, credibility of indicators also requires that they be evidence based, broadly understood, and accepted by providers. Evidence based indicators are those which are supported by clinical evidence of effectiveness in terms of leading to improved patient health outcomes. Indicators need to be broadly understood in systems that are implemented on a large scale, to ensure that all physicians aim at achievement of the same concept, as reflected by its indicator. In addition, providers need to accept that indicators selected and targets are attainable, and logically connected to quality of health care.

The second desirable characteristic of good indicators is feasibility. Feasibility along the technical dimension requires that indicators selected be under the provider's control, responsive, and attainable. A measure of patients' days of missed work, while related to the patient's health, does not satisfy these criteria. The physician's behaviour may have very little influence on the days of work missed by their patient. An indicator such as life expectancy is not responsive to physician behaviour, which affects life expectancy over a long period of time and to a small degree. Setting a target such as "100% of patients will be highly satisfied with care provided", is not attainable, and therefore not feasible.

In the technical sense, indicators satisfy the feasibility criterion only if data are easily generated, and easily gathered. Measuring each patient's blood sugar levels on a daily basis might be clinically most desirable, yet technically not feasible.

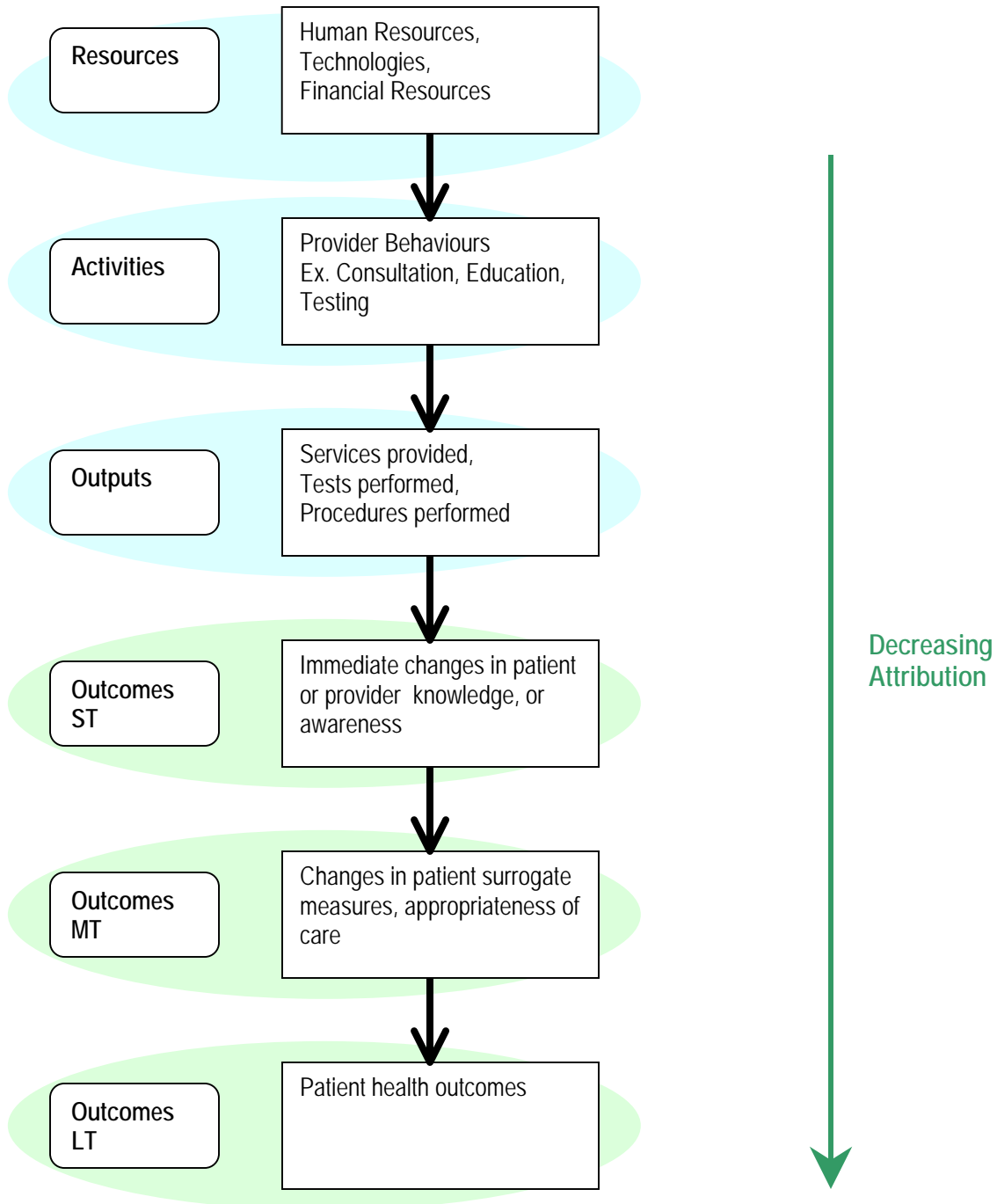
Dimensions of care and quality

Indicator selection must also be sensitive to the component of care of which the quality is to be measured. Following a logic model approach to the visualization of the health care process, one can distinguish between the process or structure of care, the outputs generated in health care, and the outcomes. While good patient health outcomes is the ultimate goal of any health care system, the measurement of care quality must take into consideration the following: The care process begins with a series of inputs into the system, including health human resources, technologies, and financial resources. These are used to create a series of activities of health care providers – their behaviours. An example might be carefully listening to patients during consultation. Each activity results in an output (although in any service industry, there is generally overlap between activities and outputs), such as longer consultations. The outputs, in turn, lead to an outcome. Using a logic model framework, outcomes can be stratified into short term, medium term, and long term. Given the example, longer consultations would lead to improved diagnosis in the short run, more appropriate treatment recommendation in the medium run, and improved patient health measures in the long run. (Wranik, 2007)

A change in financial resources channeled to physicians creates an impetus at the input level, which presumably is channeled through the chain in the logic model. Remuneration structures affect the activity (behaviour) of physicians, which leads to changes in outputs, and changes in all levels of outcome. Improvements to quality could therefore be measured at any step in this process, and in fact, various QBIP programs do target measures at various levels in the logic model. In this context, two issues warrant discussion: effectiveness and attribution.

When payment is linked to logic model components at higher levels, such as activity or output, the question of effectiveness arises. Inducing the physician directly to change her/his behaviour, or to change the outputs is desirable only insofar as those changes can effectively influence patient health outcomes. The robustness of linkages at lower levels of the logic model would have to be assumed, or proved empirically. On the other hand, payment can also be tied to lower components in the care process, such as patient outcome. As patient health is the ultimate goal, many support this approach. The question of attribution arises in this scenario. Patient health,

while most important in the health care system, is influenced by a myriad of external forces that are not related to physician behaviour, including patient characteristics, environment, and patient compliance. To what degree changes in patient health can be attributed to physician behaviour needs to be supported empirically.



Examples of quality measures

The PbR program in the U.K. is aimed at primary care providers. This is true on many U.S. P4P programs, although a large portion of the latter targets hospital quality. Measures selected are different at the physician level and at the hospital level.

QBIP programs use any subset of the following general categorization of quality measures. Quality can be assessed in four domains - clinical service, patient experience, organization, and the technical domain.

Clinical service is measured using indicators that directly relate to patient conditions, health outcomes, milestones for disease states, tests administered, treatments administered, or other health care services provided. The actual measures selected depend on two factors. First, the decision as to the component of the care logic model that should be directly linked to payment. Second, the decision on the actual health areas to be targeted in any particular health system. The latter is a crucial first step in the design of a QBIP system, as it identifies disease conditions, preventive measures, or any other clinical areas that require improvement in care quality. The decision should be based on population health needs alongside identified gaps in service provision, and/or economic estimates of benefit of specific health care procedures. (Baker et al., 2003, Milgate, 2006, Pink, 2006, Shea, 2007)

Areas of health care in need of improvement tend to be chronic disease management, as well as general early detection and prevention of illness. Examples include asthma, diabetes, coronary artery disease, breast cancer, cervical cancer, and child vaccinations. (Shea, 2006)

Patient experience assesses the interpersonal effectiveness of the patient – provider encounter, most often using satisfaction scales, and sometimes external review. The logical and most common measure of satisfaction is the patients' assessment thereof. In the U.S., patient satisfaction is generally measured via provider based survey, in combination with the Consumer Assessment of Health Plan Survey (CAHPS). Secondly, providers might be asked to rate their own satisfaction, or their perception of patient satisfaction, especially, when part of a larger health care delivery institution, such as a clinic or hospital. External measures of patient experience might include observation of staff responsiveness, method and promptness of dealing with patient concern, and other such measures, which are rare due to high cost. (Baker et al., 2003, Milgate, 2006, Pink, 2006, Shea, 2007)

The structure of the system refers to organization and method of service delivery. Specific measures vary, depending on goals of the health policy maker. They might include: the extent to which care is delivered holistically, and in an interdisciplinary fashion; the process for review of patient files; the protocol on patient follow up; waiting times; difficulty with scheduling appointments, and other measures. (Baker et al., 2003, Milgate, 2006, Pink, 2006, Shea, 2007)

Lastly, the technical domain is assessed in terms of the extent to which information technology, electronic patient records, and data collection technology has been implemented in the delivery system. As an example, measured concepts might include ability to integrate clinical data, adoption of computerized physician order entry, or use of decision support software. Physicians and practices are often offered assistance with the development of electronic records and software systems. The technical domain might be treated as separate or as a component of the QBIP program. (Baker et al., 2003, Milgate, 2006, Pink, 2006, Shea, 2007)

Quality of care in the U.S. is commonly assessed via some subset of the Health Plan Employer Data Information Set (HEDIS), which is a comprehensive list of measures, alongside a data collection software, available for purchase to all providers or health plans at <http://www.ncqa.org/programs/hedis/>. Measured domains are effectiveness of care and use of service (the clinical domain), access/ availability of care (structure), satisfaction with experience

of care (interpersonal effectiveness), as well as U.S. specific measures, such as costs of care, health plan descriptive information, and stability of health plan. The IT domain is not assessed using this tool.

Hospitals also often use measures created by Leapfrog, JCAHO, the National Quality Foundation, and Centers for Medicare/ Medicaid Services. Categories of measures include safety, compliance with evidence based medical practices, connectivity or use of IT, and medical records. (Milgate et al., 2006)

3.4. Examples

U.K. National Performance Based Reimbursement Program

The PbR remuneration system is part of a larger Quality and Outcomes Framework (QOF), which is in turn part of the General Medical Services contract introduced in April 2004. QOF measures practice achievement against a range of evidence based clinical indicators, and against a range of indicators on practice organization and management. Practices score points according to their levels of achievement of these indicators, and practice payments are calculated based on points achieved. (QOF, 2004)

Data is collected using the Quality Management Analysis System (QMAS), and further used to calculate QOF payments. Since QMAS is a single national system, consistency in the calculation of quality achievements is ensured, and meaningful feedback can be provided to providers and health authorities. (QOF, 2004)

The PbR system in the United Kingdom is a nation wide program implemented in April 2004. The program is most comprehensive in that it includes most primary care practices in the jurisdiction. The contract is voluntary, but its lucrative nature has attracted the majority of practices. Details of the program, including incentive structure, and quality measures, are described in previous sections. Below is a synthesized description of the PbR system.

The majority of GPs in the U.K. are independent contractors, practicing in groups, with an average group size of 4.4 physicians, and an average list size of 5200 patients. Ninety percent of physician groups have nationally negotiated General Medical Services (GMS) contracts, and the remaining 10 percent have locally negotiated Primary Medical Services (PMS) contracts. (Gravelle et al., 2006)

Practices with GMS contracts receive population adjusted capitation payments, as well as lump sum practice allowances. In addition, they have been receiving target incentives prior to April 2004, which have now been included in the PbR system. GMS practices have dispensing rights in areas geographically distant from pharmacies. PMS contracts pay out a lump sum for physicians to provide all services they would under a GMS, and are reimbursed for additional services provided to target groups. Target incentives are reduced under the PbR system. (Gravelle et al., 2006)

The PbR system targets four areas of PC practice: clinical, organization, patient experience, and additional services. The clinical area focuses on 11 domains, or disease types, which include Coronary Heart Disease, Stroke/ Ischemic attack, hypertensions, hyperthyroidism, diabetes, mental disorder, COPD, asthma, epilepsy, and cancer. Organization is measured along 5 domains, including records and patient information, patient communication, education and training, practice management, and medicine management. Patient experiences are garnered via patient satisfaction surveys, as well as an external assessment of consultation length. Additional services such as neonatal screening can be rewarded, as well. Furthermore, practices can gain points in the additional services category, measured along 4 dimensions, including cervical

screening, child health surveillance, maternity services, and contraceptive services. (Gravelle et al., 2006)

Providers are able to score up to 550 points for the achievement of targets for 76 clinical indicators. Points are allocated based on workload required for measured tasks. Providers can also score up to 170 points for the achievement of 56 organizational indicators. Patient experiences amount to a maximum total of 100 points. Finally, additional services can earn up to 130 points. An additional maximum of 100 points are rewarded for overall clinical achievement and holistic care delivery. In total, practices can earn up to 1050 points. (QOF, 2005, Roland, 2004)

As of 2005, each point is worth £120 to the average practice, which translates into a total potential increase in earnings of £126000 per group practice, with an average of 4-6 physicians per practice. The average physician can expect an increase in earnings of £42,000 per year. The additional revenues are to cover additional investments in quality improvement, such as the hiring of nurses. A total budget increase of £1.8 billion was allocated to high quality care, which amounted to more than 20% of the previous GP budget. (QOF, 2005, Roland, 2004)

U.S. fragmented system – program implemented by private insurance organizations

The U.S. health care market is fragmented, and a comprehensive government strategy or approach to physician remuneration does not exist, as was the case in the U.K., or Australia. P4P programs are being adopted in numerous variants across the U.S. The Leapfrog Group, a coalition of large employers, has recorded over 140 P4P programs across the U.S., many of them serving more than one state. 45% of HMOs have at least one contract with a purchaser that includes bonuses or penalties tied to performance. On average, these HMOs serve 324 thousand patients, 75% of whom are assigned to a primary care physician roster. The majority of these HMOs offer capitation payments, in addition to which, more than half the HMOs reward individual physicians on a P4P basis. HMOs themselves are paid on a P4P basis. (Rosenthal, 2006)

The Centers for Medicare and Medicaid Services (CMS) are publicly funded insurance schemes for those below a specific income threshold, or the elderly. CMS have experimented with various changes and modifications to physician improvement, and have jumped on the P4P bandwagon in the last decade. CMS are pegged as a prototype of P4P programs by Rosenthal (2004), as well as a leading payer in the current system by Shea (2007).

Medicare has various initiatives to encourage improved quality of care in all health care settings where Medicare beneficiaries receive their health care services, including physicians' offices and ambulatory care facilities, hospitals, nursing homes, home health care agencies and dialysis facilities.

CMS is collaborating with a wide range of other public agencies and private organizations who have a common goal of improving quality and avoiding unnecessary health care costs, including the National Quality Forum (NQF), the Joint Commission of the Accreditation of Health Care Organizations (JCAHO), the National Committee for Quality Assurance (NCQA), the Agency for Health Care Research and Quality (AHRQ), the American Medical Association (AMA), and many other organizations. CMS is also providing technical assistance to a wide range of health care providers through its Quality Improvement Organizations (QIOs).

CMS has developed and implemented a set of P4P initiatives to support quality improvement in the care of Medicare beneficiaries. In addition to the initiatives for hospitals, physicians, and physician groups described below, CMS is also exploring opportunities in nursing home care – building on the progress of the Nursing Home Quality Initiative – and is considering approaches for home health and dialysis providers as well. Finally, recognizing that many of the best

opportunities for quality improvement are patient-focused and cut across settings of care, CMS is pursuing P4P initiatives to support better care coordination for patients with chronic illnesses.

CMS – Hospital Quality Initiative – This is part of HHS’s broader National Quality Initiative that focuses on an initial set of 10 quality measures by linking reporting of those measures to the payments the hospitals receive for each discharge. Hospitals that submit the required data receive the full payment update to their Medicare DRG payments. Nearly all (98.3%) of the hospitals eligible to participate in this program are complying with the requirements of the provision.

CMS – Premier Hospital Quality Incentive Demonstration - The purpose of the demonstration is to improve the quality of inpatient care for Medicare beneficiaries by giving financial incentives to almost 300 hospitals for high quality. Under this demonstration, CMS is collecting data on 34 quality measures relating to five clinical conditions. Hospital specific performance will be publicly reported on CMS’s web site. Hospitals scoring in the top 10% for a given set of quality measures will receive a 2% bonus payment on top of the standard DRG payment for the relevant discharges. Those scoring in the next highest 10% will receive a 1% bonus. In the third year of the demonstration, those hospitals that do not meet a predetermined threshold score on quality measures will be subject to reductions in payment.

CMS – Physician Group Practice Demonstration – Mandated by the Medicare, Medicaid, and SCHIP Benefits Improvement and Protection Act of 2000 (BIPA), this demonstration is the P4P initiative for physicians under the Medicare program. The demonstration rewards physicians for improving the quality and efficiency of health care services delivered to Medicare fee-for-service beneficiaries. The demonstration seeks to encourage coordination of Part A and Part B services, promote efficiency through investment in administrative structure and process, and reward physicians for improving health outcomes. Ten large (200+ physicians) group practices across the country participate in this demonstration as of April, 2005. The physician group practices will be able to earn performance based payments after achieving savings in comparison to a control group. The performance payment is largely based on various quality results.

Medicare Care Management Performance Demonstration – Modeled on the “Bridges to Excellence” program, this is a three-year pay-for-performance demonstration with physicians to promote the adoption and use of health information technology to improve the quality of patient care for chronically ill Medicare patients. Doctors who meet or exceed performance standards established by CMS in clinical delivery systems and patient outcomes will receive bonus payments for managing the care of eligible Medicare beneficiaries. In contrast to the Physician Group Practice Demonstration, this demonstration, which is currently under development, is focused on small and medium-sized physician practices. It will be implemented in four states: Arkansas , California , Massachusetts , and Utah , with the support of the Quality Improvement Organizations in those states.

Medicare Health Care Quality Demonstration –This demonstration, which was mandated by section 646 of the MMA, will be a five-year demonstration program under which projects enhance quality by: improving patient safety, reducing variations in utilization by appropriate use of evidence-based care and best practice guidelines, encouraging shared decision making, and using culturally and ethnically appropriate care. Eligible entities include physician groups, integrated health systems, or regional coalitions of the same.

CMS – Chronic Care Improvement Program – This pilot program will test a population based model of disease management, whereby the participating organizations are paid a monthly per patient fee for managing a population of chronically ill beneficiaries with advanced congestive heart failure and/or complex diabetes. These organizations, which include disease management vendors and larger organizations such as insurance companies, must guarantee CMS a savings

of at least 5% plus the cost of the monthly fees compared to a similar population of beneficiaries. Payment of fees is also contingent upon performance on quality measures and satisfaction of both beneficiaries and providers. Nine sites have been selected for the pilot phase: Humana in South and Central Florida, XLHealth in Tennessee, Aetna in Illinois, LifeMasters in Oklahoma, McKesson in Mississippi, CIGNA in Georgia, Health Dialog in Pennsylvania, American Healthways in Washington, DC and Maryland, and Visiting Nurse Service of NY and United Healthcare in Queens and Brooklyn, New York. After two years, pending successful interim results, this pilot may be expanded more broadly, possibly nationally.

CMS – ESRD Disease Management Demonstration – This 3-year demonstration will test a fully case-mix adjusted payment system for an expanded bundle of end stage renal disease (ESRD) services. A portion of the payment will be linked to ESRD-related quality measures. An advisory board for the demonstration is required by the legislation and will be holding its first public meeting in February.

CMS – Disease Management Demonstration for Severely Chronically Ill Medicare Beneficiaries – This demonstration, which began enrollment in February 2004, is designed to test whether applying disease management and prescription drug coverage in a fee-for-service environment for beneficiaries with illnesses such as congestive heart failure, diabetes, or coronary artery disease can improve health outcomes and reduce costs. Three disease management organizations are participating. They receive a monthly payment for every beneficiary they enroll to provide disease management services and a comprehensive drug benefit, and must guarantee that there will be a net reduction in Medicare expenditures as a result of their services. Submission of data on a number of relevant clinical measures is required to permit evaluation of the demonstration's impact on quality.

CMS – Disease Management Demonstration for Chronically Ill Dual Eligible Beneficiaries – Under this demonstration, disease management services are being provided to dually (Medicare & Medicaid) eligible beneficiaries in Florida who suffer from advanced-stage congestive heart failure, diabetes, or coronary heart disease. The demonstration provides the opportunity to combine the resources of the state's Medicaid pharmacy benefit with a disease management activity funded by Medicare to coordinate the services of both programs and achieve improved quality with lower total program costs. LifeMasters, the demonstration organization, is being paid a fixed monthly amount per beneficiary and is at risk for 100% of its fees if performance targets are not met. Savings above the targeted amount will be shared equally between CMS and LifeMasters. Submission of data on a variety of relevant clinical measures is required to permit evaluation of the demonstration's impact on quality.

CMS – Care Management For High Cost Beneficiaries – This demonstration tests models of care management in a Medicare fee-for-service population. The demonstration will target beneficiaries who are both high-cost and high-risk. The project was launched in 2005. The payment methodology is similar to that implemented in the Chronic Care Improvement Program, with participating providers required to meet relevant clinical quality standards as well as guarantee savings to the Medicare program.

Bridges of Excellence – Coalition of large employers, health plans, the National Committee for Quality Assurance, MEDSTAT, and WebMD Health. Employers include General Electric, Ford, Humana, Procter and Gamble, UPS, BCBS, Tufts, United, and Aetna. The Bridges to Excellence program combines three distinct initiatives – Diabetes Care Link, Physician Office Link, and Cardiac Care Link. Each link tracks an established set of measures, and each measure is scored on a point system.

The Physician Office Link program offers both financial and non-financial incentives for physicians to implement specific processes that reduce errors and increase quality. Physicians are able to qualify for financial bonuses upon investing in information systems and care management tools. The latter help provide customized and integrated care over time, rather than responding to patients symptoms during office visits only. Physicians can earn up to \$55 per patient per year. Additionally, a report card is created for each physician office, describing office performance in terms of specified measures. The report card is made available to the public.

The Diabetes Care Link provides annual bonus payments to physicians, who demonstrate good control of their patients with diabetes. A reward program is available to patients, as well, to take an active role in managing their condition. Physicians can choose one of two scoring and reward tracks. The first is an annual certification track that requires the submission of outcomes measures, including data on HbA1c, blood pressure, and lipid testing. The second is a three year track that requires the submission of process and outcomes measures. By obtaining a one year certification, or a three year recognition through the NCQA's Diabetes Physician Recognition Program, physician qualify to receive a \$100 annual bonus per eligible diabetes patient treated. (Leapfrog, 2007)

Aetna is a multi state health plan, which is also a part of the Bridges of Excellence coalition. In 2004, the plan launched a network of physicians, who are reported to have better outcomes in high cost specialties, including cardiology, cardiothoracic surgery, gastroenterology, general surgery, orthopedics, and obstetrics/gynecology. The company does not pay physicians for being part of the network, but offers non-financial incentives. Physicians are pegged as higher quality performers, and patients do not need referrals to visit networked physicians. Both measures have the potential of increasing practice volume and thereby increasing income. (Leapfrog, 2007)

The Kaiser Permanente Medical Care Program is offered in nine states across the U.S. The majority of payment by the Kaiser Foundation Health Plan (KFHP) to the Permanente Medical Group (PMG) for medical services is through medical group capitation. Small additional amounts may be paid directly to PMG by KFHP for specific clinical quality achievements in some regions. For most PMGs, approximately 5% of total compensation is variable, based on payment tied to clinical quality (measured using HEDIS and other internal quality measures), and patient satisfaction (using CAHPS, and internal patient satisfaction surveys).

In turn, PMG compensates its individual physicians through a unique formula for each region that includes both fixed and variable pay components. Variable components are based on quality related contributions by physicians in three areas: specific clinical quality processes and outcomes, patient satisfaction measures, and professional contributions to the medical group. Again, roughly 5% of a physician's salary is put at risk. The amount of payment increases with increased excellence in the three quality areas. Physicians whose performance is not satisfactory receive less variable pay, no variable pay, and may face further disciplinary action, including termination of contract. (Leapfrog, 2007)

Integrated Healthcare Association (IHA) is California based association of Aetna, Blue/Cross/Blue Shield of California, CINGA Healthcare of California, Inc., HealthNet, and Pacificare. Under this plan, physician groups receive bonus payments based on previous year's performance, as evaluated on a consolidated performance scorecard. The scorecard is available to the public.

The scorecard includes three areas: clinical quality (50%)⁸, patient satisfaction (40%), and investment in IT (10%). Clinical quality assessment is based on performance in treating three chronic conditions (asthma, diabetes, and coronary artery disease), and providing three preventive services (breast cancer screening, cervical cancer screening, and childhood immunizations). Patient satisfaction is assessed in terms of communication with physician,

⁸ Weight assigned to this area in the scorecard for the year 2004.

specialty care received, timely care and service, and overall rating of care. IT is assessed on the basis of the physician group's ability to integrate data at the group level, or provide physicians with data at the point of care. Each health plan separately distributes incentive payments to physician groups based on scorecard results. The amount of payment and method of distribution varies between plans. The use of bonus payments is at the discretion of the physician group, each of which is encouraged to pass a portion or all of the bonus to individual physicians.

The above are only a few examples of the vast number of P4P programs across the U.S. Rosenthal (2004) provides a general description of P4P programs in the U.S. in terms of five main features; the percentage of patient population served under the programs, the magnitude of rewards to physicians, the targeted dimensions of quality (structure, process, or outcome), the mechanism of allocation of rewards across providers (degree of competition), and the market share of payers.

The P4P programs in the U.S. health care market can be characterized as follows. Dimensions of quality measures fall into three categories, clinical processes, structural measures, and patient experience. The latter dimension generally carries less weight. Clinical measures selected are almost always a subset of the Health Plan Employer Data Information Set (HEDIS), and relate to primary prevention (cervical cancer screening, mammography, vaccinations), and secondary prevention (asthma, diabetes, coronary artery disease). Measures at the hospital level might include in hospital mortality, complication rates, and re-admission rates. (Rosenthal, 2004)

P4P programs are predominantly competitive (56%), and often include a tiered reward system, where two or more tiers of top performers are eligible for bonus payments. The majority of non-competitive programs offer a bonus for fixed target achievement, and partial achievement is rarely rewarded. In terms of the magnitude of awards, most range from 1% to 10% of total reimbursement for achievement of clinical targets, and maximum of \$55 per patient for achievement of system implementation targets. Almost no programs reward improvement in quality *per se*. (Rosenthal, 2004)

4. Experiences to date

4.1. Theorized Effects

Arguments in support of QBIPs are straightforward. If quality of health care is not at a desired level, this can in part be attributed to physician behaviour. Tying physician reimbursement to those components of physician performance that are believed to increase quality of care, will indeed result in such increase. Aside from the practical consideration of how to measure quality, which is discussed in a later section, this simple logic is quite appealing.

There are, however, several important theoretical considerations, and potentially negative outcomes. First, it is not clear that physician reimbursement *per se* is a strong enough motivator to change performance. Second, the various structures of incentive design may result in various perverse incentives to physicians, indeed, changing their behaviour, but not in the desired fashion. Physicians have the possibility of gaming the system in pursuit of financial gain.

Furthermore, incentives must be designed with great care, as they may orient clinical services to those acts which are rewarded with a performance bonus. Accordingly, the importance of designing the performance indicators in accordance with the desired areas of emphasis cannot be overemphasized.

Given that most QBIP initiatives are less than five years old, empirical evidence of their effectiveness in terms of physician behaviour changes is limited. Likewise, there are few evaluation studies examining the success of these experiments. Since most QBIPs aim to

improve physician performance in primary prevention and chronic disease management, the effects on patient outcomes cannot be observed in the short run.

Literature has questioned the assumption that money is a strong motivator of performance, and that financial motivation is the only cause of performance. (Golden, 2006). A person's performance is a function of many factors, including motivation, capabilities and resources. Motivation is determined by financial incentives, as well as non-pecuniary job characteristics, such as professional pride and recognition, or opportunity for challenge. The professional environment, organizational culture and teamwork also contribute to performance and attainment of goals.

Furthermore, financial monitoring of physician behaviour or patient outcomes is difficult due to both being unobservable or uncertain. Rewards based on physician behaviour assume that a desired behaviour will lead to specific outcomes. Rewards based on outcomes assume that physician behaviour is the main determinant of outcomes. If these assumptions are not true, QBIPs may not be effective. If both behaviour and outcome are unobservable or un-measurable, normative controls and values based motivation must be relied upon, which has been present in the health care sector in training of health professionals. (Golden, 2006).

Lastly, pay based on performance has long been used as a remuneration method in industries other than health care. Evidence from other industries suggests that certain conditions decrease the effectiveness of payment for performance. Nearly all these conditions hold in the health care sector. They include complex technologies, ambiguous tasks, a professional culture that stresses cooperation, difficulties with quality measurement, tenuous connection between inputs and outputs, rising perception of inequality, a social culture that does not support focus on money, and more reliance on intrinsic motivation. (Golden, 2006)

The increasing use of multidisciplinary teams for delivery of primary care services is another complicating factor in the utilization of pay for performance targeted specifically at physicians.

Gravelle et al. (2006) develop an economic model of the British PbR system. As described above, the physician's performance bonus is tied to the number of points collected. For each target indicator (treatment X), the number of points collected is a function of three arguments: the number of patients eligible for treatment X, the number of eligible patients declared an exception, and the number of eligible patients given treatment X. The relationship with points collected is negative, positive, and positive, respectively. The number of points earned for a specific disease domain can be influenced by increasing the number of patients, who receive treatment X – the desirable outcome. However, an increase in the number of patients declared an exception also leads to a higher number of points, with no improvement in performance.

Furthermore, the value of each point depends positively on practice list size in relation to the average practice list size. This implies that larger practices will by default earn a higher bonus, at any level of performance achievement. (Gravelle *et al.*, 2006).

Lu and Ma (2002) provide a model of P4P programs as a framework to testing the effectiveness of the Maine Addiction Treatment System (MATS), which had adopted performance based compensation in 1995. The model, and subsequent empirical estimation, addresses the question of risk selection under various reimbursement structures. Specifically, physicians are able to treat a patient, or refer to another provider. Two physician remuneration systems are compared, a capitation system, and a capitation system with performance based contracting components. In the original system, providers have the incentive to discriminate against patients, who are costly to treat, which can be accomplished via dumping patients, offering expensive patients less desirable care, and limiting the amount of care to severely ill patients. Performance based contracting gives the incentive to create a better match between illness severity of patient, and the intensity of treatment offered. Since each treatment is offered by a different provider, each

provider has incentive to refer appropriately, and to treat appropriately. Providers are also able to game the system by selectively serving patients who are more likely to respond positively to treatment, and by reporting better patient outcomes untruthfully.

4.2. Empirical Findings

For all the attention devoted to QBIP programs, program effects in terms of provider behaviour, provider and patient satisfaction, or patient health outcomes have not been tested systematically through formal evaluations. Evaluations of P4P programs are few and far between in the health services literature, and often focused on small scale interventions, not broad programs (Rosenthal *et al.*, 2004). More formal evaluations are part of the CMS set of P4P programs and in a few years, one will be able to test more rigorously the effects of the performance-related remuneration schemes. For the moment, however, effects reported in the current literature are often case study-based observations that may or may not apply on a more global scale and may lack proper control reference groups. Some national level studies are available in the U.K.

Gravelle *et al.* (2006) use Scottish data on physician practice characteristics, patient characteristics, and disease characteristics in the population to test two questions. First, which factors influence practice QOF performance, and second, whether new incentives lead physicians to game the system. Results indicate that performance targets were reached in the first year far above expectation. On average, over 90% of practices achieved maximum thresholds, far above the projected 75%. Exception reporting in the first year ranged from 0.1 to 27.7 percent, for various indicators, and various practices. Reported prevalence of patients increased in all domains, though data does not suggest that the population became sicker.

Quality of care is positively related to the proportion of female GPs within a clinic, younger GPs, the proportion of GPs born outside the U.K., and the physician to patient ratio. Quality is lower in practices with PMS contracts, and in practices with non-competitive PbR models. (Gravelle *et al.*, 2006)

Gravelle *et al.* (2006) discover gaming of the system in the form of untruthful exception reporting. Exception reporting is positively related to several patient characteristics, such as deprivation status, ethnic minority status, and rurality. However, practice characteristics also influence exception reporting, and should not, in the absence of gaming. Positive determinants of exception reporting include proportion of foreign trained GPs, holding a PMS contract, having been a fundholding practice, being exposed to less competition from other practices, and being below minimum thresholds for targeted indicators.

The question of physicians gaming the system is further investigated by Doran *et al.* (2006). They analyze data from 8105 family practices in the United Kingdom in the first year of the PbR system, data from the U.K. census, and data on practice characteristics. It is shown that performance was moderately, but significantly, influenced by patient characteristics and practice characteristics. Exception reporting, which was not extensive, turned out to be the strongest predictor of high performance (Doran *et al.*, 2006). In a world of no gaming, practice characteristics and exception reporting should have little influence on performance, save for sound clinical reasons.

Lu and Ma (2002) test the theoretical model described in the previous section using data on 18972 patients in the Maine Addiction Treatment System. The match between patient condition and treatment intensity, as well as referrals and outcomes data are compared in two time periods, before the P4P program, and after it's implementation. They find that the P4P program is positively associated with a higher number of referrals. The referrals were appropriate, since the P4P program seems to have improved the matching of patients to treatment. Some patient dumping was discovered, but was not statistically significant. (Lu and Ma, 2002)

In terms of health system output measures, it was observed that during the first two years of a P4P program for gynecologists and obstetricians, the rate of C-sections declined by 2%, length of stay decreased by 85%, the use of endometrial biopsy increased to over 85%, and the pre-certification rate rose to over 85% (Hanchack, 1997). Assessment of changes in health outcomes resulting from changing health care outputs are difficult, since outcome measures depend as much on patients' co-morbidities and compliance with treatment regimens as they do on the doctors skill or effort (Herndon, 2005).

Petersen, *et al.* (2006) provide a systematic literature review to evaluate empirical evidence of the relationship between explicit financial incentives designed to improve health care quality, and quantitative measures of health care quality. In total, 17 studies were examined. 12 found partial or positive effects of P4P on measures of quality, generally relating to preventive care. Four found unintended consequences, though not all investigated them.

Rosenthal *et al.* (2005) investigate the impact of the PacifiCare P4P program on quality of care, as measured via three process indicators: cervical cancer screening, mammography, and hemoglobin A_{1c} testing. In total, the plan paid \$3.4 million in bonus payments in one year. For all three measures, physician groups with high baseline performance at or above threshold improved the least, yet collected the highest share of rewards. Results indicate that paying clinicians to reach a common goal, a fixed performance target may produce little gain in quality given the cost of the program. Rewards are channeled primarily to physicians with higher performance at baseline.

Qualitative interviews with 36 primary health care professionals revealed that the PbR system led to major changes in clinical practice, due to financial incentives, and due to other considerations, including professional drive, desire to maintain autonomy, and peer pressure. These results support the notion that financial incentives are but one determinant of performance. A crucial enabler to quality improvement is the alignment of managerial leadership, professional values, and financial incentives. (Spooner *et al.*, 2001)

Public disclosure on provider quality changed provider behaviour more so than patient behaviour. Only PacifiCare of California seems to have affected patient choice by releasing information on quality of physician groups. Patient plan designs encouraging choice of high quality providers have more impact in this respect, but are mostly still in the planning or development stages. Paying providers differently based on performance seems to motivate quality improvements (MedPac, 2003).

The NHS system has been subjected to assessment by the UK Audit Commission in its report "Early Lessons from Payment by Results" (2005). The report highlights several interesting findings. First, the amount of administrative complication and burden during the first year after implementation of the PbR system was greater than foreseen. The system proved to be complex, time consuming and challenging, as well as expensive (with an estimate of £50 million nationally). The energy devoted to planning and reporting arrangements, negotiation, and attention to detail might have detracted from focus on actual quality improvements.

The NHS has also experimented with bonus payments at the specialist level. U.K. hospitals employ contracts with consultant surgeons in various specialties, such as general surgery, urology, trauma and orthopaedics, ENT, and ophthalmology. Consultant surgeons can be hired on a full time basis, where their private practice income is limited to 10% of their NHS income; or they can be employed on a maximum part time basis, meaning that the specialist is paid ten elevenths of full time salary, but faces no limit on amount earned through private practice. Under both types of arrangements, consultant surgeons are offered bonus payments based on earning discretionary points and distinction awards. These are not earned directly through clinical activity volume, but instead might be based on teaching excellence, hence might increase or decrease clinical activity. Variation in

clinical activity between surgeons is substantial. Surgeons under 40 years, and above 60 years are the least active. Some variation is explained by region of practice. Maximum part time contract holders have substantially higher activity rates than full time contract holders. Bonus payments are associated with higher activity rates (Bloor *et al.*, 2004). It is discovered that a reduction in private sector activity does not necessarily lead to an increase in public sector activity, therefore incentives should not be aimed at the former. Rewards and bonuses should be based on NHS activity, not private activity. Currently, maximum part time contract holders are not eligible for bonus payments.

While a sufficient number of formal evaluations is not available, observation suggests that P4P programs might be a necessary but not sufficient condition to health care quality improvement. These should instead become part of a broader strategy of promoting health care quality through measuring and reporting performance, providing technical assistance and evidence-based guidelines, and giving consumers incentives to choose high quality providers (Rosenthal *et al.*, 2004).

5. Lessons Learned and Recommendations

Despite efforts to foresee and mitigate problems, there remain challenges related to funding, ease of administration, data collection, reporting of actionable information to providers, identifying outcomes, and savings results, provider acceptance, and transparency of results for consumers (Baker *et al.*, 2003). Lessons learned from QBIP programs to date are categorized into those related to measures, data collection, program structure, and implementation. Lessons learned from the QBIP programs are related to the reality of implementing similar programs in Manitoba.

5.1. Quality Measures

Aside from being valid, reliable and reflective of underlying concepts of quality, measures must satisfy a set of practical and logistical criteria, so as to be operationalized. Measures need to reflect program goals, and must be relevant to the program's major constituents, such as health plans, purchasers (employers), and providers. Measures also need to focus on outcomes that are in the providers' control. For acceptance, understanding, relevance, measures must be specialty specific and developed by the physician community. In addition, measures should be reviewed and evaluated periodically. (Baker *et al.*, 2003, Herndon, 2005, Milgate *et al.*, 2006).

Most importantly, measures need to be related to population health goals and established health objectives. They need to be aligned with overall planning and direction of the health system, and also have the capacity to allow for regional variation. As such, there needs to be some flexibility required to their implementation.

RECOMMENDATION 1

Quality indicators selected should be reflective of population health needs in the region. They should also be relevant to the physician, and under her/his control.

The PbR system has been criticized in terms of the determination of weights assigned to various indicators within each clinical domain. Indicator weights are based on workload estimates of performing specified procedures, but not on potential health gains from procedures, nor population health needs. Fleetcroft *et al.* (2006) examine a subset of eight preventive interventions (38 of the 76 clinical indicators in the QOF). Potential health gain from each procedure is compared to the weight assigned to the indicator in the QOF. It turns out that payments in the PbR contract do not reflect likely population health gains. The system skews physician behaviour toward activities with high workloads, which do not result in maximum potential health benefits. (Fleetcroft *et al.*, 2006)

RECOMMENDATION 2

Weights for indicators within a composite index should be based on degree of need in terms of population health, or based on potential cost savings, but preferably not on workload or other cost estimates.

5.2. Data Collection

From the perspective of the Manitoba Health Care system, perhaps the most important lesson is the following. Prior to program design and implementation, it is crucial to establish baseline information. The PbR system in the U.K. was overwhelmed with the number of providers achieving targets, and costs to the system were higher than anticipated. The reason was a lack of baseline information, which would have revealed that many providers were already at, or not far away from, established targets. In the end, the large expenditure did not result in an equally large quality improvement. Similar observations were made in the U.S. Providers with high baseline performance reaped the greatest financial rewards, while contributing the least to quality improvement (when payment was linked to target, not improvement). To effectively implement this programme, it would be important for Manitoba to collect accurate baseline information prior to the setting of targets, as otherwise targets might be aimed too low, or too high.

RECOMMENDATION 3

Baseline information on current quality measures needs to be collected, in order to align targets with gaps in service provision.

A major concern with data collection is that providers need to be able to actually perform this task in addition to their clinical care responsibilities. Data collection must be reliable, standardized, easy for physicians to record and report, based on a clinical data set and in a manner acceptable to the physician community, and cannot create a burden on practices. In the U.S., it must also be in compliance with requirements under the Health Insurance Portability and Accountability Act (Herndon, 2005). QBIP programs cannot be overly complicated or expensive to administer. It would be advisable for them to be well structured and simplified.

While most programs require data collection and reporting, such as for instance the Medicare Advantage Plan, a series of difficulties arise in this arena. The large number of programs with varying guidelines makes it difficult to maintain administrative records to verify compliance with guidelines, performance achievements, and to compare among different programs. Individual providers face a challenge when they admit patients from more than one insurer, as various purchasers may have different performance targets. Conciliation of goals of various purchasers is not always possible. Furthermore, IT requirements such as electronic records or computerized prescriptions are large capital investments that might not be affordable for small physician groups or solo practitioners (Herndon, 2005, Milgate *et al.*, 2006).

RECOMMENDATION 4

Providers require financial and personalized support in implementation and operation of new IT systems, and electronic records.

5.3. Program Structure

An important decision to make in QBIP program design is the degree of competitiveness of the program. The literature is not conclusive in terms of a recommendation. Competitive programs reward high performers, while penalizing low performers, where performance is evaluated relatively to all other providers. Competitive programs are advocated as more likely to result in improvement, since a stationary threshold target is easily mis-specified (see above). Competitive programs can also potentially be budget neutral. On the other hand, competitively structured P4P programs redistribute income from low quality to high quality providers, and might have the

beneficial effect in the marketplace of weeding out low quality providers.⁹ Yet, low quality providers are arguably the ones in most need of resources to improve quality (Rosenthal *et al.*, 2004). This may then create internal management issues, particularly if quality is not measured properly and a low quality rating results from the patients' lower than average morbidity.

Proponents of non-competitive programs (Herndon, 2005, Milgate *et al.*, 2006) argue that any type of system that rewards providers by improving patient care and outcomes should not be subject to budget neutrality or used as a physician volume control. QBIP programs must not be punitive. When designing QBIP programs, focus should be given to rewarding providers who achieve benchmarks, and giving incentives to improve performance. Funding should be set aside, so that QBIP payments are made from a set aside additional budget. Arguments are not substantiated with empirical evidence by the authors.

Rosenthal (2004) identifies several points of caution, given her synthesized description of P4P programs. First, the majority of plans reimburse based on good performance, but do not based on improvement. While this rewards past investments in quality and high performance, it is not fully consistent with the goal of quality improvement. Second, competitive P4P models, which dominate the U.S. market with a small majority, redistribute reimbursement from low quality to high quality providers. It might be argued that low quality providers are in most need of support. Third, the dimensions of care that receive most attention tend to be those easily measured, and not those most valued or needed. And lastly, one must be careful not to create a situation, where targeted interventions become the primary focus of care delivery, while other types of care is neglected. This is particularly troubling in situations, where integrated holistic care is appropriate. (Rosenthal, 2004)

RECOMMENDATION 5

QBIP program designers must carefully weigh pros and cons of competitive versus non-competitive programs.

RECOMMENDATION 6

Payments should be tied directly to improvement in quality measures, not solely to absolute achievement of a fixed target.

5.4. Program Implementation

For QBIP programs to operate effectively, they need to be accepted and supported by those affected by them. To successfully implement a program, leadership and cooperation from medical staff is critical. Implementation of QBIP programs should be phased in and pilot tested on a voluntary basis first (Herndon, 2005, O'Hare, 2005).

To increase provider acceptance and uptake, P4P programs should be accompanied by care delivery protocols and clinical pathways. Additionally, desirable features to ensure acceptance include programs being clinically relevant, offering meaningful incentives, providing sufficient patient volume, and using measures that do not create administrative burden (Baker *et al.*, 2003, O'Hare, 2005).

Evidence also suggests that the most acceptable performance management systems are those which are developed collaboratively among government, physicians, health authorities and institutions and which have been the object of discussion, deliberation and agreement

⁹ If P4P is budget neutral (as in the case of Medicare), then low performers are penalized. (Herndon, 2005)

RECOMMENDATION 7

Proposed measures, targets, and program structures must be developed collaboratively with physicians, health authorities, and health care institutions.

6. Conclusions

Quality Based Incentive Payment programs have developed in several health care systems within the context of comprehensive strategies for quality improvement in health care, including reforms in the NHS, improvements to CMS, changes in management style in Kaiser Permanente. Implementation of a QBIP system in isolation may not be effective. Primary Health Care Renewal strategies across Canada, including the Manitoba Physician Integrated Network create a good context for establishment of QBIP programs.

Such implementation, however, should not proceed hastily. Experiences in other jurisdictions reveal that a gradual approach is preferable in terms of ability to improve quality of health care – which is the ultimate goal. First, the design of the incentive structure proved more challenging than anticipated, since each system has created some potential for physicians to game the system. Second, weights assigned to various quality indicators need to be rooted in population health needs and gaps in service provision, as informed by baseline data. Third, the lack of baseline data in the U.K., coupled with higher than expected achievements, has raised suspicion that target levels were set too low. The result was a high cost to the system and improvements in quality below potential.

In addition, pay for performance systems require organizational support and change to a culture in which health objectives are clearly established and wherein measures are presented in the context of an evidence-based framework. Effective managerial leadership is required at multiple levels to ensure understanding, implementation, and the utilization and application of information on performance. Organizational culture seems to be changing across Canada, as exemplified by the several new models of primary care: Family Health Networks, Family Health Groups, and Chronic Care Models in Ontario, Groupes de Médecine Familiale in Quebec, as well as some specialist models, such as the SEAMO, with an emphasis on delivering better quality health care within some context of performance measurement.

Availability and quality of data have been emphasized and improved in Canada, as well. First, the creation of CIHI has allowed an increasing amount of reliable information to be collected and analyzed. Furthermore, there have been attempts at analyzing the data to measure quality and performance in many provinces (Ontario, Québec, Manitoba, Nova Scotia; see also Ontario Hospitals reports, Katz *et al.* (2004) study paper on quality indices). There have been several new models of primary care

References

- Ashworth, M., Armstrong, D., 2006, The relationship between general practice characteristics and quality of care: a national survey of quality indicators used in the UK Quality and Outcomes Framework, 2004-5, *BioMed Central Family Practice*, 7:68-75
- Baker, G., Houghton, J., Mongroo, P., 2003, Pay for Performance Incentive Programs in Healthcare: Market Dynamics and Business Process, downloaded February 14, 2007 from http://www.leapfroggroup.org/media/file/Leapfrog-Pay_for_Performance_Briefing.pdf
- Bloor, K., Maynard, A., Freemantle, N., 2004, Variation in activity rates of consultant surgeons and the influence of reward structures in the English NHS, *Journal of Health Services Research and Policy*, 9(2):76-84
- Bodenheimer, T., May J.H., Berenson R.A., Coughlan J., 2005, Can money buy quality? Physician response to pay for performance, *Issue Brief Cent Stud Health Syst Change*, 102:1-4
- Bonis, P.A., 2005, Quality incentive payment systems: promise and problems, *Clinical Gastroenterology*, 39(4 Suppl 2):S176-82
- Chassin, M.R., Galvin, R.W., 1998, The Urgent Need to Improve Health Care Quality, *National Roundtable of Health Care Quality, JAMA*, 280:1000-1005
- Curry, T., 2005, Payment by results, *Nursing Management*, 12(3):16-9
- Department of Health, 2002, Reforming NHS Financial Flows: Introducing payment by results, downloaded February 14, 2007 from <http://www.dh.gov.uk/assetRoot/04/06/04/76/04060476.pdf>
- Doran, T., Fullwood, C., Gravelle, H., Reeves, D., Kontopantelis, E., Hiroeh, U., Roland, M., 2006, The New England Journal of Medicine, 355(4):375-384
- Ettner, S.L., Thompson, T.J., Stevens, M.R., Mangione, C.M., Kim, C., Steers, W.N., Goewey, J., Brown, A.F., Chung, R.S., Narayan, K.M.V., and the TRIAD Study Group, 2006, Are Physician Reimbursement Strategies Associated with Processes of Care and Patient Satisfaction for Patients with Diabetes in Managed Care?, *Health Services Research*, 41(4), Part 1:1221-1241
- Feasby, T.E., Gerdes, C., 2006, Pay-for-Performance – Can It Work in Canada?, *Healthcare Papers*, 6(4):47-50
- Fleetcroft, R., Cookson, R., 2005, Do the Incentive Payments in the New NHS Contract for Primary Care Reflect Likely Population Health Gains?, *CHE research Paper 3*
- Golden, B. R., 2006, The Cost of Pay-for-Performance in Healthcare: An Alternative View, *Healthcare Papers*, 6(4):40-6
- Gravelle, H., Sutton, M., Ma, A., 2006, Paying for quality : British general practitioners' performance under the quality and outcomes framework, downloaded from <http://www.nhh.no/sam/conf/HealthProgramme-nov2006.html>
- Guthrie, B., 2003, Quality measurement of care for people with type 2 diabetes in Tayside, Scotland: implications for the new UK general practice contract, *British Journal of General Practice*, 53:709-713
- Guthrie, B., McLean, G., Sutton, M., 2006, Workload and reward in the Quality and Outcomes Framework of the 2004 general practice contract, *British Journal of General Practice*, 56:836-841
- Hanchak, N.A., 1997, A Performance-Based Compensation Model for Obstetricians/ Gynecologists, *Clinical Obstetrics & Gynecology Amniotic Fluid*, 40(2):437-445
- Herndon, J.H., 2005, Pay for performance: Coming your way soon. *The American Academy of Orthopaedic Surgeons, Bulletin*, www.aaos.org/wordhtml/bulletin/aug05/cover.htm
- Landon, B.E., 2006, Is Pay-for-Performance Moving North? P4P Prospects in the Canadian Healthcare System, *Healthcare Papers*, 6(4): 28-33
- Leger, P.T., 2000, [Quality Control Mechanisms under Capitation Payment for Medical Services](#), *Canadian Journal of Economics*, 33(2):564-86

- Lindenauer, P.K., Remus, D., Roman, S., Rothberg, M.B., Benjamin, E.M., Ma, A., Bratzler, D.W., 2007, Public Reporting and Pay for Performance in Hospital Quality Improvement, *The New England Journal of Medicine*, 356(5):486-496
- Lix, L. Yogendran, M., Burchill, C., Metge, C., McKeen, N., Moore, D., Bond, R., 2006, Defining and Validating Chronic Diseases: And Administrative Data Approach, Manitoba Center for Health Policy, Faculty of Medicine, University of Manitoba, Winnipeg, MB
- Lu, M., Ma, C.A., Yuan, L., 2002, Risk selection and matching in performance-based contracting, *Health Economics*, 12: 339-354
- McLean, G., Sutton, M., Guthrie, B., 2006, Deprivation and quality of primary care services: evidence for persistence of the inverse care law from the UK Quality and Outcomes Framework, *Journal of Epidemiology and Community Health*, 60:917-922
- McNamara, P., 2005, Quality-based payments: six case examples, *International Journal for Quality in Health Care*, 17(4):357-362
- Milgate, K., Cheng, S.B., 2006, Pay-For-Performance: The MedPAC Perspective. *Health Affairs*;25(2):413-9
- National Health Service in England. 2005. Early lessons from payment by results. Audit Commission. Retrieved from www.audit-commission.gov.uk in August 2006
- O'Hare, P.K., 2005, Pay for performance: will your hospital be ready? *Healthcare Finance and Management*, 59(5):46-8
- Pink, G.H., Brown, A.D., Studer, M.L., Reiter, K.L., Leatt, P., 2006, Pay-for-Performance in Publicly Financed Healthcare: Some International Experience and Considerations for Canada, *Healthcare Papers*, 6(4):8-26
- Roland, M., 2004, Linking Physicians' Pay to the Quality of Care – A Major Experiment in the United Kingdom, *The New England Journal of Medicine*, 351(14):1448-1454
- Rosenthal, M.B., Fernandopulle, Song, H.R., Landon, B., 2004, Paying For Quality: Providers' Incentives For Quality Improvement, 23(2): 127-141
- Rosenthal, M.B., Landon, B.E., Normand, S.-L.T., Frank, R.G., Epstein, A.M., 2006, Pay for Performance in Commercial HMOs, *The New England Journal of Medicine*, 355(18):1895-1902
- Rowland, D. 2004. Linking physicians' pay to the quality of care – a major experiment in the United Kingdom. *The New England Journal of Medicine* 351(14): 1448-1454.
- Shea, M.A., 2007, Pay for performance's small steps of progress., PricewaterhouseCoopers, downloaded February 14, 2007 from <http://pwchealth.com/pdf/p4p>.
- Smith, P.C., York, N., 2004, Quality incentives: the case of U.K. general practitioners, *Health Affairs*, 23(3):112-8
- Street, A., Miraldo, M., 2006, Reform of hospital funding and the introduction of patient choice in England, downloaded on February 14, 2007 from <http://www.nhh.no/sam/conf/Paper%20Health/AndrewStreet.pdf>
- Strunk, B.C., Hurley, R.E., 2004, Paying for quality: health plans try carrots instead of sticks, *Issue Brief Cent Stud Health Syst Change*, 82:1-4
- Vounasis, F., Dubinsky, I.L., 2005, Are Physician Executive Pat-for-Performance Programs the Future of Physician Leader Compensation in Canada?, *Healthcare Quarterly*, 8(4):86-90
- Wranik, D., 2007, Quality Based Incentive Payments – Canada's Health Care Solution, Working Paper, School of Public Administration, Dalhousie University, Halifax, NS