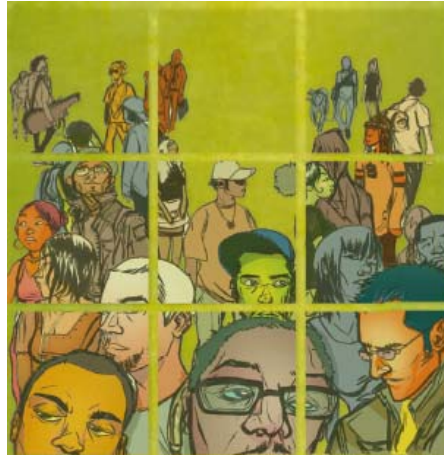




Performance incentives for health in high-income countries key issues and lessons learned

Riku Elovainio

**World Health Report (2010)
Background Paper, 32**



© World Health Organization, 2010

All rights reserved.

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted lines on maps represent approximate border lines for which there may not yet be full agreement. The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by the World Health Organization in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters. All reasonable precautions have been taken by the World Health Organization to verify the information contained in this publication. However, the published material is being distributed without warranty of any kind, either expressed or implied. The responsibility for the interpretation and use of the material lies with the reader. In no event shall the World Health Organization be liable for damages arising from its use. The findings, interpretations and conclusions expressed in this paper are entirely those of the author and should not be attributed in any manner whatsoever to the World Health Organization.



Performance incentives for health in high-income countries – key issues and lessons learned

World Health Report (2010) Background Paper, No 32

Riku Elovainio¹

¹ World Health Organization, Geneva, Switzerland

The author would like to thank Jean Perrot for comments and suggestions and Gary Humphreys for editorial support.

1 Introduction

During the last decade several new types of health provider performance incentive schemes have emerged under designations such as "Pay-for-Performance (P4P)", "Results Based Financing", "Performance Based Financing" or "Performance Based Contracting". First adopted in high-income countries, these schemes have received growing attention in low- and middle-income countries where performance incentives are advocated as a means to improving key health indicators and increase access to care.

The schemes are highly diverse. Indeed, while it is true that there is a unifying rationale behind them, i.e. the financial reward or sanction of specific behaviours and results, there are also significant variations in their logic and in their implementation mechanisms. This diversity is mainly related to the fact that performance incentives are used to solve different problems in different contexts.

It could be argued that performance incentives are an integral part of a health system architecture - in all contexts. Chaix-Couturier et al. (2000), for example, stress the fact that a wide variety of different incentive methods have been used in health systems in the past. According to the authors these incentive measures include: capitation; managed care; sanctions/bonuses according to generic drug prescriptions; incentives for early retirement of health workers; fund holding; etc. These measures are all linked to more or less comprehensive policy choices that modify the incentivizing elements in health systems; for example, choosing one type of provider payment method over another also implies a choice of one type of incentive structure over another. From this perspective, it is clear that health service providers are given incentives derived from the general health system architecture and from the diverse policies and strategies which are destined, directly or indirectly, to influence their behaviour.

The performance incentives discussed in this document are not linked to these "structural" incentives. Rather, we are focusing on a wide variety of distinct interventions/schemes which directly target providers, as individuals or as organizations, and which establish a link between provider remuneration and a set of predefined performance measures. The distinction made is thus between the "system" incentives, or the "organizational incentive regime" (WHO, 2000), which operate mainly on a macro scale and the "explicit"

performance incentives which are often micro level in scope and are linked to a bottom-up approach to influencing provider behaviour and results.

The objective of this background paper is to study the provider performance incentive schemes of high-income countries with the specific purpose of identifying lessons learned that could feed into the discussion on using provider performance incentives in different contexts - including in low-income country settings. These lessons learned cover the key issues of implementation architecture and strategy, and they are often related to the "moving parts" which can define the successes and failures of the schemes.

It should be clear from the outset that it is impossible to provide hard evidence on provider performance incentives as such. The available literature concerns often only singular interventions (often small scale pilot projects) that are based on context-specific arrangements. We are still far from possessing a critical mass of evidence that would allow us to draw any solid conclusions on the usefulness of provider performance incentives in general. Moreover, there are few stable and invariable elements allowing comparison between the different performance incentive schemes. New schemes are implemented in different contexts, by different "sponsors", with different objectives and under different nomenclatures². Similarly, as new evidence is gathered, the subject of study has already moved on (new ways of "doing" provider performance incentives come up all the time), rendering the earlier evidence, at least partially, irrelevant. Any type of holistic view is thus problematic. It for this reason that we here focus on some key implementation issues and build evidence case by case.

We shall include in our analysis only supply side incentives that target health care providers : hospitals, clinics, health centres, practitioners or networks. This approach puts performance incentive schemes targeted at the administrative level outside the scope of the study. It also means that we are not looking at the different types of demand side incentives such as Conditional Cash Transfers for health service utilization or incentives to encourage patient compliance.

² That said, there are also several counter forces that are destined to increase harmonization in the use of performance incentives. In the USA, for example, P4P schemes are often implemented through facilitating organizations that engineer a P4P "package" which can then be used by several P4P sponsors. The World Bank based Health Results Innovation Trust Fund (for Result Based Financing) that focuses on performance incentive schemes in low-income countries also aims to provide a larger institutional framework for the differently implemented schemes.

2 Provider performance incentives in the high-income countries - an overview

2.1 The general picture

The literature on provider performance incentives in high-income countries is divided between a fairly substantial volume of literature on schemes in the United Kingdom and the United States of America (USA), and studies of schemes that have either just started or are in a planning phase, in countries like France, Italy, Estonia and Spain. The former tends to be descriptive but includes some evaluations of the experiences; the latter is limited to description only. There are thus in general few rigorous evaluations of provider performance incentive schemes in high-income countries; and overall, limited documentation covering high-income country schemes outside the United Kingdom and the USA.

There are several reasons to think that performance incentives might actually have a bigger role in high-income countries than could be concluded from the volume of relevant literature, one of the obvious being that such schemes are simply left undocumented. Indeed, most high-income countries have several autonomous institutions that can act as purchasers and decision-making bodies : municipalities, counties, autonomous hospitals, private insurances, etc. These institutions have the possibility, capacity and financial leverage to implement performance incentive schemes at the operational level. This will necessarily happen on a small scale, through the (contractual) relationships these institutions enter into with the health service providers. In these instances there is seldom a documented case study of the "scheme".

The situation is quite different in low-income countries where performance incentive schemes often take the form of pilot initiatives that have been more or less thoroughly documented and evaluated for the dual purpose of financier/donor accountability and in order to explore the possibilities for scaling up. In high-income countries these considerations rarely pertain, which means that a lot of the initiatives implemented at a micro level are done with a "business as usual" rationale. There is however some evidence on the use of performance incentives in these micro contexts in the high-income countries : in Portugal, for example, some autonomous hospitals have introduced performance based pay for their staff (Barros and Simões, 2007); in the Netherlands private insurance companies have started to introduce performance related bonus payments to general practitioners (Wynand et al., 2009); and in Sweden the contracts between county councils

and private care providers can include performance related payment methods (Dagens Nyheter, 2009). However in order to get a good picture of these "micro schemes", one would need to go through a thorough research of material such as municipal council decisions on contracting with private providers. Research of this kind was beyond the scope of this study.

A second reason for the low apparent prevalence of performance incentive schemes in high-income countries could be the fact that the focus in these contexts is on results-based payment methods within larger multisectoral policy reform.

Indeed, in several high-income countries, overarching policy changes have recently taken place in the general labour remuneration rationale. This is true for the private sector but also for civil servants. Pay for performance policy for civil servants has been adopted, in one form or another, in two-thirds of OECD countries (OECD, 2005). This general policy has subsequently been broken down into sector-specific approaches. For example, in the United Kingdom there is a long history of performance related pay in the public sector that has trickled down from the central government to the sectoral level. However, it seems that in the British case a central government policy has had a very limited impact on health workers (Arowsmith et al., 2001). In Finland there is also a general policy of result based pay for public sector workers - established through collective bargaining between the public sector workers' unions and the public sector employers' organizations. The policy has been introduced into the health sector through a sectoral Collective Agreement where the principle of result based pay, called the "personal part of remuneration" is included (Commission of Local Authority Employers, 2009). However, the Collective Agreement does not specify any indicators or methods to determine personal performance; it is left up to each organization to establish the rules in local negotiations between the management, the staff and the union representatives.

2.2 Provider performance incentives in the U.S. health system - P4P

The fragmented U.S. health system has a limited number of elements that create policy coherence; individual actors (purchasers, administrators, etc.) have considerable elbow room to introduce new initiatives. This contextual particularity has led to an interesting situation where there has been a global impetus to introduce provider performance incentives which nonetheless take different forms around the country.

Provider performance incentive schemes, referred to as Pay-for-Performance (P4P) initiatives³ in the U.S. context, have existed since the late 1980's. However it is only at the beginning of this century that the exponential growth of such schemes took place. One of the decisive moments behind the development of the P4P "movement" was the 2001 publication of the "Crossing the Quality Chasm" report from the Institute of Medicine (IOM, 2001). This report has often been identified in the literature as the major catalyst for P4P initiatives in the U.S. (Peteresen et al., 2006; Rosenthal et al. 2004; Christianson et al., 2006; Berwick, 2002). The report identified six challenges for the U.S. health system : patient safety, effectiveness, patient-centeredness, timeliness, efficiency and equity. One of the key conclusions of the report was the notion that provider payment mechanisms were one of the most promising entry points for effective health care reform. The report recommends that policy makers:

"Align financial incentives with the implementation of care processes based on best practices and the achievement of better patient outcomes. Substantial improvements in quality are most likely to be obtained when providers are highly motivated and rewarded for carefully designing and fine-tuning care processes to achieve increasingly higher levels of safety, effectiveness, patient-centeredness, timeliness, efficiency, and equity." (IOM, 2001:184).

The IOM report is not the only piece in the P4P puzzle but it has surely been an important factor in the rapid diffusion of P4P in the whole U.S. health system; an evolution that has been described as "arguably the most striking change in the U.S. health care system since the managed care era." (Rosenthal et al., 2007).

The general P4P movement has been translated into a wide variety of interventions and it is estimated that there are in total more than 200 schemes implemented⁴. The sponsoring institutions of these schemes are mainly commercial health insurance companies or health maintenance organizations but publicly funded schemes, administered through Centres for Medicaid and Medicare Services (CMS), are growing in importance.

3 In the US context the provider performance incentives are usually referred to as Pay-for-Performance initiatives. We shall therefore use this term from now on when referring to the provider performance incentives in the US context.

4 This is an estimation based on available survey results from the Leapfrog Group <http://www.leapfroggroup.org/> and Med Vantage® <http://www.medvantage.com/>; it has also been estimated that half of the commercial health maintenance organizations use P4P incentives (Rosenthal et al., 2006)

A survey of twenty-seven well documented cases revealed that most of the P4P schemes targeted primary care physicians (PCP), around 90% also targeted specialists (with cardiologists and general surgeons topping the list) and 80% targeted a hospital (Rosenthal et al., 2007). In general, all the larger P4P schemes included all three components.

There has been a marked increase in the use of outcome measures : in 2003 only 59% of the studied schemes used them but by 2006 94% of the schemes included them. The most commonly included outcome indicators were related to cholesterol levels or blood pressure ("outcome" in this context refers to an intermediate outcome). All schemes included process indicators, the most cited being compliance with asthma and diabetes care guidelines.

Almost all of the surveyed schemes increased the number of indicators over time, but there were also indicators that were dropped. For example, patient satisfaction indicators were dropped in 17% of the schemes during the four year survey period; the lack of variation between provider results and the cost of collecting the data were given as reasons for dropping the indicator. Also, there were several indicators that seemed to be consistently "too high", this category of indicators included, for example, MMR vaccination or tobacco cessation counselling.

In the state of California several P4P schemes have been implemented by health care financiers (employers, insurers, etc.). A large proportion of these schemes are implemented through a joint-venture organization - Integrated Health Care Association (IHA) - that has established a common measurement set and payment modalities for most of the P4P schemes in the state. The IHA methodology establishes a mechanism where about 50% of the indicators are linked to clinical domains, 40-30% to patient experience and 10-20% to IT development⁵.

The P4P schemes seem to have relatively little 'heft' in monetary terms. According to one survey, the average value of the total bonuses paid to the providers amounted to just US\$1.4 per insured member per month (ranging from twenty cents to fifteen dollars); this represents just 2.3% of the average reimbursements to the providers (Rosenthal et al., 2004). The IHA payments for physician groups in California in 2008 totalled US\$52 million, representing an average 1% of the total yearly income of the physician groups.

5 http://www.iha.org/financial_transparency.html (consulted 15 November 2009)

The P4P movement in USA has been reshaping itself in the last couple of years. There has been a push towards a more comprehensive approach where the performance related payments are included in a wider provider payment reform (Rosenthal, 2008). These reforms can include, for example, the introduction of a capitation payment that is modulated by a performance component, or a savings sharing scheme for fee-for-service beneficiaries, where integrated physician groups can earn bonuses for demonstrating slower growth in spending relative to peers and for increased quality (quality and savings have to happen simultaneously) (Trisolini et al., 2008).

2.2.1 The CMS Premier Hospital Quality Incentive Demonstration (PHQID)

In order to get a better understanding of the different architecture details of the P4P schemes that have been introduced recently, we here present one initiative in more detail : the Centres of Medicare and Medicaid Services (CMS) run Premier Quality Incentive Demonstration (PHQID) - scheme.

The Centers of Medicare and Medicaid Services (CMS) is the biggest health care purchaser in the USA. The CMS administers public funds through which health services are purchased for low income and elderly Americans. In 2007 the CMS spent US\$570 billion on health care⁶. The CMS sponsors a wide variety of P4P initiatives with different scopes and objectives and using several operational methods.

The PHQID initiative targets around 230 hospitals with a specific objective of tracking five clinical areas:

- 1) acute myocardial infarction,
- 2) heart failure,
- 3) coronary artery bypass graft,
- 4) pneumonia,
- 5) hip and knee replacement.

Each of the five clinical areas have their own Composite Quality Score (CQS) and all the CQS from all the clinical areas are then aggregated to give an overall score for each hospital. The score is then used to determine four types of awards:

⁶ <http://www.cms.hhs.gov> (consulted 7 May 2009). These figures will obviously grow with the new Health Care legislation voted in March 2010

1. Attainment award (40% of the total available awards) - this award is paid to hospitals that attain or exceed the aggregate median level performance for all the hospitals; those hospitals that have reached the median level performance or better share this award on a pro rata, per discharge, basis. It is thus a tournament type of award⁷ that divides the hospitals into two groups of equal size - those who get the award and those who get nothing.

2. Top performer award - this incentive payment rewards the top 20% of the hospitals with an extra bonus that comes as a top-up of the attainment award.

3. Improvement award - this incentive payment rewards those hospitals in the top 20% that have registered the largest quality improvements. It is important to note that the award is only destined to the top performers, no improvement award can be paid to the hospitals outside the top 20% performers, even with significant performance improvements.

4. Threshold Penalty - this is a penalty that affects those hospitals that do not score above the 9th or 10th decile threshold in a given clinical area. These hospitals receive a 1% reduction of their Medicare payments if outside the 9th decile threshold and a 2% reduction if outside the 10th decile threshold.

According to an internal monitoring study of the PHQID scheme, composite performance scores increased significantly during the first four years in the participating hospitals.

Table 1 Performance evolution for the PHQID scheme - the Composite Quality Scores of participating hospitals

	<i>Inception</i>	<i>End of Y3</i>	<i>End of Y4</i>
Acute myocardial infarction	87.5	96.1	96.3
Coronary artery bypass grafting	84.8	97.4	98.5
Heart failure	64.5	88.7	92.2
Pneumonia	69.3	90.5	92.6
Hip and knee replacement	84.6	96.9	97.2

Source: Centers for Medicare and Medicaid Services, 2009

The total improvement in average composite quality scores for all the clinical areas over PHQID's first three years was 17.2 percentage points. Between PHQID's third and fourth years,

⁷ The concept of "Tournament" in the P4P literature refers to a mechanism where the incentive payments are based on the ranking of performance of the providers covered by the scheme. There is thus a maximal total amount of bonuses that will be distributed among the providers.

the average increase was 2.2 percentage points. One can thus observe that by the end of year 3 the hospitals were reaching very high scores and that there was little room for substantial improvements. It could be argued that at these levels of performance, the incentive payments would start to lose their impact. However, the scheme has, at least partially, taken this into account by using incentive mechanisms that rely heavily on a competition between the hospitals and on the quality improvements for the top 20% of performers (the Improvement Award), thus incentivizing marginal improvements at the top end.

During the first two years of the PHQID initiative the bonuses paid to the hospitals totalled US\$17.55 million (for two years) and for the fourth year of the initiative, payments amounted to US\$12 million (for one year)⁸. As there were 230 hospitals included in the scheme, the total incentive payments for the fourth year amounted to US\$52 000 per year, per hospital, a very low portion of the total reimbursements that the hospitals receive on average.

However, the PHQID scheme does seem to have had a notable impact. The internal monitoring report shows that significant improvements in performance have been achieved. Moreover, as noted above, given that the amount of incentives paid to the hospitals appears to be quite low, it could be argued that the intervention has been cost-effective from the financier's point of view.

Nevertheless, the monitoring of results tell only one side of the story and there are several caveats to consider. First, the monitoring figures rely only on a 'before and after' methodology, without comparing the results to a control group. Second, the monitoring results give only the performance of the hospitals for the five clinical areas covered by the incentive scheme; no information is given on the unrewarded aspects of the hospitals' activity.

There have been some external evaluations of the PHQID that have tried to address these issues. Lindenauer et al. (2007), for example, compared the PHQID hospitals with hospitals that had voluntarily published their performance data under the Hospital Quality Alliance (HQA) initiative. The study found that after adjustment for differences in baseline performance and other characteristics, the hospitals within the PHQID scheme were outperforming the HQA hospitals by around 3-4 percentage points. The indicators followed by the study were those of the HQA

⁸ <http://www.cms.hhs.gov/HospitalQualityInits/downloads/HospitalPremierFactSheet200907.pdf>

scheme which often coincided with the PHQID indicators but not always; the study does not reveal if the PHQID did better or worse for the indicators not rewarded.

Another study followed the performance evolution of the hospitals only against acute myocardial infarction (AMI) related indicators (Glickman et al., 2007). The study followed not only AMI process and outcome (in-hospital mortality rate) indicators that are rewarded under the PHQID scheme but also 8 other process indicators that have been deemed clinically pertinent for this condition, but which were not rewarded under the PHQID programme. The study compared performances between P4P hospitals and control (non-P4P) hospitals. The study observed that the overall performance score for the AMI indicators for P4P hospitals increased by 7.2 percentage points (from a baseline of 87.0 to 94.2) during the three year period, while the increase was 5.6 percentage points (from 88.0 to 93.6) at control hospitals - the difference of 1.6 percentage points was not however statistically significant. Interestingly, the P4P hospitals outperformed the control hospitals also in most of the indicators not included in the incentive payments.

2.3 Provider performance incentives in the United Kingdom - the Quality and Outcomes Framework (QOF)

As might be expected the differences in the U.S. and British health systems find expression in the P4P landscapes of the two countries. Compared to the USA, where several performance incentive schemes coexist, the United Kingdom has a much more uniform and centralized approach to performance incentives. In the British context the provider performance incentives are encapsulated in the Quality and Outcomes Framework (QOF) initiative which was put in place in 2004. The QOF is a performance incentive mechanism for family practitioners – or general practitioners (GPs). In essence, the QOF rewards GPs based on their adherence to a set of best practices and to a disease management logic.

Several factors influenced the way the QOF was implemented. First, was a 1997 general election pledge made by Labour to increase health spending. One of the concrete objectives was to increase the remuneration of GPs. In addition, in the decade prior to QOF implementation, there was a growing focus on quality of care and the piloting of several quality-improvement initiatives mainly relying on information guidance. There was also a focus on reinforcing the national inspection system (Campbell et al., 2001; Timmins, 2005).

Negotiations between the professional representatives, the British Medical Association (BMA) and the NHS constituted the first step of the QOF process. During the 18 month negotiations the indicators and payment/bonus mechanisms were chosen (with input from academic advisers that were especially involved in fine tuning elements such as the weightings given to the indicators). The negotiation process was lengthy and comprehensive. There was a historical reason for this: in the mid 1980's the NHS tried to introduce a financial incentive scheme called the Good Practice Allowance (GPA) which failed because of strong opposition from the medical profession. The BMA stated that the GPA initiative "was political and provocative, prepared by a policy unit whose main contact seemed to have been with philosophers, privateers and trendy professors" (Roland 2004). In order to avoid similar problems, the NHS relied on extensive negotiation with the objective of getting the health professionals on board from the beginning.

The negotiations resulted in an agreement on a set of indicators that formed the QOF framework which was ready for implementation in 2004. The QOF was designed to cover three areas : (i) clinical care, (ii) practice organization and (iii) patient experience. In later versions, new categories were added. In the original framework, the 76 clinical indicators represented 55% of the total maximum score, and were related to ten different chronic conditions that were chosen according to their importance in terms of burden of disease, such as hypertension, coronary heart disease, diabetes, etc.

In the current version of QOF (the indicators are renegotiated every year) the clinical indicators cover a wider variety of clinical dimensions and include nine new domains: smoking, obesity, Atrial fibrillation, chronic kidney disease, learning disabilities, palliative care, heart failure, dementia and depression. The introduction of new domains has only slightly increased the number of clinical indicators meaning that there are now fewer indicators per clinical domain/condition than in the beginning. The weight of the clinical indicators in the current version has gone up to 65%. The clinical indicators are a mix of process and intermediary outcome indicators.

Alongside the clinical indicators are three other domains:

- The "practice organization" indicators cover five areas: education and training, medicines management, patient communication, practice management and patient records. The weight of the "practice organization" score in the current QOF is 16.75%.

- The "patient experience" component relies on one consultation length indicator and on a patient survey. The patient surveys are administered by the practitioners themselves; the survey does not cover the aspect of patient satisfaction as such, and is focused on items such as patients being able to obtain a consultation in a given time frame. The "patient experience" score represents 14.65% of the total QOF score.
- Finally, the "additional services" component is composed of eight indicators related to cervical screening, child health surveillance, contraceptive services and maternity services. These indicators represent 3.6% of the total score⁹.

The QOF also includes an exclusion mechanism that allows GPs to exclude patients who will not be accounted for when calculating the final score. The exclusion criteria apply for example to patients who fail to attend after repeated invitations or patients with contraindications for the treatments rewarded. This exclusion mechanism was one the key elements to come out of the negotiations between BMA and the NHS.

Most of the evaluation studies seem to indicate that the QOF initiative has reached at least some of its objectives (Campbell et al., 2007; Doran et al., 2008). There has been a notable performance increase for most of the clinical indicators. Asthma and diabetes care related indicators for example have improved significantly since the introduction of QOF. However, QOF seems to have no effect on coronary heart disease indicators which have followed the same trend as before the introduction of QOF (Campbell et al., 2007). According to one study, there were no significant difference in the rates of improvement between clinical indicators for which financial incentives were provided and those for which they were not provided (Campbell et al., 2007).

When QOF was implemented, there was a common belief that the general score for the practitioners would be somewhere near the 75% mark. However, for the first round of QOF the average score was 91%. For year two, after the first round of measurement modifications, the score went up to 96% and for year three (2006/2007) the score settled at 95%. The remuneration of British GPs increased by 21% in real terms between 2003/4 and 2004/5 and by 7% between 2004/5 and 2005/6 (Fujisawa & Lafortune, 2008); most of this increase is due to QOF payments.

⁹ It is interesting to note that the items in the "additional services" domain are pretty much the same as the ones in the low-income settings that are very often the core focus of the performance incentive schemes. This is thus a very obvious example of the fact that contextual factors strongly orient the way performance incentives are used.

A study done after the first year of QOF revealed that the average yearly QOF payment per individual GP was £23 000 (Doran et al., 2006); this figure however does not represent a net income for GPs who had to invest for example in administration in order to meet the targets. But it is clear that British GPs can gain thousands or even tens of thousands of pounds per year through the QOF programme.

There has been widespread criticism of the QOF metrics; for some, the high achievement results showed that the indicators were focused on a set of "low hanging fruit" and that the thresholds were too low (Fleetcroft & Cookson, 2006). Others have expressed concern over the fact that as there were no baseline references, the result might represent an evolution that was already taking place before the QOF was implemented. Finally it has been argued that the results mainly reflect an improvement in reporting and that no real behaviour change occurred.

If any of these arguments merits consideration the resources put into QOF could be considered wasteful, especially given the high overall cost of the QOF programme - about £1 billion per year, or about 15% of all expenditure on primary care.

Some commentators have also argued that the clinical evidence behind the QOF indicators is not solid (Mangin & Toop, 2007), an argument that relies on the fact the QOF is a top-down system that interferes with the clinicians' freedom, creating "a real risk that general practice will lose its ability to deconstruct evidence and apply it critically in a biopsychosocial model".

There have also been studies evaluating the way incentive payments are distributed within the country and looking for underlying patterns in the dispersion of payments between practices. Such studies have shown for example that practices with lower achievements have a higher proportion of low-income patients or patients over 65 years (Doran et al., 2006; Ashworth et al., 2007). The incentive rewards disbursed are also lower for GPs who are 50 years of age or older. However the various regression models used in different studies offer only a weak indication of the reasons for variation in payments allocated to GPs.

2.4 Other incentive schemes in high-income countries

2.4.1 Australia

In Australia, the Practice Incentives Program (PIP) has been implemented by Medicare Australia since 1998 (Scott, 2007). The PIP is destined to accredited GPs. The scheme is composed of 12 different components that target different aspects of clinical and organizational practice. The components are:

1. After Hours Incentive	2. Practice Nurse Incentive (PNI)	3. Quality Prescribing Incentive (QPI)	4. Teaching Incentive
5. Rural loading	6. Cervical Screening Incentive	7. Asthma Incentive	8. Diabetes Incentive
9. Procedural GP Payment	10. Domestic Violence Incentive	11. Aged Care Access Incentive	12. eHealth Incentive

Source: Medicare Australia <http://www.medicareaustralia.gov.au/provider/incentives/pip/index.jsp>

Each of the different components translates into one or several indicators.

For example, the diabetes incentive is composed of three indicators:

- an indicator that is based on the compliance for diabetes reporting and on the existence of a patient recall/reminder system. The GP has to fulfil these criteria in order to get a one-off signing award that depends on the size of the practice
- an indicator that concerns those practices with at least 2% of their patients diagnosed with diabetes mellitus - if at least 20% of these diagnosed patients complete a cycle of care (defined through the Australian standards) the GP gets a AUS\$20 reward per patient
- an indicator that calculates the absolute number of diabetes treatment cycles completed - every completed treatment cycle is awarded AUS\$40.

Some of the indicators are very specific to the Australian context; the Practice Nurse Incentive (PNI) for example is a reward that is related to the employment of practice nurses and/or Aboriginal health workers. The PIP also tries to address the problem of retention of rural health workers in Australia, the PIP system has a Rural Loading component that multiplies PIP awards according to the remoteness of the practice.

Unlike most other performance incentive schemes, the Australian PIP initiative does not include a composite score that would add up all the different components; instead, each component has more or less its own calculus rationale and reporting mechanism. The complexity of the PIP has been well described in the literature; the initiative has been criticized for the large amount of paperwork associated with it and for the constant changes in criteria and indicators (Oldroyd 2003). The complexity of the mechanism makes it costly to the practices: one practice manager summed up that while he receives in average AUS\$80 000 as incentive payments a year, he also has to hire a practice nurse for AUS\$55 000 a year, has to upgrade his IT platform for AUS\$30 000 - \$40 000 and has to apply for a specific accreditation at a cost of AUS\$5000. On the other hand the manager acknowledges that "the PIP payment definitely allows us the opportunity to afford to employ a practice nurse or keep the IT up to date".¹⁰

2.4.2 Spain

In the autonomous region of Catalonia, Spain, the local health administration finances and regulates public and private providers through contracting. Quality of care is controlled through a set of monitoring indicators that are defined in the contracts. If the providers do not achieve the targets set, they face a deduction in the payments from the Catalan local health administration (Benavent et al., 2009). This (partially) performance based payment system has sparked some management reorganization initiatives at the provider level. The Catalan Institute of Health (ICS), which is the largest public provider of primary and hospital care services, has implemented a Management by Objectives (MBO) policy which covers physicians, nurses and managers. Physicians and nurses are remunerated according to indicators relating to accessibility, quality of clinical care, respect of standards in drug prescription, management of sick-leave in the population, budget execution and user satisfaction. Manager remuneration is also tied to these indicators but in addition it is determined by performance against indicators of professional quality of life (QPL) among the staff managed; this indicator is measured by a voluntary questionnaire survey among the staff.

An evaluation study concluded that after a year of implementation the global score on the QPL composite measure showed some slight improvements for the nurses but for the physicians no global change could be detected. However for the latter there were some divergent evolutions

¹⁰ *Australian Doctor*. Sharing PIP. 28-Jun-2006
<http://www.australiandoctor.com.au/news/bd/0c0412bd.asp> (consulted 5 Nov. 2009)

inside this composite indicator, with an improvement for the components linked to "support from management", "career evolution possibilities" and "getting feedback". Meanwhile there was a deterioration in work stress related components (Gené-Badia et al., 2007). The study also found that during the first year of the incentive scheme, user satisfaction did not change significantly. Finally, the study observed that there was a high positive correlation between user satisfaction and the QPL scores (better user satisfaction score correlated with a better QPL score). This correlation between professional satisfaction among health workers and patient satisfaction has been noted in the literature (Haas et al., 2000; Weisman and Nathanson, 1985). The result suggests that it is counterproductive to introduce a high level of result pressure on staff through an incentive scheme, since this can lower staff motivation and thus negatively impact patient satisfaction.

2.4.3 Italy

In Italy the public health care system is highly decentralized. The financially and managerially autonomous regions have been pushing to include performance incentives in the operational level contractual relationships between the Local Health Authorities and GPs. This has resulted in the implementation of a large variety of mechanisms that include performance related payment methods (Lippi Bruni et al., 2009). For example in the Emilia Romagna region, districts have introduced performance incentives for diabetes care and follow up. GPs that take on their patient lists diabetes type-2 patients and assume regular follow up, including periodic measure of HbA1c (glycosylated haemoglobin) levels, receive financial bonuses. The strategy aims to increase the quality of diabetes care and to transfer the burden of diabetes care from hospitals to GPs. The latter objective can be seen as a cost containment strategy.

2.4.4 France

In France, the National Health Insurance Organization (Caisse Nationale d'Assurance Maladie des Travailleurs Salariés (CNAMTS)) in April 2009 introduced a new performance based contract for French GPs, called the CAPI (le Contrat d'Amélioration des Pratiques Individuelles).

There are two explicit targets behind the initiative : quality of care and efficiency. The CAPI contract, which the GPs can sign individually on a voluntary basis, has two dimensions :

preventive care and follow up of chronic diseases; and "optimization of medicine prescriptions". The preventive care and chronic diseases component makes up 60% of the reward maximum and the medicine prescription component 40%. The preventive care and chronic disease component focuses on five dimensions: influenza vaccination coverage for population over 65 years; breast cancer screening; diabetes care and follow up; blood pressure and prevention of adverse effects of medicine use. Three of the conditions are followed with one indicator, adverse effects of medicines is followed by two (the level of prescription of vasodilators and benzodiazepine) and the diabetes component has four indicators linked to it. The drug prescription component is composed of seven indicators. The total number of indicators in the CAPI contract is thus sixteen. GPs receive a year end incentive bonus that depends on their aggregate performance score and number of patients registered

The French Medical Association recently voiced its opposition to the CAPI, one of the arguments being that the CAPI was not negotiated with the professional associations and boards and that it breaks the tradition of having collective agreements¹¹.

2.4.5 Elsewhere

In addition to the experiences discussed above, there are some high-income countries that are at the discussion stage regarding the implementation of provider performance incentive schemes; this is happening for example in Estonia (Maynard, 2008) and in Canada (Pink et al., 2006). The current literature on those countries relates to how provider performance incentive principles and mechanisms would be applied given the local situation

¹¹ <http://www.conseil-national.medecin.fr/article/%C2%AB-capi-%C2%BB-le-non-du-cnom-tels-qu-ils-sont-ils-contreviennent-la-deontologie-654>

3 Lessons learned

The objective of this chapter is to provide a discussion framework that will identify the main issues related to the use of provider performance incentives in high-income countries. The aim is not to be comprehensive but to concentrate on a couple of key elements that determine the success or failure of a given performance incentive scheme.

The starting point of our work, in the previous section, was to provide a general overview of some of the schemes in high-income countries. Different type of provider performance incentive schemes have also been implemented in the low-income country contexts. These include pilot projects and national scale Performance Based Financing schemes in Burundi (Busogoro and Beith, 2010) and in Rwanda (Meessen et al., 2007; Basinga et al., 2010); the Haiti case of performance contracting with local NGOs for primary care services (Eichler, Auxila & Pollock, 2001) or the pilot project introducing performance incentives for reproductive health services in Egypt (Huntington et al., 2009). These schemes are described elsewhere in the literature and we are not going to discuss them here; for a general overview of low-income country schemes one can look for example in Eichler et al. (2009), Perrot et al. (2010).

Discussing high- and low-income country schemes together and providing evidence that can transcend the contextual differences is not an easy task. The performance incentive schemes in high-income countries are often destined to solve different types of problems than those implemented in low-income country contexts. It could be argued that the performance incentive schemes in these two contexts evolve within two different "paradigmatic" frameworks. This document will not attempt to address to this complicated comparability question. However, in this second part of the document the focus is on some elements and issues that may be relevant when implementing performance incentives in high-income as well as in low-income country settings.

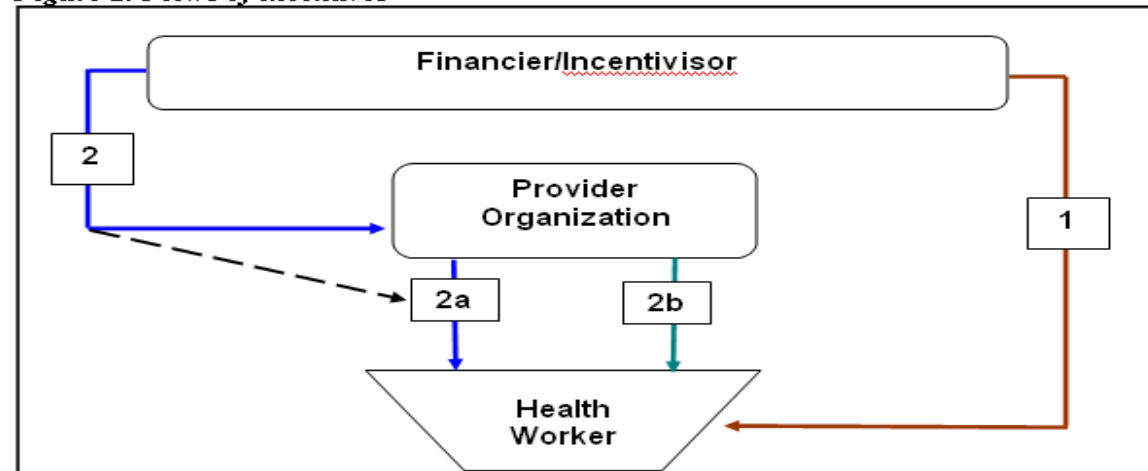
3.1 Individuals and organizations - who is rewarded and how?

One of the most important questions for performance incentive scheme architecture is the way it targets institutions and individuals. The provider incentive schemes can basically have either an individual or an organizational focus. However, in reality there are no clear-cut models and in most cases both organizational and individual targets are included. The CAPI mechanism in

France, for example, is based on a standard form (voluntary) contract that is established directly between the National Health Insurance organization and the individual GPs (French GPs, for the most part, work alone in practices that do not employ other clinical or administrative staff). The CAPI contract thus targets only the individual performance of the GPs who are the sole receivers of the eventual reward payments which come as a top-up to the standard fee-for-service reimbursements from the NHI. The CAPI programme is one of the “purest” examples of a scheme that has a direct individual focus.

There are also incentive schemes that primarily target organizations. The PHQID (see p. 6), which targets hospitals, is an example of such a scheme. Incentive schemes that target organizations introduce a new layer through which the incentives will have to flow. As already stated, performance incentives are built around the notion of influencing provider performance through extrinsic motivation. However, organizations do not have feelings and they do not respond to incentives, humans do. This means that the incentives targeting organizations have to be translated in one way or another to the staff level.

Figure 1. Flows of incentives



In the above figure three different itineraries of incentives are described. Path 1 represents the situation of direct incentives to health workers (the French CAPI for example). Path 2 goes through the organization; in the case 2a the role of the organization’s management is less important if the modalities of dealing with the scheme at the organizational level are already fully or partially integrated in the initial contract with the financier. In case 2b the contract between the financier and the provider organization does not cover the question of organizational level issues, thus leaving the organization free to define these modalities. The latter case refers to a “black box” approach where the financier is interested only in the results and gives the provider organization freedom of choice on how to achieve these results.

The lessons learned concerning the focus of the incentives do not relate to choosing between the targeting of individuals or of organizations, since it is obvious that the way the payments are targeted will strongly depend on how the health system is organized in general and on the scope and objectives of the scheme. For example, in France GPs are mainly operating from single handed practices, thus a scheme targeting GPs will necessarily be based on a direct contract between individual GPs and the financier. On the other hand if a scheme targets the performance of highly specialized hospitals it would be almost impossible to target each and every staff member, or even staff category; thus the target will be the organization and its performance.

Contextual constraints aside, there are still several important issues concerning the targeting of a performance incentive scheme that need to be figured out when building the operational mechanisms.

3.1.1 Building a mechanisms that incentivizes the right people

One interesting lesson that can be derived from a study of the literature is that there can sometimes be mismatches between the objective of the incentive scheme and the way it is targeted. The QOF programme in the United Kingdom, for example, is somewhat ambiguous about its target. Basically the QOF relies on a direct incentive mechanism where GPs are rewarded on their performance. However, even if the "spirit" of the scheme clearly targets GP performance, technically the QOF mechanism is based on standard form contracts that target the practices and not the individual practitioners. This means that the scheme also targets other clinical and non clinical staff¹². This set-up leaves open the question of who inside the practices should be rewarded and by how much.

It seems that there are no clear rules on how QOF bonuses are redistributed inside the practices, but there is some evidence that most of the financial benefits are kept by the GPs (or the GP partners to be exact), regardless of who actually contributed to the performance of the practice (Campbell et al., 2008). This would suggest that the incentives do not cascade to the administrative and other clinical staff, even though the individual performances of the staff in these categories partly define the organizational

12 In fact, in average only 23% of the staff working in practices are GPs (NHS information centre, GP workload survey, 2007), see for example: http://www.ic.nhs.uk/cmsincludes/process_document.asp?sPublicationID=1183384474670&sDocID=3830 (consulted 5 Nov. 2009)

performance of the practices. This has been observed to create some resentment among non-GP staff and to pose a possible threat to the practice team relationships (Hutchison, 2008). This issue is particularly important for nurses since an increasing volume of (routine) care procedures are delegated to them, especially in the management of chronic diseases (Wiener-Ogilvie et al., 2008). As these procedures actually form the main components that determine the QOF reward, logically nurses should be directly incentivized and their performance rewarded.

The Australian PIP incentive system is similar to the QOF with respect to targeting. Most PIP payments are provided as a top-up to the fee-for-service payments of the practices - without any individual earmarks (Keleher et al., 2007). There are few formal rules in the PIP system on how rewards should be distributed inside the practices. It seems that in most practices monetary rewards are channelled to GPs, especially in areas where there are GP shortages and where the PIP payments can be used to attract them.¹³

The absence of any systematic approach to rewarding non-GP staff in these two incentive schemes may be linked to the fact that both in Australia and in the United Kingdom, one of the explicit objectives was to encourage practices to employ more staff with the additional income. It could be argued that the scheme provided the necessary conditions for the non-GP staff to have a job in the first place and a basic salary that was relatively high. It is possible be however, that in the longer run, when the initial hiring effect fades away, the frustration over the bonus reward mechanisms among the non-GP staff may grow.

- ***Lesson learned # 1 : there can be too much emphasis on rewarding one category of staff over another; this is against the "meritocratic" underpinning of performance incentives and can be harmful in the long run; the problems are aggravated if there are no explicit rules and the reward sharing is done in an implicit manner***

3.1.2 Building adequate organizational arrangements

The QOF in the United Kingdom and the Australian PIP schemes can be considered as having a quasi-organizational scope since the incentives flow through the organizational level – in this case the practice - but at the same time this organizational layer is relatively thin. In the schemes that

¹³ *Australian Doctor*. Sharing PIP. 28-Jun-2006
<http://www.australiandoctor.com.au/news/bd/0c0412bd.asp> (consulted 5 Nov. 2009)

target larger organizations such as hospitals the organizational issues becomes much more complicated.

In high-income countries, apart from some micro-level schemes, the major performance incentive initiatives concerning hospitals are found in the USA. Once again it is difficult to draw any general conclusions but it seems that the question of how to incentivize staff has revealed some interesting issues.

It seems that the PHQID programme, for example, was accompanied with very little guidance from the financier; the main responsibility for implementing the plan at the hospital level was left to the hospitals themselves. A national study on the implementation of PHQID at the hospital level reveals that hospitals mainly relied on a strategy based on getting key staff engaged in making the necessary system changes (Gregg et al., 2006). The main objective of this approach was to keep key individuals "in the loop" and consequently to get their approval and support for the changes. During implementation the internal arrangements relied heavily on a managerial approach where the work of physicians in particular was monitored more closely than before; some of the hospitals also used peer review committees that assessed the individual performance of the staff and intervened if a problem was detected (for example a discussion with the individual whose performance was deemed unsatisfactory). It is thus interesting to note that incentives that were positive at the organizational level were translated into negative incentives - monitoring and sanctions - at the staff level.

However, in addition to the monitoring and sanction approach, some of the hospitals also resorted to internal positive incentives. One approach used was to establish project goals for key managers and physicians and to link bonus payments to the attainment of these goals. One hospital established an end of year bonus that was paid to all the staff members if the hospital met the performance goals.

Finally a common method for the PHQID hospitals to induce organizational change that would support the performance effort was to invest resources in those support factors that could help the staff to better respond to the new expectations. One of the main interventions of this kind was the upgrading of the IT systems (Gregg et al., 2006).

There are no studies that providing evidence that allow for comparison between the different organizational/managerial methods. However, it may be stated that that there does not seem to be one single approach to translating an incentive scheme from the organizational level to the staff level; managers can draw on a multitude of sanction and reward related actions to achieve their objectives. The positive results obtained from the PHQID intervention (CMMS, 2009; Lindenauer et al., 2007)(see p. 8) suggest that the wide range of methods employed were, in average, successful.

Perhaps the lesson to be learned here is that managers can be creative in the way they build organizational arrangements; above all it is important that some action be taken and that this be done in an explicit manner. Gregg et al. (2006) report that interviewed hospital managers, looking back at their experience, said that they had not given enough emphasis to informing or engaging the whole staff of the hospital in the process and that if they could do everything over again they would put more emphasis on this aspect.

It is interesting to note that very few of the studied PHQID hospitals implemented an organization wide bonus distribution system. One of the major reasons for this might be that the financial reward in the PHQID programme was fairly low in comparison to the total financial flows of the hospital (on average hospitals received ~ US\$50 000 as bonuses). If distributed to all the staff, the individual bonus would have been too small to make a difference. Some of the hospitals resorted to rewarding only some key individuals, and this could indeed be a way to get some leverage out of the rewards offered; but at the same time it might lead to an increase in tensions between the different categories of staff.

- *Lesson learned # 2: There is a clear need to translate the incentives targeting organizations to the staff level; there is no single method for achieving this but it seems that the different measures used will call for a sharpening of the managerial functions and an explicit dialogue/communication with staff.*
- *Lesson learned # 3: The performance incentive schemes targeting hospitals (or other larger health care organizations) have thus far provided only very small financial awards relative to the total income flows of these organizations. In these cases it is probable that the financial bonuses paid to the organization cannot be redistributed to all of the staff since the individual share would be too small to make a difference. This suggests that financial rewards should be used in a way to maximize leverage - for example by using them to upgrade the IT system, recruiting new staff or buying some items that improve the working conditions (for example a couch in the staff lounge).*

3.2 Different performance objectives

It seems that there is no universally agreed definition of what constitutes good performance for a health provider. There are different points of view that emphasize different aspects, some linking performance to quantity of output, some to quality (clinical quality, but often also the quality of processes), and for others the key consideration is efficiency.

3.2.1 *Quantity vs. quality*

In high-income countries performance incentive schemes explicitly focus on the quality of care. In fact, in some of the literature performance incentives are actually referred to as Quality Based Payments (Bodenheimer et al., 2005) or Paying For Quality (Rosenthal et al., 2004). In the USA the P4P "movement" was largely catalysed by the IOM report Crossing the Quality Chasm that put the gaps in quality of care on the front page. As its name indicates, the Quality and Outcomes Framework in the United Kingdom is also focused on quality issues. The other high-income schemes are also focused to a large extent on the quality.

In low-income countries the quality of care aspect is included in most performance incentive programmes, but the main focus has been on quantitative aspects - on the volume, of health service delivery. In other words, in low-income countries provider performance incentives mainly address the problem of underutilization of health services (Eichler et al., 2009). In order to increase service utilization providers are incentivized to maximize their efforts and consequently to increase the volume of activity. The schemes in low-income country settings are thus based on the hypothesis of a supply driven strategy that aims to increase the utilization of health services.

The volume of services produced by health care providers is not generally a problem in high-income countries where provider payment mechanisms, especially for primary care, are based, fully or partially, on fee-for-service mechanisms. Under these arrangements, practices and/or physicians are already incentivized to provide a high volume of services. In addition, a social protection system coupled with a high capacity to pay have also kept demand at a high level. These two supply and demand factors, added with the availability of a trained workforce, has lead to an overall high level of utilization of health care : on average, individuals living in OECD countries register 6.7 physician visits per year¹⁴. From this perspective it is possible to argue that, at an aggregate level, providers in high-income countries are performing well since they are

¹⁴ OECD Health Data 2009

producing high volumes of health care services. However, it has been argued that the "real" performance of the providers is less than satisfactory since the quality of care has not matched the quantity; the fee-for-service payment structure, for example, rewards numerous quick consultations rather than those clinical services that take time and are often related to prevention. In other words, structural incentives related to the provider payment mechanisms have been unable to provide quality of care.

In the USA, for example, it has been estimated that only a little over 50% of the population are treated with the recommended preventive and curative health care procedures (McGlynn et al., 2003; Kerr et al., 2004). In most cases previous quality improvement interventions have been limited to the issuing of guidelines, standards and other information-based guidance strategies. In this context the basic rationale for the use of performance incentives is to effectively align providers' incentives with the quality goals set.

In most high-income country schemes quality is defined in terms of adherence to intervention protocols, norms and standards. In most schemes some sort of outcome measures are also included, mainly through measures of intermediary outcomes that basically reflect the changes in the biological status of individuals (for example, blood pressure results for hypertensive patients). Patient experience has also been used as a quality measurement element. Quality is also linked to the providers' responsiveness. However, there have been mixed experiences concerning the measurement of patient experience and in the USA several schemes dropped this measure since it did not discriminate sufficiently between providers and in addition surveys were often costly.

➤ ***Lesson learned # 4 : a focus on quality or quantity is strongly related to context, with the focus being more on increasing utilization in low-income countries and enhancing quality of care in high-income countries. In high-income countries the organization and financing methods related to supply (provider payments, etc) and the factors related to demand (social protection, capacity to pay, etc.) have induced high volumes of services but with little by way of structural incentives for quality - this is why performance incentives schemes have tended to target quality issues.***

3.2.2 Cost and cost-effectiveness

When talking about incentive reforms in high-income health systems, one often refers to discussions of structural reforms designed to influence provider behaviour in order to achieve efficiency gains and/or cost containment. Most often these reforms are related to provider

payment mechanisms (WHO, 2007). For example, the introduction of Disease Related Groups (DRGs) has been seen as a provider payment mechanism reform that has the overall objective of containing the rise of health care (hospital) costs.

However, the introduction of explicit provider performance incentives, such as the P4P programmes, in high-income countries is not directly linked with the cost saving agenda included in other, overarching reforms (Rosenthal et al., 2007). In fact performance incentive schemes in high-income countries are either cost neutral or actually come with an increase in cost.

Incentive schemes can in theory be cost neutral if the providers' volume of service/effort stays the same, but there will be a shift from one set of activities to another - for example from curative to preventive care. In other words, this would imply that the financier transfers (some of) its funding from one type of services to another.

In most cases however it seems that the schemes implemented carry a cost increase. The British QOF programme is a good example: it has a general yearly price tag of £1 Billion and has resulted in an increase in the level of payments to GP practices of about 25% in the last five years. The French CAPI system will also necessarily result in a rise in the reimbursement payments from the NHI since the system is based on a top-up of the basic reimbursements.

Moreover, along with the increased payments to providers there are also costs associated with implementing the schemes. For the financiers this is a clear programme cost, in other words a cost for setting up and administering the scheme. But there will unavoidably also be charges on the providers that are targeted by the schemes; providers need administrative support to keep up with reporting and other requirements that come with a performance incentive scheme. These costs at the provider organization level will in one way or another be recovered from the financiers/purchasers over the medium to long term.

Nevertheless, there are instances where performance incentives can lead to a decrease in the financiers' costs: notably where incentive schemes encourage direct efficiency gains. For example the French CAPI mechanism is partly focused on what is called the "optimization of medicine prescriptions". In theory if the incentive scheme achieves its objective, there should be fewer unnecessary prescriptions which will in turn lower the total reimbursements (France has the third highest pharmaceutical per capita spending among the OECD countries). Savings can also be

achieved where performance incentive schemes have the implicit objective of reducing the need for curative care over the long term. For example, the QOF rewards GPs for smoking cessation counselling. It is obvious that if this measure lowers the prevalence of smoking in the population, there will be fewer curative services needed for lung cancer treatment in the future. For the moment we lack the data required to evaluate these hypotheses.

➤ ***Lesson learned # 5 : most performance incentive schemes will increase costs for the financier/payer/purchaser since they introduce new payments and administrative costs; however, there can be long savings.***

The experiences from USA shows that there is very little correlation between the amount of incentives paid and the success of the scheme (Rosenthal et al., 2007b). On the other hand the British experience shows that a performance incentive scheme can indeed give positive results but that it can also be costly. Today the cost-benefit interpretation of the QOF remains a hot political issue in the United Kingdom (Campbell et al., 2007). The NHS agency National Institute for Health and Clinical Excellence (NICE), which manages the QOF indicator development process, has suggested that the current price value of the QOF indicators do not reflect the health gains obtained¹⁵.

NICE has put together an expert committee that has the objective of creating a methodology that allows the systematic inclusion of cost-effectiveness in further revisions of the QOF. Moreover, the use of incentive money has given rise to some lively discussions; there has been some concerns from the administrators and financiers (Department of Health, NHS) side that most of the money has gone to GP pay and not enough to investments at the practices. This may mean that the long term "investment yields" for the financier will not be as high as expected.

As already noticed the (explicit) focus of the performance incentive schemes has not been on cost-effectiveness, and it is for this reason there are very few cost-effectiveness studies. Nahra et al. (2006) have studied the cost effectiveness of a P4P scheme that targeted heart conditions. They converted care improvements into quality-adjusted life years (QALYs) gained. The hypotheses made were built on literature on clinical effectiveness of heart condition interventions and their impact on survival rates. The study found out that an estimated 24,418 patients received improved care as a result of the P4P scheme. This resulted in a QALY gain that ranged from 733 to 1,701 -

¹⁵ NHS National Institute for Health and Clinical Excellence : <http://www.nice.org.uk/aboutnice/qof/qof.jsp>

depending on the underlying assumptions. Cost per QALY gained was calculated to be between US\$12,967 and US\$30,081.

Another study focused on the savings of a P4P scheme against a previous trend for health care costs (Curtin et al., 2006). The P4P scheme studied, implemented in the Rochester area of the state of New York, was rewarding primary care physicians according to their compliance with a set of care processes. The scheme was cost neutral from the reimbursement point of view since the performance related pay came from a "pool" fund comprised of a 10% withhold of the normal reimbursements to the physicians. However, the scheme itself was not cost neutral since the programme costs (creating the measurement system, administering the scheme, etc.) were evaluated at US\$1.15 million per year. The study, which followed only developments in diabetes care, compared a projection of a baseline trend based on historic cost and an actual trend with the P4P implemented. The study found that by comparing the trends, the P4P scheme had resulted in savings of US\$1.9 million the first year and US\$2.9 million for the second year. Thus the investment in the P4P programme yielded positive returns for the financier (it is not known how much of the savings were redistributed to the physicians).

The return on investment for the financier was as high as 160% for the first year and 250% the second year. The study has some important limitations, notably the fact that it compared the baseline trend with a trend line that was based only on the two first years of the P4P implementation. There was also no control group comparison in the study. Finally, it is important to underline the fact that the "return on investment" calculated here refers to a situation where the scheme was cost neutral for the financier when it comes to the reimbursements/payments to the physicians. This means also that the physicians, through their association, were ready to put at risk a part of their payments.

- ***Lesson learned # 6 : the performance incentive schemes in high-income countries do not target overall efficiencies, consequently there is very little literature on their cost-effectiveness or cost-benefits. Some studies have shown that there are possible positive effects, but the results are highly programme and context specific.***

3.3 Choosing the indicators

The indicators are arguably the most important functional elements of the performance incentive schemes. Indicators are used to manifest the schemes' objectives; if the scheme targets a set of clinical conditions, for example, these must be translated into indicators that can be used to measure performance. The passage from a programme objective to an indicator is not always easy, and sometimes it is the availability and/or the reliability of indicators that dictates what the rewarded elements are.

The basic rationale of indicator selection is that the adopted indicators be evidence-based. However, if there were unanimity on what constitutes an objective set of "evidence-based" indicators there would be only little variation between different schemes. This is of course far from the reality and is indicative of the fact that there are several different factors that can influence the choice of the final set of indicators included in a scheme.

Finally, the mechanics of indicator setting and the construction of an indicator set imposes further complications. Kahn et al. (2006) for example argue that although the individual indicators can be seen as evidence-based, transforming those indicators into composite scores, for which different weighting and/or different "prices" are used, requires choices for which there is no evidence based foundation.

3.3.1 *How to choose the indicators?*

The process of choosing indicators is not straightforward. In the United Kingdom, the QOF indicators were chosen during the 18 month negotiations between the health administration (representing the financiers) and the British Medical Association (representing the providers). This lengthy negotiation process was designed to mollify possible opposition from the medical profession. At the end of the negotiations the parties came to an agreement on the first set of 146 indicators that were include in the QOF.

As already discussed, GP practices have been reaching very high levels of performance if measured using the indicators used for the QOF scheme. There has been some discussion on the fact that the BMA negotiators got a "good deal" by influencing the choice of the indicators so that the final QOF mechanism included indicators and reward levels that increased the probability of

attaining high scores. For example for the 2006/2007 QOF, 5.1% of primary care practices obtained the maximum score of 1000 points, and a further 68.5% had a score between 950 and 1000 points; only 6.9% of the practices had a score below 850 points (or 85% of the maximum).¹⁶

In the French CAPI mechanisms, the negotiations between providers (the GP associations) and financiers (the Social Health Insurance administration) were not as comprehensive as was the case with the QOF. The choice of indicators was based on a top-down approach. This has already created some tensions, and the French National Medical Association (Le Conseil National de l'Ordre des Médecins) has been openly hostile towards the CAPI scheme¹⁷.

In the USA several schemes are built with technical assistance from a third party organization that acts as the developer and manager of the schemes. For example in California the Integrated Health Care Association (IHA), which is an independent joint-venture of several insurers, health plans, universities and providers, operates the P4P programmes on behalf of the financiers. The independent structure of the IHA and the fact that it represent all the major stakeholders makes it a useful mechanism for enabling dialogue and achieving compromise. Moreover, such intermediary organizations can delegate some of the indicator choosing and definition tasks to another organization which will do the work. This chain of delegation approach, in theory, reduces the likelihood of vested interests skewing the indicator choosing process.

In the United Kingdom the National Institute for Clinical Excellence (NICE) and its independent Primary Care QOF Indicator Advisory Committee was commissioned to produce a menu of indicators that could be used to feed into the QOF 2010/2011 revision. In the end the NICE indicators were not included (at least not immediately), since the BMA and NHS negotiations did not approve them. This serves as a reminder that the choosing of the indicators can be an independent and objective process (driven by an evidence-based rationale), but only up to a certain point; the final decision will always be arrived at by the financier and the providers through negotiation. (and may reflect differences in the power or 'clout' of the parties involved).

¹⁶ Source: *The Health and Social Care Information Centre*

<http://www.ic.nhs.uk/webfiles/QOF/2006-07/QOF%202006-07%20Statistical%20Bulletin.pdf>

¹⁷ <http://www.conseil-national.medecin.fr/article/%C2%AB-capi-%C2%BB-le-non-du-cnom-tels-qu-ils-sont-ils-contreviennent-la-deontologie-654>

- *Lesson learned # 8 : there are no standard procedures for selecting the indicators; different types of organizational and procedural factors will interfere and thus the "evidence base" of the indicators will get "adjusted" during the process.*
- *Lesson learned #9 : the financier and the providers will have different objectives (as the agency theory demonstrates) thus compromises need to be reached; if the providers' voice is not heard this might result in their disengagement, if the providers get too much of a say the scheme can become costly to the financier. The use of intermediary organizations can facilitate the reaching of compromises and can keep indicators linked to objective criteria.*

3.3.2 *How many indicators?*

The number of indicators used in different schemes has varied quite substantially. In France the CAPI programme is constructed around 16 indicators; on the other hand the QOF mechanisms relies currently on 134 indicators. In the USA the number of indicators varies greatly between schemes.

Once again, the number of indicators used depends on the scope and objectives of the scheme, which makes it impossible to draw any meaningful conclusions about what is the optimal number. There are some important observations that can nevertheless be made. Rosenthal et al. (2007) for example found that in the set of P4P schemes that they had followed over a three year period, in 99% of cases the number of indicators had increased. This happened notably with indicators in domains such as IT development and clinical outcome. The IHA indicators have also been steadily increasing in number during the five years of its existence; in 2003 they were 23 and in 2009 the number had reached 68¹⁸. This suggests that there is a tendency to add new indicators to schemes as they mature, and that there is often an evolution towards comprehensiveness and decreasing concern about the financial and administrative burden imposed by large numbers of indicators.

At the same time some indicators get dropped during the iterative process of redefining indicators at regular intervals (usually a year). The QOF has seen indicators come and go during the first five years of its existence; the original version included 146 indicators and the current version has

¹⁸ http://www.iha.org/pdfs_documents/p4p_california/P4PWhitePaper2_June2009_FullReport.pdf

134. In most schemes the indicators used have not stayed the same from one year to the next and it seems that as a rule there are at least some changes in indicators each time they are revised.

- ***Lesson learned #10 : the number of indicators used is necessarily specific to the different schemes, however there are some indications that as programmes evolve the number of indicators increases, meaning that the comprehensiveness of the schemes becomes more of an issue as time goes by and that the financial and administrative burden becomes less important once the scheme has built up speed.***

3.4 Reward mechanisms

There are several options when it comes to creating the reward mechanism for a performance incentive scheme. These options translate into additional fine tuning elements that can be used to strategically align the scheme's architecture with its objectives.

3.4.1 Improvement vs. nominal result

The conceptual difference between a reward system that is based on improvement and a mechanism based on a nominal result is easy to grasp. The improvement based reward system functions on the basis of an improvement in time between two points of observation, while the nominal measurement will simply reward the result at one point in time.

In high-income countries both methods have been used. While it seems that most of the programmes target nominal results, Rosenthal et al. (2007) observes that US P4P programmes have been increasingly using improvement measures; in 2003 all of the surveyed schemes were referring to a nominal performance measurement but in 2006 nearly a quarter of the schemes were rewarding an improvement in the performance. Petersen et al. (2006) found that the providers with the lowest baseline performance often improved the most and thus got bigger rewards than when the reward was given on the basis of nominal performance. This reward mechanism thus has the advantage of getting all providers more or less on the same starting line. On the other hand where nominal measurement is chosen providers with a low baseline score might be less eager to invest resources and effort since the likelihood of getting a reward is low. However, improvement based reward can have a negative effect on high performers since they will need higher marginal investments to improve their performance which is already at a high baseline level.

Of course schemes can use both nominal and relative payment systems. The PHQID scheme, for example, rewards target attainment as well as improvement. The attainment award is accessible to all those who attain a certain threshold of performance; but in addition to this, improvements in performance are also rewarded. However, in the PHQID example the improvement award is accessible only to the top 20% of hospitals. This keeps the top performers focused on improvement since they will be measured against a pool of high performing hospitals. The French CAPI system also has attainment and improvement rewards; those GPs with lower baselines are rewarded for improvements and those with higher baselines are rewarded for nominal result - a mechanisms designed to even out possible disparities in incentive payments between GPs.

- ***Lesson learned #11 : the choice between rewarding based on performance targets or improvements in performance can have a direct influence on how the providers will effectively be incentivized; those with low baseline performance will be more responsive to improvement related reward while the high performers will be more responsive to rewards for target attainment; in some cases schemes have resorted to a mixture of the two methods.***

3.4.2 Competing mechanisms vs. individual rewards

Incentive rewards based on competition mechanisms rely on the notion of a finite divided cake - the best providers getting the biggest pieces and the worst getting the small or nothing at all. Individual rewards, on the other hand, are not taken from a common pool, and each provider will get their reward according to their individual performance - the performance of one provider will not have an influence on the reward of the others.

The above mentioned experience from Rochester NY (p. 26) is an interesting example of a competing reward system. In this scheme 10% of total reimbursements were held in a common pool (the cake) that was then distributed to the physicians, an arrangement that meant that there were going to be losers and winners from the beginning. Indeed the bonuses awarded to the physicians varied between .5 and 1.5 times the average payment, meaning that some physicians were getting 95% of the income that they would of have gotten without the P4P programme while others were getting up to 105% of their "normal" reimbursements (the performance multiplier affected of course only the 10% that was withheld, hence the distribution was +/- 5%) .

The QOF scheme is built on an individual reward mechanism where the reward paid to one practice does not reduce the potential reward to the others.

One of the most notable differences between these two mechanisms is their impact on the cost for the financier. In the competing model the cost of the programme can be better managed by the financier since the total sum of the reward bonuses will be known in advance. The competition based reward system can be cost neutral, as in the Rochester case (excluding the programme costs), or they can be a defined sum that will be added to the general pool of payments. If, on the other hand, the reward is defined separately for each provider, the cost of the aggregated rewards is unknown to the financier in advance. In the QOF scheme the total cost was in fact a great deal higher than expected, and the financier, the NHS, had to increase the budget allocated to QOF after the first implementation year.

It has to be underlined that even in the case of individual rewards, some level of competition can still enter into consideration, for example through the publication of the results.

➤ ***Lesson learned #12 : the choice between a competing or an individual reward system can have important repercussions for the financier; in theory a competition based reward mechanism is better for the financier since it will allow him to predict the financial impact. However because this system creates winners and losers, it is not easy to get the providers' acceptance. Both types of reward mechanism have been used.***

3.5 Possible negative effects

There is no clear evidence from high-income countries on the possible negative effects of performance incentive schemes. The evidence on the negative effects has to be gathered from different individual studies that have often focused on only one aspect of possible negative effects, and even then, the literature is fairly thin.

The list of possible negative effects associated with performance incentives could, potentially, be quite long. The Price Water House Cooper survey on performance incentives identified two major categories of potential negative effects (as perceived by the interviewed key stakeholders) : (i)

providers targeting higher reward procedures and patients; and (ii) providers decreasing focus on areas not rewarded¹⁹.

3.5.1 Targeting higher rewarding procedure and patients

Under the QOF programme GPs' practices were allowed to exclude patients from their performance count under certain conditions. These conditions could be for example a contraindication to a certain treatment or the fact that the patient had recently registered at the practice. There has been widespread suspicion that the high levels of practice performance achieved was partly due to the misuse of exemption reporting that allowed for gaming of the system. In the first years of QOF implementation wide variations in exemption reporting was observed, the exclusion rate varying between 2% and 25% (Doran et al., 2008).

It is impossible to attribute these variations to deliberate gaming, although it has been reported that only a small proportion of practices (1.1%) excluded more than 15% of their patients (Doran et al., 2006). Behind the variation, the average rate of exclusion was only 5%. It is impossible to judge whether this figure is high or low, but in some studies this has been seen as a plausible figure and Doran et al. (2008) conclude that no evidence of widespread gaming can be found from the exclusion figures. On the other hand Gravel et al. (2007) have found structural differences in the exclusion reporting that do point to some gaming behaviour. Based on a regression analysis, they estimate that exemption rates were on average 10.9% higher than they theoretically should of have been.

One of the ways to game the system is to concentrate on reporting rather than on the actual intervention itself. Roski et al. (2003) report that a scheme in the USA that was rewarding tobacco cessation advice resulted in an increased documentation of tobacco use status but the provision of advice to quit smoking did not increase. This result highlights the possible problem of gaming behavior, which in this case manifests itself in improvements in documentation rather than a change in the quality of health care delivered to patients.

¹⁹ Source: PricewaterhouseCoopers' Health Research Institute. *Paying for Performance: Incentives and the English Health System*. 2008
<http://www.pwc.com/gx/en/healthcare/paying-performance-incentives-english-health-system.jhtml>

3.5.2 Lack of focus on services not covered by the incentives

Some studies in the United Kingdom suggest that the un-incentivised aspects of care have not improved as much as those that are directly rewarded (Campbell et al., 2007). On the other hand other studies have reported a 'halo-effect' where non-rewarded services improve alongside the rewarded services at the same place (although this finding may also point to the fact that the change in behaviour is not be attributable to the incentives in place). A study from Glickman et al. (2007) on the PHQID programme concluded that the care processes for acute myocardial infarction were not skewed towards processes that were rewarded and that the providers resorted also to other, non-rewarded, and clinically pertinent care processes.

3.5.3 Gaming and fraud not a problem?

There is relatively little literature on the issue of gaming and fraud in performance incentive schemes. There is even less literature backed by substantial evidence on the occurrence of these negative behaviours. This could lead one to think that gaming and fraud are not important problems in performance incentive schemes, an hypothesis supported by the intuitive argument that, as the stakes for the providers are fairly low (as the schemes often offer only a 1-2% increase in income) the gain from gaming or fraud is also low. The QOF scheme is an exception here since it has a major impact on the GPs revenues. The QOF also include a yearly audit and possibly severe sanctions but it is not known how comprehensive or effective the yearly audits have been. Outside the QOF, sanctions seem to be relatively weak in most cases, excluding the implicit sanction of reputation loss.

It is fair to say that in general verification and monitoring mechanism have not been a central issue in the performance incentive schemes' overall design. This is in part understandable since there are already several mechanism in place that are used for provider monitoring. However it is clear that performance incentive schemes are not always compatible with these monitoring mechanisms and there is a need to think about new mechanisms to identify and sanction unwanted behaviour.

- ***Lesson learned #13 : There are some concerns that gaming can be a problem and there is some evidence of its existence. There is a clear need to put more emphasis on studies that observe performance incentive schemes from this perspective.***

4 Conclusion

Provider performance incentives can be best described as an "approach" or a "concept" that has its basis in a set of theoretical principals, such as the agency theory, information asymmetry or the distinction between extrinsic and intrinsic motivation factors. However, these theoretical principals give little guidance on how to actually build a meaningful intervention. In practice there exists a multitude of scheme setups which all have their own intervention logic, objectives, and particularities.

The wide variety of provider performance schemes has naturally turned discussion of them towards the practicalities of implementation. The different architectural elements each with its different implications offer an interesting challenge in analysing the "moving parts" of provider performance incentive schemes. In this document we have tried to put forward some of the key issues related to the implementation and the architecture of provider performance schemes in the high-income countries. We have put forward a number of lessons learned that can be useful for future schemes or for improving existing ones in different contexts and in both low- and middle-income countries.

There are however limitations to the drawing of conclusions based on the current evidence. What evidence there is is still relatively thin and sometimes contradictory. Similarly, because schemes are often only recently implemented we do not yet have the historical perspective needed to really understand their impact at various levels. At best there exists partial evidence on some specific elements of the schemes from specific contexts, but there is generally a lack of evidence that would allow us to make far reaching conclusions on the schemes and their implementation strategies. This also limits the possibility for providing solid cases of lessons learned.

The second limitation relates to context-specific nature of performance incentive schemes. The wide heterogeneity observed reflects the fact that in different contexts schemes are implemented in different ways in order to respond to different sets of problems. There are several fundamental differences for example in the financing of the schemes and in the purchaser-provider relationships.

Taken into account these limits and the complex and "unstable" nature of the object of study, it is impossible to draw far reaching conclusions on the use of provider performance incentives based

on the experiences in high-income countries. However, the proliferation of such schemes and the ever widening and deepening literature on them already gives some very useful insights into some of the key implementation issues : what to do and what not to do. This will surely lead to the development of more robust and sophisticated schemes in the future. But will provider performance incentive schemes, even at the height of their sophistication, constitute a panacea for all the problems health systems face around the world? That is another question.

References

Arrowsmith J., French S., Gilman M., Richardson R. Performance-related pay in health care. *Journal of Health Services and Research Policy*. 2001;6(2): 114-9.

Ashworth M., Seed P., Armstrong D., Durbaba S., Jones R. The Relationship between Social Deprivation and the Quality of Primary Care: A National Survey Using Indicators from the UK Quality and Outcomes Framework. *British Journal of General Practice*. 2007;57: 441–48

Basinga P., Gertler P.J., Binagwaho A., Soucat A.L.B., Sturdy J.R., Vermeersch C.M.J., *Paying Primary Health Care Centers for Performance in Rwanda*. 2010. The World Bank. Policy Research Working Paper 5190

Benavent J. et al. Pay-for-performance in primary healthcare centres in Spain. *Quality in Primary Care*. 2009;17: 123–31

Berwick D. A User's Manual For The IOM's 'Quality Chasm' Report. *Health Affairs* 2002;21(3): 80-90

Bokhour B.G., Burgess J.F., Hook J.M., White B., Berlowitz D., Guldin M.R., Meterko M., Young G. Incentive Implementation in Physician Practices: A Qualitative Study of Practice Executive Perspectives on Pay for Performance. *Medical Care Research and Review* 2006;63: 73S-95S

Bodenheimer T., May J.H., Berenson R.A., Coughlan J., Can money buy quality? Physician response to pay for performance. 2005. *Center for Studying Health System Change*. Issue Brief No. 102

Busogoro J-F and Beith A. Pay for Performance for Improved Health In Burundi. USAID. 2010. <http://www.hs2020.org/content/resource/detail/2575/>

Campbell S.M., McDonald R., Lester H. The Experience of Pay for Performance in English Family Practice: A Qualitative Study. *Ann Fam Med*. 2008;6(3): 228–234.

Campbell S., Reeves D., Kontopantelis E., Middleton E., Sibbald B., Roland M. Quality of Primary Care in England with the Introduction of Pay for Performance. *N Engl J Med*. 2007;357(2): 181-190

Campbell S., Wilkin D., Roland M. Primary care groups: improving quality of care through clinical governance. *BMJ*. 2001;322: 1580-2.

Chaix-Couturier C, Durand-Zaleski I, Jolly D, Durieux P. Effects of financial incentives on medical practice: results from a systematic review of the literature and methodological issues. *Int J Qual Health Care*. 2000; 12:133-142.

Commission for Local Authority Employers. The Collective Agreement for Municipal Physicians.

<http://www.kuntatyonantajat.fi/files/EA246B2349794EC1A9355A7AFC0738C6> (consulted 29 October 2009)

Conrad D.A & Christianson J.B. Penetrating the "Black Box": Financial Incentives for Enhancing the Quality of Physician Services. *Medical Care Research and Review*. 2004; 61: 37S-68S

Curtin K., Beckman H., Pankow G., Milillo Y., Greene R.A. Return on investment in pay for performance: a diabetes case study. *Journal of Health Care Management*. 2006;51(6): 365-374

Dagens Nyheter, *Snabba läkare belönas med bonus*, 2009-06-10.
<http://www.dn.se/sthlm/snabba-lakare-belonas-med-bonus-1.888110> (consulted 5 November 2009)

Dielemann M., Viet Cuong P., Vu Anh L., Martineau T. Identifying factors for job motivation of rural health workers in North VietNam. *Human Resources for Health*. 2003;1: 1-10.

Doran T., Fullwood C., Gravelle H., Reeves D., Kontopantelis E., Hiroeh U., Roland M. Pay-for-Performance Programs in Family Practices in the United Kingdom. *N Engl J Med*. 2006;355: 375-84.

Doran T., Fullwood C., Reeves D., Gravelle H., Roland M. Exclusion of Patients from Pay-for-Performance Targets by English Physicians. *N Engl J Med*. 2008;359: 274-84.

Eichler R, Levine R and the Performance-Based Incentives Working Group. *Performance Incentives for Global Health*. Center for Global Development. Washington D.C. 2009

Eichler R., Auxila P., Pollock J. Promoting Preventive Health Care. Paying for Performance in Haiti. in *Contracting for Public Services: Output Based Aid and Its Application*. eds. Brook P.J. & Smith S. The World Bank. Washington D.C.

Figueras J., Robinson R., Jakubowski E. (eds). Purchasing to improve health systems performance. *The European Observatory on Health Systems and Policies*. 2005
<http://www.euro.who.int/Document/E86300.pdf>

Franco L.M., Bennett S., Kanfer R. Health sector reform and public sector health worker motivation a conceptual framework. *Social Science and Medicine*. 2002;54: 1255-1266.

Fujisawa R. & Lafortune G. The remuneration of general practitioners and specialists in 14 OECD countries: What are the factors influencing variations across countries? *Organization of Economic Co-operation and Development*. Health Working Papers 2008.

Gene-Badia J., Escaramis-Babiano G., Sans-Corrales M. Impact of economic incentives on quality of professional life and on end-user satisfaction in primary care. *Health Policy* 2007;80: 2-10.

Glickman S.W., Ou F.-S., DeLong E.R., Roe M.T., Lytle B.L., Mulgund J., Rumsfeld J.S., Gibler W.B., Ohman E.S., Schulman K.A., Peterson E.D. Pay for performance, quality of care, and outcomes in acute myocardial infarction. *JAMA*. 2007;297: 2373-2380.

Gravelle H., Sutton M., Ma A. Doctor behaviour under a pay for performance contract: Evidence from the quality and outcomes framework. *Centre for Health Economics*. York. 2007; CHE Research Paper no. 28.

- Gregg W., Moscovice I., Remus D. The Implementation of Pay-For-Performance in Rural Hospitals: Lessons from the Hospital Quality Incentive Demonstration Project. *Upper Midwest Rural Health Research Center*. 2006. Working Paper 2 www.uppermidwestrhrc.org
- Huntington D., Zaky H.H.M., Shawky S, Fattah F. A. Impact of provider incentive payments on reproductive health services in Egypt. 2009. World Health Organization. http://whqlibdoc.who.int/hq/2009/WHO_RHR_09.04_eng.pdf
- Hutchison B. Pay for Performance in Primary Care: Proceed with Caution, Pitfalls Ahead *Healthcare Policy*. 2008;4(1): 10–15
- Institute of Medicine*. Rewarding provider performance: aligning incentives in Medicare. Washington DC. National Academy Press, 2007.
- Institute of Medicine*. Crossing the Quality Chasm: A New Health System for the Twenty-first Century. Washington DC. National Academy Press, 2001.
- Jennifer S. H., Cook E.F., Puopolo A.L. Burstin H.R., Cleary P.D., Brennan T.A. Is the Professional Satisfaction of General Internists Associated with Patient Satisfaction? *J Gen Intern Med*. 2000;15: 122–128.
- Kahn C.N. III, Ault T., Isenstein H., Potetz L., Van Gelder S. Snapshot Of Hospital Quality Reporting And Pay-For-Performance Under Medicare. *Health Affairs*. 2006;25(1): 148-162
- Keleher H., Joyce C.M., Parker R., Piterman L. Practice nurses in Australia: current issues and future directions. *MJA*. 2007;187(2): 108-110
- Kerr E.A. et al. Profiling the quality of care in twelve communities: results from the CQI study. *Health Affairs*. 2004;23(3): 247–256.
- Lindenauer P.K., Remus D., Roman S., Rothberg M.B., Benjamin E.M., Ma A., Bratzler D.W. Public Reporting and Pay for Performance in Hospital Quality Improvement. *N Engl J Med*. 2007;356: 486-96
- Lippi Bruni M, Nobilio L, Ugolini C. Economic incentives in general practice: The impact of pay-for-participation and pay-for-compliance programs on diabetes care. *Health Policy*. 2009; 90:140-148
- Mangin D., Toop L. The quality and outcomes framework: what have you done to yourselves? *Br J Gen Pract*. 2007;57: 435-7
- Maynard A. Payment for Performance (P4P): International experience and a cautionary proposal for Estonia. *World Health Organization - Regional Office for Europe*. 2008
- Meessen, Bruno, Jean-Pierre Kashala, and Laurent Musango. Output-Based Payment to Boost Staff Productivity in Public Health Centers: Contracting in Kabutare District, Rwanda. *Bulletin of the World Health Organization* 2008; 85(2): 85–160.
- McGlynn EA. et al. The quality of health care delivered to adults in the United States. *N Engl J Med*. 2003; 348: 2635-2645

- Nahra T.E., Reiter K.L., Hirth R.A., Shermer J.E. and Wheeler J.R. Cost-Effectiveness of Hospital Pay-for-Performance Incentives. *Medical Care Research and Review*. 2006;63 (1): 49S-72S
- Oldroyd J., Proudfoot J., Infante F.A., Davies G.P., Bubner T., Holton C., Beilby J.J., Harris M.F. Providing healthcare for people with chronic illness: the views of Australian GPs. *MJA*. 2003;79: 30-33
- Organization of Economic Co-operation and Development*. Paying for Performance: Policies for Government Employees. 2005. Policy Brief
- Organization of Economic Co-operation and Development*. Health at a glance 2007. OECD indicators. 2007
- Perrot J et al. *Performance incentives for health care providers*. Geneva, World Health Organization, 2010 (Health Systems Financing Discussion Paper, HSS/HSF/DP.E.10.1; http://www.who.int/contracting/DP_10_1_EN.pdf, accessed 7 July 2010).
- Petersen L.A., Woodard D.L., Urech T., Daw C., Sookanan S. Does Pay-for-Performance Improve the Quality of Health Care? *Ann Intern Med*. 2006;145: 265-272
- Pink G.H., Brown A.D., Studer M.L., Reiter K.L., Leatt P. Pay-for-performance in publicly financed healthcare: some international experience and considerations for Canada. *Healthc Pap*. 2006;6: 8-26.
- Pita Barros P. & de Almeida Simões J. Health Systems in Transition. Portugal: Health System Review. *The European Observatory on Health Systems and Policies*. 2007
- PricewaterhouseCoopers' Health Research Institute*. Paying for Performance: Incentives and the English Health System. 2008
- Roland, M. Linking Physicians' Pay to the Quality of Care — A Major Experiment in the United Kingdom. *N Engl J Med*. 2004;351: 1448-1454
- Rosenthal M.B. Beyond pay for performance – emerging models of provider–payment reform. *New England Journal of Medicine*. 2008;359: 1197–200.
- Rosenthal M.B., Landon B.E., Howitt K., Ryu Song H.S., Epstein A.M. Climbing Up The Pay-For-Performance Learning Curve: Where Are The Early Adopters Now? *Health Affairs*. 2007;26(6): 1674–1682
- Rosenthal M.B. & Dudley A.R. Pay-for-Performance: Will the Latest Payment Trend Improve Care? *JAMA*. 2007b;297: 740-744
- Rosenthal M.B., Fernandopulle R., Ryu Song H.S., Landon B. Paying For Quality: Providers' Incentives For Quality Improvement. *Health Affairs*. 2004;23(3): 127-141
- Roski J. et al. The impact of financial incentives and a patient registry on preventive care quality: increasing provider adherence to evidence-based smoking cessation practice guidelines. *Prev Med*. 2003;36: 291-9

Sabri B., Siddiqi S., Ahmed A.M., Kakar F.K., Perrot J. Towards sustainable delivery of health services in Afghanistan: options for the future. *Bulletin of the World Health Organization*. 2007; 85(9): 712-718

Scott I. A. Pay for performance in health care: strategic issues for Australian experiments. *MJA*. 2007;187: 31–35

The Future of Health Care Task Force. Pay for Performance in Health Care: An analysis of the arguments for and against. The American Association of Clinical Endocrinologists. 2005

Timmins N. Do GPs deserve their recent pay rise? *BMJ*. 2005;331: 800

Trisolini M., Aggarwal J., Leung M., Pope G., Kautter J., The Medicare Physician Group Practice Demonstration: Lessons Learned on Improving Quality and Efficiency in Health Care. *The Commonwealth Fund*. February 2008

Van de Ven W.P.M.M. & Schutman F.T. Competition in the Netherlands: still work-in-progress. *Health Econ*. 2009;18: 253–255

Weisman C.S. & Nathanson C.A. Professional Satisfaction and Client Outcomes. A Comparative Organizational Analysis. *Medical Care*. 1985;23(10): 1179-1189

World Health Organization. Provider Payments and cost-containment. Lessons from OECD countries. WHO/HQ/HSS/HSF/HFP Technical Brief for Policy Makers. 2007
http://www.who.int/health_financing/documents/pb_e_07_2-providerpay_oecd.pdf