During the past decade health care insurers in the United States have been integrating a model of provider reimbursement called pay-for-performance (P4P) into the traditional reimbursement system, where the models have been fee-for-service, capitation or budgets, case-related reimbursement (for example, diagnosis-related groups – DRGs), and salaries. Under pay-for-performance, providers are given financial incentives to encourage and reinforce pre-established targets for health care delivery. The overall intent is to improve the quality of health care provision and enhance patient outcomes.

In Europe pay-for-performance within health care is still a relatively new concept, with the UK being one of the first European countries to implement a system of this kind. Given its novelty in Europe, it is important to discuss whether the outcomes of P4P in the US have been as expected and their relevant implications for European health policy.

Introduction to health insurance and delivery in the US

Health insurance coverage in the US is fragmented between public and private payers. The main public organisations are Medicare, Medicaid, and other public insurers like the Veterans Administration and TRICARE. Medicare is mainly available to certain disabled people and individuals aged sixty-five and above. Physician reimbursement is based on a fee schedule that is adjusted according to factors such as the type of service provided, where the service is performed, and geographical location. Hospitals are reimbursed according to the DRG methodology, which is a system that classifies hospital cases into groups with similar expected hospital resource use. Medicaid covers certain low-income individuals, and the programme is overseen by the Federal government and administered by the states. While Medicaid programmes have considerable leeway in determining methods of provider reimbursement, most states reimburse physicians via fee-for-service (adjusted by the specific service performed) or capitation. Hospital reimbursement methodologies include per diem rates, cost-based payments, DRG payments, and capitation.

Private insurance coverage encompasses group and non-group coverage and managed care and fee-for-service programmes. Even within these categories, physician and inpatient reimbursement methodologies may vary considerably. Physician reimbursement is generally fee-for-service or capitation, while inpatient reimbursement may be per diem, capitation, or case rates (including DRGs).

The rationale behind P4P

One disadvantage of the traditional reimbursement models is the lack of emphasis on quality. For instance, there may be overuse of medical treatment under fee-for-service, while there may be underuse under capitation or salaries. The US Institute of Medicine (IOM) highlighted the importance of provider reimbursement in a 2001 report, Crossing the Quality Chasm. The IOM argued that the quality of US health care was well below established benchmarks based on the best available evidence, and that payment mechanisms were an important contributor to the poor quality of care. Other researchers have indicated that adherence to evidence-based practices is variable across regions and providers, implying that US insurers and providers have considerable scope to implement quality improvements. In response, health plans and purchasers have put a considerable amount of effort into profiling providers and publicly reporting information on the quality of different providers, but evidence indicates that consumers fail to use this information when making health care decisions.

The nature of health care delivery,
whereby a physician acts as an agent for the patient and a third-party payer covers the majority of costs, implies that consumer choice alone may be insufficient for bringing about quality improvements. Pay-for-performance provides financial incentives to improve quality even when consumers are unresponsive to quality information. This reimbursement model is also expected to encourage ongoing quality improvements, for instance, through the installation of health information technology and more health care staff. Importantly, P4P may not reduce health care costs, but it is meant to ensure better value for money.

**P4P methodologies**

There are multiple ways in which models can be designed. The main features centre around the aspects of quality that are targeted, the allocation of the rewards, and whether rewards are based on quality improvements or target achievement. The targeted aspects of quality may be structure, process, outcome, or a combination of all three. Structure encompasses the resources needed to deliver care (for example, labour, facilities, and materials), process covers the completion of specific tasks or suggested procedures, and outcome includes the effect of treatment, mainly the patient’s experience and health status. Each aspect has its own advantages and disadvantages, but essentially the choice of which aspects to target influences the level of quality improvements, equity, and the amount of risk that providers experience.

Third-party payers can also choose whether to allocate rewards to individual providers or groups of providers. The difficulty in allocating rewards to individual providers is that some providers may have small patient groups leading to skewed outcomes, while incentives to improve quality may be muted in provider groups. The intuition is that within provider groups, some providers may ‘free-ride’ on the efforts of other providers. However, by rewarding provider groups, P4P programmes may create an incentive for coordinated care, for instance, through the creation of multi-disciplinary provider groups. Another important aspect of P4P is whether providers are rewarded for good performance or for improving quality. The former method rewards providers that are already high-quality but gives little incentive to low-quality providers to improve.

**The uptake of P4P in the US**

Pay-for-performance systems have diffused relatively rapidly throughout US health care insurers within the past decade, albeit on a small scale. Within private insurance markets more than half of health maintenance organisations (HMOs) have instituted P4P programmes, coverage that represents more than 80% of HMO enrollees in the US. Pay-for-performance is also growing in importance within the public insurance arena. By July 2006 more than half of the state Medicaid programmes had implemented pay-for-performance schedules, and it is estimated that by 2011, almost 85% of states will operate Medicaid P4P programmes. Medicare is also embracing P4P, although more detail on Medicare initiatives is given later in the article.

Despite the widespread use of P4P, these programmes typically make up a small proportion of provider reimbursement. Rosenthal and colleagues previously indicated that most payers only put 5% or less of provider compensation at risk of profit or loss from the P4P system.

**Examples of P4P in the US**

While there are hundreds of different P4P activities running, it is useful to discuss the main initiatives on the part of public and private payers. Medicare began running a P4P demonstration project jointly with Premier, Inc. (a nationwide organisation for non-profit hospitals) in 2003. The programme rewarded hospitals according to performance in five critical treatment areas for older people: acute myocardial infarction, heart failure, pneumonia, coronary artery bypass graft (CABG) surgery, and hip and knee replacement. The quality determinations were based on process and outcome measures. All hospitals were scored and ranked by treatment area. Hospitals in the top 10% of performers for each condition received a bonus of 2% of their Medicare payments. Hospitals in the next decile received a 1% bonus. The demonstration has ended, and Premier, Inc. and Medicare are currently discussing whether to extend the experiment. In 2007 Medicare also launched the Physician Quality Reporting Initiative, providing a 1.5% (of total allowed charges for eligible services) bonus for participating providers that successfully report the designated set of seventy-four quality indicators.

Pay-for-performance programmes within Medicaid are generally state-specific. However, nine Medicaid programmes have joined with other groups to improve pay-for-performance activities. A specific example is the Oregon Health Care Quality Corporation, a group of organisations (the state government, health plans, medical groups, insurers, purchasers, providers, and consumers) that are working together to incorporate standardised performance measures into P4P programmes. Interestingly, many Medicaid agencies are encouraging participation rather than performance in their P4P programmes; the intent is to incentivise providers to adopt technologies such as electronic health records and electronic prescribing.

In California, a group of health plans and physician groups developed the California Pay for Performance programme, which entails a set of quality performance measures, public report cards and financial incentives. The programme is now the largest of its kind in the US. Performance is measured on three main domains: clinical events (preventative measures and chronic care management), patient experience, and information technology investment. The California programme does not define the level of financial reward; instead this decision is left up to the individual participating health plans.

Another P4P programme backed by large employers is Bridges to Excellence (BTE), which aims to incentivise physicians in a number of target markets across the US to improve health care quality. BTE has four distinct initiatives: the Diabetes Care Link, the Cardiac Care Link, the Spine Care Link, and the Physician Office Link. Participants are awarded points for achieving quality measures within each of these links, and points are translated into financial awards that are specific to each link.

**Evidence from the literature**

Empirical evidence on P4P programmes is still relatively scant, but based on one review of the literature it is possible to draw preliminary conclusions regarding the outcomes and design of P4P programmes. This review indicated that most studies found partial or positive effects of P4P financial incentives on quality measures, whether the activity was at the individual physician or physician group level.

In addition, the design of incentives is important. A few studies have determined that documentation, as opposed to actual
use of preventative services, improved under P4P. While documentation is important, quality improvements also need to come from better use of services. The design of incentives may also have influenced risk selection in that providers may have avoided sicker patients where this was possible. Given that P4P programmes can target individual providers or provider groups, the authors of the review also indicated that the effect of P4P at the provider group-level is small, whereas the effect seems less muted for rewards at the individual physician level. Thus, it appears that some providers may ‘free-ride’ on the efforts of other providers within the group.

As discussed earlier, there is debate over whether providers should be rewarded for meeting benchmarks or for improving performance. Interestingly, providers with the lowest baseline performance may improve the most even if they receive the smallest amount of performance pay, implying that a P4P programme should consider incentives for both improvement and target achievement. The size of the reward may also influence whether P4P has any effect on the achievement of target indicators. Some studies that have found no relationship between the P4P programme and quality may have obtained this result because of small bonuses.

The review of P4P programmes found only one article that examined their cost-effectiveness. This study considered incentives to improve access to nursing homes and patient outcomes within the nursing home. This indicated that the P4P programme saved an estimated $3,000 per stay. However, not only is that study over fifteen years old, but the savings that the study found may not have accrued to the third-party payer. Thus, there is insufficient evidence available to draw any conclusions on the cost-effectiveness of P4P programmes.

**Pay-for-performance in Europe**

The P4P model has already made its way into Europe with the implementation of the large-scale Quality and Outcomes Framework (QOF) for primary care in the UK. Under QOF general practitioners are rewarded for chronic disease management, practice organisation, patient experience, and extra services (for example, child health and maternity services). Overall, Europe lags behind the US in implementation of P4P programmes, and thus a number of comments can be offered to European policy makers considering P4P. One important but unanswered question is whether P4P is cost effective. While literature on this topic is lacking, it seems prudent for European policy makers to determine if a P4P programme provides value for money in a pilot setting before implementing it on a large scale. Related to the costs of the programme, the reward structure that a third-party payer imposes is important. If rewards are insufficient, providers may not change their behaviour. While there are few guidelines available on the size of rewards, some suggest that 5% of the physician’s capitation income is necessary for behavioural change. In addition, policy makers must decide whether the reward is for meeting a benchmark, improving the influence of quality on outcomes, or both. The choice of the reward is related to whether the aim is to improve the performance of the lowest-quality providers, to maintain the performance of the highest-quality providers, or to achieve a combination of both.

Risk selection is also important to consider. If it is more difficult for providers to improve performance with certain types of patients (for example, sicker patients), then providers have an incentive to select healthier patients if this is possible. Since many European health care systems explicitly aim for equity, disincentives for risk selection should be built into any European P4P activities. England has addressed the possibility of risk selection by allowing providers to exempt reports on patients whose targets were more difficult to meet.

There is also the question of whether providers shift their focus to targets under the P4P system and pay insufficient attention to medical issues that are not part of the system. This may lead to a focus on the disease rather than on the patient. The question for policy makers is whether this influences the overall quality of care provided. There is also a consideration of whether it is possible to design a P4P system that encompasses more quality measures or whether this type of system would create too much of an administrative burden.

Overall, the success of a P4P programme appears to depend crucially on the programme design, as this influences the achievement of quality improvements. Importantly, there may be a trade-off between cost and quality, although more research on cost-effectiveness is needed to make any definitive statements regarding a trade-off. The variety of P4P measures that can be adopted also implies that European policy makers need to clearly define the aims of any P4P programme and adopt specific P4P measures to meet these aims.

**REFERENCES**