Financial incentive schemes in primary care

Stephen Gillam
Department of Public Health and Primary Care, Institute of Public Health, University of Cambridge, Cambridge, UK

Abstract: Pay-for-performance (P4P) schemes have become increasingly common in primary care, and this article reviews their impact. It is based primarily on existing systematic reviews. The evidence suggests that P4P schemes can change health professionals' behavior and improve recorded disease management of those clinical processes that are incentivized. P4P may narrow inequalities in performance comparing deprived with nondeprived areas. However, such schemes have unintended consequences. Whether P4P improves the patient experience, the outcomes of care or population health is less clear. These practical uncertainties mirror the ethical concerns of many clinicians that a reductionist approach to managing markers of chronic disease runs counter to the humanitarian values of family practice. The variation in P4P schemes between countries reflects different historical and organizational contexts. With so much uncertainty regarding the effects of P4P, policy makers are well advised to proceed carefully with the implementation of such schemes until and unless clearer evidence for their cost–benefit emerges.

Keywords: financial incentives, pay for performance, quality improvement, primary care

Introduction

There is widespread international interest in pay-for-performance (P4P) systems in primary care. Their introduction has aimed in particular to address well-attested variation in performance and quality of care.1 Primary care by its very nature is likely to encompass substantial variation in the delivery and organization of care (for example, with different staff mixes and training levels). This also reflects differing contextual factors (sociodemographic breakdown, geography, etc).2 Surveying the P4P systems developed in a number of countries, Schoen et al demonstrated variation in the degree to which financial incentives are employed to improve the quality of primary care. Doctors in the US (33%), Sweden (10%), and Norway (35%) were less likely to be receiving financial incentives as compared with doctors in countries such as the UK (89%), the Netherlands (81%), and New Zealand (80%).3

This paper examines the current evidence on P4P schemes and the extent to which such schemes can improve the quality of primary care services. For these purposes, I have borrowed on recent systematic reviews. A literature search was conducted in PubMed, using MeSH terms “Reimbursement, Incentive” or “Pay-for-performance” or “Financial incentive” and “Primary care or general practice”. Systematic reviews were filtered out to yield 34 studies of which seven were suitable for inclusion. Primary research is also quoted illustratively below. My own work has focused, in particular, on the Quality and Outcomes Framework (QOF) introduced to the UK in 2004, a scheme that has generated much of the recent research.4 We reviewed 575 articles in...
2012 of which some of the most informative were based on qualitative rather than quantitative analyses.

Before considering the evidence base, we need to consider what constitutes “quality” in primary care. What, after all, are financial incentives devised to improve?

**Defining quality in primary care**

There have been many attempts to define quality in primary care. The task is made more complex by the lack of any universal definition of primary care. As comparative international studies demonstrate, there are distinct differences between health systems as to what constitutes primary care and definitional complexities regarding terms such as family practice, general practice, primary medical, and primary health care.6,7 This paper draws heavily on experience in the UK where primary care is equated with general practice. However, care must be exercised in universalizing both definitions of quality and the impact of P4P.

Starfield famously identified four unique features of primary care service (four Cs): first contact access, person-focused care over time, comprehensiveness, and coordination.7 Hogg et al suggested that other important aspects of primary care include patient–provider relationships as defined by communication, holistic care, and an awareness of the patient’s family and culture.8 Primary care performance needs to be considered within the context of wider health care system and the organization of the practice itself.

There is also a distinction to be made between a quality health service and quality care for individuals. This point is particularly relevant in examining quality in primary care where Campbell et al have argued that individual care is more important than broader health system measures of quality. They offer their own definition of quality as “whether individuals can access the health structures and processes of care which they need and whether the care received is effective”.9 They focus on two aspects of individual care that they see as relevant to primary care, access and effectiveness, both of which are linked to improved health outcomes.9 However, too much emphasis on individual factors ignores key population aspects of primary care identified by Starfield et al.10

Definitions of quality include those of Donabedian11 and Maxwell12 and statements such as those by the Institute of Medicine, which defines quality as “the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge”.13 All recognize that quality is a complex and multidimensional concept.14

Wilson et al suggested four broad areas upon which general practice should be measured: equity, quality of clinical care, responsiveness to patients, and efficiency.15 Their review of practice suggested that UK practices score highly in all four domains although there have been recent concerns about a lack of support for self-care and for people with long-term conditions.16

In their review of outcome indicators, Sans-Corrales et al found that improved satisfaction and health outcomes were associated with continuity of care, patient-centered care, longer appointments, and a good patient–doctor personal relationship. These factors were also associated with lower overall health costs. Continuity of care is consistently reported as a key attribute and quality indicator of good primary medical care.17

Performance in P4P schemes tends to be measured in two main domains: aspects of health care delivery and the technical quality of clinical care.8 Most schemes focus on the latter but performance against process indicators should not be conflated with health outcomes.18 “Not all that is measurable is of value, and not all that is of value can be measured”.19 Of particular concern is the way P4P schemes tend to redefine how quality is conceptualized in practice. For example, several recent articles reporting on the UK experience appear to equate quality with the P4P criteria.20,21 For reasons that will become clearer as we consider the evidence base, this is potentially harmful.

**P4P improving quality – what is the evidence?**

In short, the evidence for the effectiveness of financial incentives is inconsistent.22 A recent Cochrane review of seven studies in primary care found that financial incentives were effective for some outcomes in some settings but concluded that there was “insufficient evidence to support or not support the use of financial incentives to improve the quality of primary health care”.23 Similarly, previous systematic reviews24,25 have concluded that P4P contracts do affect physician behavior and increase the range of primary care services provided but that their impact is often limited.26,27

The modest effects of financial incentives tend to be measured in terms of improvements in the processes of chronic disease management. Most P4P schemes focus on processes (such as measuring blood pressure) and intermediate outcomes (controlled blood pressure) for which there is either evidence or professional consensus and that can be measured and rewarded. Treatment and secondary prevention are thus
favored over primary prevention and less well-researched conditions can be marginalized.

The actual effect of financial incentives appears to depend on factors such as the age and sex of physicians, previous experience of financial incentives, the uptake of continuing professional education, the payment method, the type and severity of the conditions targeted through incentives, the volume of activity, and the location and type of organization.\textsuperscript{24}

Our review of the QOF found that quality of care for incentivized conditions during the 1st year of the framework improved at a faster rate than the preintervention trend but subsequently returned to prior rates of improvement.\textsuperscript{4}

A key concern that recurs in the literature is whether financial incentives generate dysfunctional physician behavior,\textsuperscript{28} negatively affect motivation,\textsuperscript{29} or aggravate “inverse care” effects.\textsuperscript{30} Research conducted in the USA found that the size and structure of incentives do seem to be important in promoting effective physician activity. However, incentives have to be large enough to influence behavior\textsuperscript{31} and designed in such a way that they cannot be “gamed”.\textsuperscript{32} The size of incentive may be less important in improving care processes than giving public recognition for scoring well on quality measures.\textsuperscript{33}

Some commentators have argued that there is a risk of undermining the morality and trust, which are central to the physician’s professional ethos.\textsuperscript{34} General practitioners (GPs) in the UK were anxious that “biomedical” targets might undermine holistic, continuity of care as treatment becomes increasingly divided within larger teams of health practitioners and as GPs offload routine tasks to nursing staff.\textsuperscript{35}

A further concern about externally structured incentives such as financial inducements is that they might “crowd out” professional self-esteem and a sense of self-determination.

On the other hand, one study in the UK found that an externally imposed system of incentives did not appear to damage the internal motivation of GPs. The authors attributed this to the fact that the indicators within the QOF aligned with what GPs themselves considered good clinical care objectives.\textsuperscript{27} Another study found that GPs felt that, while professional autonomy had decreased and workload increased, they were paid more and their job satisfaction levels had increased under the QOF.\textsuperscript{20} Nurses also report that their specialist skills have been enhanced.\textsuperscript{4}

Questions remain about the likely individual and population health gain from P4P schemes. An overview of four systematic reviews in health care found that none had examined the effect on patient outcomes.\textsuperscript{36} Evidence of physician activity does not always correspond to better health outcomes.\textsuperscript{25} The evidence of a relationship between incentive payments, physician activity, and health gain is weak.\textsuperscript{26} It is difficult to detect patterns from the diverse range of definitions of quality and the outcome measures used by researchers. The most commonly adopted measure – mortality – is unreliable because it is affected by many factors beyond the control of the physician or provider.\textsuperscript{25}

There is some evidence that P4P can reduce health inequalities resulting from socioeconomic disadvantage. For example, Doran et al found that the gap in median achievement comparing practices from the most deprived and least deprived quintiles in the UK narrowed from 4.0% to 0.8% between 2004 and 2007.\textsuperscript{37} On the other hand, achievements incentivized under the QOF have not reduced premature death in the population\textsuperscript{38} and inequalities have persisted.\textsuperscript{39}

Theoretically, the costs of the incentives might be repaid by reduced costs, but evidence on cost-effectiveness is limited.\textsuperscript{40} There is some evidence that emergency hospital admission rates (and hence costs) for ambulatory care sensitive conditions may have reduced following the introduction of P4P.\textsuperscript{41} Walker et al concluded that some QOF incentive payments were cost-effective, although they took no account of the costs of administering the scheme.\textsuperscript{42} On the other hand, Cecil et al linked an 8% rise in unplanned admissions of children usually managed by GPs to the QOF’s introduction. Less than 3% of the targets introduced in 2004 applied to children’s care, which was therefore unincentivized.\textsuperscript{43}

What does drive clinicians’ behavior?
Studies have paid insufficient attention to effect modifiers such as the nature and complexity of the target behavior, the size and method of the incentive, the health professional group being targeted, and the organizational environment. While incentives for individuals have been extensively examined, group rewards are less well understood.\textsuperscript{40} There is growing acceptance that financial rewards can undermine motivation and worsen performance of complex cognitive tasks.\textsuperscript{44} Economic research suggests that although financial incentives promote simple repetitive tasks, they can be counterproductive for tasks requiring more complex mental processes.\textsuperscript{45}

Financial incentives may encourage delivery of care that follows a simple algorithm, but algorithms are hard to apply meaningfully in the real world of individuals with unstructured symptoms and expectations. The complicated conceptual process of integrating suitable care for people with chronic conditions may not be enhanced by financial incentives. NICE (National Institute of Health and Care Excellence) guidelines may have limitations when applied to UK populations in
primary care.\textsuperscript{46} It may be possible to adapt guidelines to cater for people with multimorbidities – for example, through systematic cross-referencing\textsuperscript{47} – but they will always have shortcomings.\textsuperscript{48}

Money is only one of the many internal and external influences on clinical behavior, and many factors will moderate the size and direction of any response. The evidence on whether financial incentives are more effective than other interventions is weak. New P4P programs should incorporate research to examine the impact, downsides, and cost-effectiveness of incentives, and this should include evaluation of the comparative effectiveness of different strategies in different contexts. Few studies have gathered data on potential unintended consequences such as attention shift, gaming, and loss of motivation. Surprisingly little is known of patients’ views on P4P. Health professionals may perceive that the person-centeredness of consultations and continuity are negatively affected.\textsuperscript{49} Finally, such research should also include long-term follow-up, since behavior may revert when incentives are withdrawn.\textsuperscript{50,51}

### The future

Proponents of privatization and health care markets tend to assume that financial incentives improve outcomes. While many policy makers believe financial incentives can promote evidence-based changes to clinical practice, they have many limitations. Glasziou et al have proposed a checklist aimed at guiding implementation of P4P past some of these pitfalls (Figure 1).\textsuperscript{24} A decision to implement an incentive scheme should plainly include a critical assessment along these lines beforehand.

Future incentives should be balanced against sanctions for poor practice, as some evidence supports the use of penalties alongside rewards.\textsuperscript{52} In any event, the burgeoning research literature suggests that benefits to patient care are at best modest. With so much uncertainty regarding the effects of performance related pay, policy makers are well advised to contain expenditure on such programs.\textsuperscript{53,54}

### Disclosure

The author reports no conflicts of interest in this work.

### References


