Background
The recent release of The Royal Australian College of General Practitioners clinical quality indicators has sparked renewed debate about the role of pay-for-performance (P4P) programs and their potential usefulness in Australian general practice.

Objective
This article seeks to update recent evidence about the impact of P4P programs on the quality of primary care and presents the evidence-based viewpoint that there are potential problems with P4P, which may limit its usefulness.

Discussion
P4P programs are attractive to funders as they suggest a theoretical link between provider performance and improvements in healthcare quality—and potentially improved control over costs. Although P4P programs in primary care appear to have an effect on the behaviour of general practitioners, there is little evidence that these programs in their current form improve health outcomes or healthcare system quality. Further research is needed into the effect of these programs on healthcare quality before they are introduced into Australian general practice.

Keywords
general practice; quality of health care; health care economics
**Does P4P change provider behaviour?**

Early reports from the UK QOF suggested that healthcare workers did change their behaviour to fulfill requirements of these programs. Some studies have suggested a positive effect of P4P systems (particularly QOF) on risk factor recording and reporting, while others failed to find a significant effect. A recent systematic review for the Cochrane Collaboration, which looked at the effect of financial incentives on doctor behaviour, did find a significant change in behaviour in response to P4P. The authors found these programs effective in improving process measures, referrals and admissions and prescribing costs, but less effective in others areas such as adherence to guidelines or consultation rates.

**Does P4P improve quality of care or health outcomes?**

Initial results from the QOF suggested that these programs might be capable of improving quality of care. However, much of this early evidence failed to account for either the pre-existing trend for improvement in healthcare quality or other funding measures introduced at the time. When the trend of improving care was controlled for, there appeared to be an initial statistically significant improvement in management of some health conditions (diabetes and asthma) but not others (heart disease). However, this improvement appeared to be temporary: when this longitudinal study was continued for a further 2 years, it showed improvement reducing to earlier levels.

Recent systematic reviews found limited positive effect, mixed effects or no effect of these programs on improving care quality or health outcomes. Another recent Cochrane Collaboration review summarises current opinion. This review found an inconsistent effect of P4P incentives on healthcare quality and stated ‘there is insufficient evidence to support or not support the use of financial incentives to improve the quality of primary healthcare provision’. A retrospective review of administrative data collected during a physician P4P program in the USA suggested that financial incentives did not improve the quality of preventive care. The previously mentioned systematic review for the Cochrane Collaboration found that while there is evidence of financial incentives affecting the behaviour of healthcare providers, there is no evidence that they improve health outcomes.

**Possible undesirable effects of P4P programs**

A number of authors have expressed concerns about the potential unintended consequences of P4P systems, including around ‘gaming’ where practitioners manipulate results in order to maximise performance. In the QOF, there is some evidence that this type of manipulation has occurred. Other concerns regarding the QOF relate to ‘crowding out’ of the consultation where patient reason for consultation is minimised in order to complete prescribed P4P tasks, and ‘crowding in’ where care for incentivised diseases may dominate at the expense of other nonincentivised conditions. Concerns have also been raised about the impact of QOF on the nature of the GP consultation and dynamics of the primary care team, resulting in decreasing continuity of care and a loss of professionalism. In addition, two studies in the UK have concluded that the QOF has not contributed to any reduction in healthcare inequalities and there is evidence of QOF performance differing with ethnicity, patient age and being worse for nursing home residents.

**Not everything can be measured**

While P4P programs in primary care appear to have an effect on the behaviour of GPs, there is little evidence that these programs in their current form improve health outcomes or healthcare system quality. Possible explanations for this discrepancy include the concept that not everything can be measured. Current P4P programs focus on clinical and organisational measures, which may be relatively easy to measure through objective data or observation. However, there are other aspects of general practice that are less easily quantified and are only briefly considered in many P4P programs. These include continuity of care, ease of access to care, strength of the patient-doctor relationship and patient satisfaction. These aspects of care have been shown to be important for healthcare quality and also important to patients and providers. Importantly, by decreasing the emphasis on less easily measured aspects of care, P4P programs in their current form could have the potential to worsen overall care quality, as providers focus their activity on measureable outcomes and result in primary care moving toward achievement of performance indicators and away from holistic ‘general practice’.

A focus on valuing and measuring all dimensions of quality is becoming more important, as the results of these schemes are increasingly being used as a proxy for health system performance. For instance, in the UK, discussions about the quality of primary care focus heavily on the results in the QOF. It is important to remember that while these clinical measures are important, they do not provide a complete measure of care. Similarly, while the RACGP clinical indicators are helpful as a quality improvement guide, they are not able to measure all dimensions of healthcare quality.

**Different health systems: different effects**

It is important to note that before the introduction of the P4P program, the UK health system was a relatively underfunded capitation system with strict patient registration. The introduction of the P4P program resulted in a significant increase in overall funding and a pay increase for GPs. Introduction in a fee-for-service setting such as Australia is likely to be very different. Importantly, in most P4P programs (with the QOF being the notable exception) the magnitude of payment from P4P programs is less than 2% of practitioner income. This may provide insufficient incentive to change behaviour or to overcome the underlying incentives of the existing health system. In terms of evaluation, it is also very difficult to isolate the effect of P4P programs from the other incentives within a health system.

Importantly, with the emergence of Medicare Locals, there may be increasing need for regional and individual performance data in order to assist health service planning. Pay-for-performance programs could be used as an incentive to obtain this information from GPs. However, this is a very different focus for P4P than for quality improvement.

**Newer models**

Other payment models are emerging. One model of increasing focus in the US is based on the ‘patient centred medical home’. This model appears similar to capacity payments of the Australian PIP scheme and rewards practices that meet criteria associated...
with better outcomes (e.g., increased continuity of care or ease of access to providers). Pilots are underway using this model and early evidence suggests it may improve patient experience and care quality while controlling costs. 35

Another variation on the P4P model is ‘pay-for-participation’. Italian researchers found that paying providers to take responsibility for complete diabetic care reduced emergency hospitalisation, even without any monitoring. 36 This looks quite similar to the PIP sign-on payments. The long term effect of such payments needs further research.

Other authors have suggested that quality could be encouraged by paying GPs to focus on ethical or more aspirational attributes (honesty, judgement), so-called ‘pay-for-virtue’. 37

Conclusion

Pay-for-performance programs are attractive to funders as they provide increased transparency and evidence of activity. However, while P4P programs in primary care appear to have an effect on the behaviour of GPs, there is little evidence that these programs in their current form improve health outcomes or healthcare system quality. In addition, these programs may lead to undesired consequences, which need to be considered before their introduction. Before such programs are introduced in Australia, more research is needed to see if payment for broader measures of quality (such as continuity of care) results in improvement in the quality of care or health outcomes. New models of primary care are emerging and may improve quality of care more effectively than P4P.

Author

Michael Wright MBBs, FRACGP, is Visiting Research Fellow, London School of Hygiene and Tropical Medicine, London, United Kingdom. wrighthdmc@gmail.com.

Conflict of interest: none declared.

References