

ANALYSIS



Can pay for performance improve the quality of primary care?

 OPEN ACCESS

Martin Roland and **Frede Olesen** explore what other countries can learn from the UK's experience with the Quality and Outcomes Framework

Martin Roland *professor of health services research*¹, Frede Olesen *professor*²

¹Institute of Public Health, Cambridge CB2 0SR, UK; ²Department of Public Health, 8000 Aarhus, Denmark

For many conditions and in many countries, medical care falls short of what should be provided. Although high quality care depends crucially on the professionalism of doctors, this has not proved sufficient to ensure universal high quality care, and as a consequence payers have looked towards incentives as one way of further improving quality. In 2004 the UK National Health Service introduced the world's largest healthcare related pay-for-performance scheme into primary care, the Quality and Outcomes Framework (QOF). Primary care doctors were paid up to 25% more if they met a complex set of clinical and organisational indicators.¹ The scheme was a national one with no opportunity for local variations to tackle particular local healthcare needs.

Although the scheme was initially popular, general practitioners (GPs) have become increasingly disenchanted, especially with the substantial administrative workload at a time of rising clinical demand, and there are concerns about reduced time for aspects of care that are not readily measured. Parts of England have been allowed to opt out of the scheme, and Scotland is dropping QOF in favour of quality circles (groups of GPs working together to improve quality).² What can other countries learn from the UK's experience of pay for performance? We set out some principles and caveats for the development and introduction of pay for performance as a tool for quality improvement.

How should quality of care be measured?

The first problem in developing any pay-for-performance scheme is to decide how to measure quality of care. The development of clinical indicators is itself a science. Considerable skill is needed to develop clear reliable and valid indicators with clinical relevance that also meet administrative or managerial needs, and there are hundreds of possible quality indicators (the Brazilian pay-for-performance scheme in primary care has over 1000). However, if the scheme is too large and cumbersome it will be complex to administer and difficult for

clinicians to understand—a factor which itself can be demotivating.³ In contrast, a scheme with a small number of indicators may lead doctors to prioritise a limited range of aspects of care. Quality improvement schemes that depend on measurement will inevitably tend to prioritise aspects of care that are easy to measure, potentially at the expense of care which is equally or more important.

The UK's scheme started with a focus on clinical care, organisational aspects of care, and patient experience (box 1). The clinical indicators developed and expanded to include more conditions over the years, but the organisational indicators were dropped in 2012 when virtually all practices scored 100%. The indicators for patient experience were changed several times but were never really successfully implemented (see examples in box 2).

There is no optimum size for a pay-for-performance scheme; it will depend on the nature and scale of the challenges to be addressed. For example, the UK's first scheme providing incentives for childhood immunisation and cervical cytology coverage had just three indicators, was effective in maintaining population coverage, and ran for over 10 years from 1991 without the introduction of any other clinical targets.⁴

Who should decide on the indicators?

Clinicians whose performance is to be assessed should be involved in selecting indicators, using systematic appraisal of evidence and working with experts in technical construction of valid and reliable indicators. The indicators themselves should have strong face validity, either through an evidence base or widespread professional consensus. The UK's QOF ran into problems in the late 2000s partly as a result of indicators being introduced that had little professional consensus (eg, completing a symptom inventory for patients with depression), seemed to have a managerial rather than a clinical agenda (eg, incentives designed to reduce emergency admissions), or had poorly

Box 1: How the Quality and Outcomes Framework works

- The original scheme included 76 clinical indicators covering 10 conditions
- Data on clinical quality were extracted automatically from practice electronic records
- Doctors could exclude patients from individual clinical indicators (exception reporting) for specified reasons including clinical inappropriateness, intolerance of medication, and patient dissent
- Organisational indicators included medical records, information for patients, education and training, practice management, and medicines management
- Patient experience indicators related to conducting and acting on the results of patient experience surveys and offering booked appointments of at least 10 minutes

Box 2: Examples of QOF indicators*Clinical process indicators*

- Percentage of patients with chronic obstructive pulmonary disease who have had a review in the preceding 12 months, including an assessment of breathlessness using the MRC dyspnoea scale
- Percentage of patients on lithium therapy with a record of serum creatinine and thyroid stimulating hormone in the preceding 9 months

Clinical intermediate outcomes

- Percentage of patients with coronary heart disease in whom the last blood pressure reading (measured in preceding 12 months) is $\leq 150/90$ mm Hg
- Percentage of patients with diabetes whose last measured total cholesterol (measured within the preceding 12 months) is ≤ 5 mmol/L

Organisational indicators

- The practice has up-to-date clinical summaries in at least 70% of patient records
- A record of all clinical staff having attended training in basic life support skills in the preceding 18 months

Patient experience indicators

- Percentage of patients who indicate in a national survey that they were able to obtain a consultation with a GP within two working days (England only)
- Practice will have undertaken an approved patient survey each year and, having reflected on the results, will produce an action plan that sets priorities for the next two years, describes how the practice will report the findings to patients, and describes the plans for achieving the priorities

constructed formulas linking performance to pay (eg, data from patient surveys). The result was that GPs lost confidence in the scheme.

Should doctors be able to exclude patients?

It may be impossible to meet the requirements of quality indicators if some patients decline treatment or follow-up. Treatments or management suggested by clinical guidelines may also be clinically inappropriate for individual patients. QOF allows GPs to exclude (exception report) patients on the basis of their clinical judgment or because patients fail to attend for follow-up appointments. This was an important part of getting buy-in from UK family doctors and reduces the risk of doctors being given incentives to manage the patient in a way they believe to be inappropriate.⁵ An average of around 5% of patients are excluded from individual QOF indicators by their GP, though the figure varies widely between indicators.⁶ If doctors are allowed to exclude patients, this should be monitored, with doctors being prepared to justify the decisions made.

Is performance a useful way of paying doctors?

The experience of pay-for-performance schemes across the world is that they have less effect than payers hope for, and the schemes should therefore be seen as one part of a wider quality improvement strategy.⁷ However, no payment system is perfect and all have advantages and potential disadvantages (table). This fact lies behind attempts to develop “blended” payment systems designed to provide population coverage and good access without compromising quality of care.

Pay for performance will also not generally be suitable for improving care in patients with more complex problems, or the wider determinants of personal health and health in society. For important aspects of care that are not easily measurable (eg, for patients with mental health problems and those with multimorbidity and complex conditions), better mechanisms are needed to ensure high quality care. Potential methods include allowance for the additional time taken to provide good care for such patients (as occurs in the Australian health service) and encouraging doctors to engage more widely with the health and social needs of patients—for example, through social prescribing schemes.⁸

The formula linking performance to pay needs to be simple and transparent and should encourage improvement at all levels up to an agreed maximum threshold, ensuring that low performing practices have an incentive to improve. “Tournament” schemes providing payments for the highest scoring practices or groups (eg, those performing in the top 10%) should be avoided as they inappropriately penalise doctors if all practices reach high standards of care (eg, if all practices meet 95% on a particular quality marker). They also don’t provide incentives when all providers need to improve or allow providers to plan and budget ahead as the reward is dependent on the performance of other providers.

Good data recording and reporting are essential for any pay-for-performance scheme, and one positive effect of QOF has been the development of comprehensive electronic records and explicit collection and use of data for measuring quality. This can be used to educate doctors and increase public transparency about quality of care. However, it is not clear how much payments add to feedback and public reporting of performance in improving quality. Although small financial rewards are probably less effective than large ones, the UK’s

experience is that GPs have tried hard to get maximal points, sometimes expending effort that was clearly beyond any substantial benefit for themselves or their patients. Clinical ambitions and reputations are important to doctors, and public reporting of data on quality of care may itself act as a driver of performance.

How often should schemes be revised or changed?

The UK's QOF was changed around every two years. Some changes were minor (eg, changing the thresholds of individual indicators) and some major (eg, dropping the organisational indicators entirely). Experience from the UK suggests that substantial changes should not occur more frequently than every 2-3 years in order to keep motivation among doctors. However, there is also an argument for rotating conditions or indicators to allow a wider spectrum of clinical conditions to be included. In the UK, there has been constant pressure from patient and professional groups to include more conditions in QOF as this is seen as an effective way of gaining the attention of GPs.

A key question is when indicators that have been effective in improving quality should be dropped. If pay for performance is intended to improve quality, should payments be stopped when maximum achievement on the indicators has been sustained for some years? Or will quality then decline? The answer is unclear.^{9 10} One solution may be to reduce quality payments as indicators are rotated or withdrawn but require limited ongoing data collection for a reduced set of indicators, and at the same time require (and maybe incentivise) continued and transparent data collection, potentially including areas where there are no longer incentives for quality improvement.

What about unintended consequences?

Any incentive scheme, financial or otherwise, can have unintended consequences. These can range from neglect of conditions that are not included to gaming or outright fraud. Policy makers and payers must understand the potential for unintended or unexpected consequences, anticipate them as far as possible, and be vigilant for them once a scheme is introduced. Some of the QOF indicators that produced unintended consequences are described in a linked paper.¹¹ Larger incentives carry greater risk of perverse outcomes, and the UK decided that deriving 25% of GPs' income from QOF was too much; it put too much focus on performance targets, reduced the importance of the patient's agenda, and potentially reduced patient choices about the treatments they should receive. Remuneration from QOF was reduced to around 15% of income in 2013 and is likely to be reduced further. Box 3 sets out good design principles that can help avoid adverse consequences.

Conclusions

There is no perfect system of paying for medical care, nor one without potential unintended consequences, and there is an argument for including some element of pay for performance in medical care. However, the risks and unintended consequences always need to be taken into account, as does the organisational context into which incentives are introduced—for example, whether to reward the individual, team, clinic, or groups of practices; how that is likely to affect the behaviour of individual clinicians; and how it will lead to improved care. Research is needed on more sophisticated ways of using indicators to improve quality of care and to evaluate the introduction of pay for performance in different settings.¹²

Competing interests: We have read and understood BMJ policy on declaration of interests and declare that MR advised the BMA and NHS Employers on the development of the Quality and Outcomes Framework from 2001 to 2003.

Contributors and sources: Both authors contributed to the development and authorship of this article. MR is the guarantor.

- 1 Roland M. Linking physicians' pay to the quality of care—a major experiment in the United Kingdom. *N Engl J Med* 2004;351:1448-54. doi:10.1056/NEJMhpr041294pmid:15459308.
- 2 Beyer M, Gerlach FM, Flies U, et al. The development of quality circles/peer review groups as a method of quality improvement in Europe. Results of a survey in 26 European countries. *Fam Pract* 2003;20:443-51. doi:10.1093/fampra/cmg420pmid:12876119.
- 3 McDonald R, Harrison S, Checkland K, Campbell SM, Roland M. Impact of financial incentives on clinical autonomy and internal motivation in primary care: ethnographic study. *BMJ* 2007;334:1357-9. doi:10.1136/bmj.39238.890810.BEpmid:17580318.
- 4 Baker D, Middleton E. Cervical screening and health inequality in England in the 1990s. *J Epidemiol Community Health* 2003;57:417-23. doi:10.1136/jech.57.6.417pmid:12775786.
- 5 Roland M. Should doctors be able to exclude patients from pay-for-performance schemes? *BMJ Qual Saf* 2015;30:bmjqs-2015-005003. doi:10.1136/bmjqs-2015-005003pmid:26717988.
- 6 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. doi:10.1056/NEJMsa0800310pmid:18635432.
- 7 Roland M, Dudley RA. How financial and reputational incentives can be used to improve medical care. *Health Serv Res* 2015;50(Suppl 2):2090-115. doi:10.1111/1475-6773.12419pmid:26573887.
- 8 University of Westminster. Report of the annual Social Prescribing Network conference. 2016. www.westminster.ac.uk/file/52171/download
- 9 Lester H, Schmittiel J, Selby J, et al. The impact of removing financial incentives from clinical quality indicators: longitudinal analysis of four Kaiser Permanente indicators. *BMJ* 2010;340:c1898. doi:10.1136/bmj.c1898pmid:20460330.
- 10 Dreischulte T, Donnan P, Grant A, Hapca A, McCowan C, Guthrie B. Safer prescribing—a trial of education, informatics, and financial incentives. *N Engl J Med* 2016;374:1053-64. doi:10.1056/NEJMsa1508955pmid:26981935.
- 11 Roland M, Guthrie B. Quality and Outcomes Framework: what have we learnt? *BMJ* 2016;354:i4060.
- 12 Kristensen SR, Meacock R, Turner AJ, et al. Long-term effect of hospital pay for performance on mortality in England. *N Engl J Med* 2014;371:540-8. doi:10.1056/NEJMoa1400962pmid:25099578.

Published by the BMJ Publishing Group Limited. For permission to use (where not already granted under a licence) please go to <http://group.bmj.com/group/rights-licensing/permissions>

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

Box 3: Design principles for pay for performance schemes

- Pay for performance should be seen as part of wider quality improvement efforts.
- Public reporting of information on quality of care may be an effective driver of change without pay for performance
- Alternative strategies should be used to improve quality for aspects of care not easily measured to avoid them being neglected
- Single disease indicators may not be appropriate for important patient groups such as complex patients with multimorbidity
- Clinicians should be strongly represented among those selecting indicators and designing the programme
- Technical expertise in developing and implementing indicators is needed so that they measure what they are intended to measure and reduce unintended consequences
- Indicators should represent aspects of quality that command wide professional support and, where available, be based on strong scientific evidence
- Payments should be large enough to change behaviour but not so large as to divert excessive effort onto incentivised aspects of care
- The payment formula should encourage improvement at all levels up to an agreed maximum, ensuring that low performing providers have an incentive to improve
- Unexpected consequences should be anticipated and continuously monitored
- The effect on inequalities in delivery of care should be monitored

Key messages

Pay for performance can be a useful part of wider programmes to improve quality of care, though quality gains from financial incentives are generally modest

Clinicians must be closely involved in the development of quality indicators and pay-for-performance programmes

All incentive schemes (financial and non-financial) have the potential to produce unexpected or perverse outcomes, and these should be anticipated and continuously monitored

Table

Table 1 | Intended and unintended consequences of four methods of paying doctors

Payment method	Intended outcomes	Potential unintended consequences
Salary: pay independent of workload or quality	Provides a basic secure income, may be especially important in hard to serve areas.	Provides no incentive to provide population coverage, improve access, provide high quality care, or work beyond contracted hours or duties
Capitation: pay according to number of people on a doctor's list	Provides coverage for the whole population: promotes equity.	Provides no incentive to provide good access if patients are technically registered, or to improve quality of care
Fee for service: pay for individual items of care	Prioritises meeting demand	Tendency to over-provide, whether or not care is needed
Pay for performance: pay for meeting quality targets	Improvement in quality of care	Neglect of aspects of care not measured or incentivised