HEALTH POLICY REPORT

Mary Beth Hamel, M.D., M.P.H., *Editor*

Successes and Failures of Pay for Performance in the United Kingdom

Martin Roland, D.M., and Stephen Campbell, Ph.D.

In 2004, the United Kingdom introduced one of the world's largest pay-for-performance programs, the Quality and Outcomes Framework, with a quarter of the income of family practitioners tied to measures of their performance.[1] Since then, much research has been conducted to answer the question "Does pay for performance improve the quality of health care?" This research has included more than 20 systematic reviews and now 1 systematic review of systematic reviews.[2] As a summary of this increasing body of work, it is clear that pay for performance can be effective. However, the effects are sometimes only short-term and are often not as large as payers wish. The effect of incentives is dependent on the context in which they are introduced, and pay-for-performance programs always have the potential to produce unintended consequences. This leads to many unanswered questions about how such programs can best be introduced[3] and suggestions from a range of countries on the key design issues that need to be considered.[4-7]

There has been a growing awareness in the United Kingdom among both professionals and politicians that the Quality and Outcomes Framework needs major change, and some changes have recently been adopted.[8] On the basis, in part, of our experience in advising successive governments on the design and content of the program, we describe some of its successes and failures.

The success or failure of a pay-for-performance program is predicated on its clarity of purpose and criteria for assessing performance.[9] The Quality and Outcomes Framework was originally designed in part to give family practitioners a substantial pay increase, since government officials had agreed that this was necessary to boost recruitment in primary care. The profession offered quality in return. However, the amount of program-associated money (25% of family practitioners' income) became increasingly regarded as a distraction, diverting their gaze onto limited parts of clinical practice and reducing the focus on the patient's agenda during the consultation. Changes were recently adopted to reduce the pay-for-performance element of family practitioners' income by around a third and to redistribute that money to other mechanisms of payment — principally, capitation payments. This change has been welcomed by many practitioners, since it recognizes that much clinical work cannot be measured and that incentives have the potential for unintended consequences. All payment mechanisms (e.g., salary, fee for service, and pay for performance) have drawbacks,[10] and many payers are looking to blended payment systems to balance the advantages and disadvantages of various ways of paying physicians. As the percentage of physicians' pay that is tied to performance increases (e.g., above 10%), the effect of the program is likely to increase, but so are the risks of unexpected or perverse consequences.

It is clear that financial incentives change doctors' behavior. In 2004, when the Quality and Outcomes Framework was introduced, much changed overnight. Family practitioners and practice staff started using full electronic medical records because without these systems, payments could not be made. They also changed the structure and staffing of their practices in two key respects. First, there was an increase in nursing staff, with the management of major chronic diseases such as diabetes increasingly moved out of regular response-mode consultations into nurse-run, protocol-driven clinics. Second, there was an increase in administrative staff so that family practitioners could have rap-

id access to data on their performance. These changes were already under way, with many family practitioners realizing that the systematic monitoring of chronic diseases needed to be more organized. Over time, however, the program became more intrusive into regular consultations with family practitioners. This was partly because the number of conditions in the framework increased to such an extent that nurse-led clinics could no longer be used for all the conditions that were covered. Moreover, family practitioners resented constant electronic reminders of "boxes to be ticked," which led to a more biomedical focus in consultations with less attention being paid to patients' concerns.[11,12] Table 1 summarizes the key features of the original Quality and Outcomes Framework and subsequent changes.

CLINICAL INDICATORS

Clinical care probably improved after the introduction of the Quality and Outcomes Framework,[13] though the effects were not compelling and were difficult to disentangle from other ongoing quality-improvement initiatives (e.g., national guidelines and public release of information on quality of care). For major chronic conditions, it seemed that the incentives maintained or increased preexisting trends in quality improvement but that the effect plateaued as physicians gained the maximum rewards available (Fig. 1A).

The program may have had a greater effect in clinical areas that were not previously measured and were low priorities for family practitioners before the program was initiated (e.g., annual review of patients with a learning disability). However, for the major chronic conditions, such as diabetes and heart disease, financial incentives should be seen as one part of a wider strategy of quality improvement. This philosophy is consistent with the general literature on quality improvement, which suggests that there is no "magic bullet" but that multiple interventions that are sustained over time can produce major improvements in care.[14]

Not all new indicators had professional support. One example shows the challenges of introducing an indicator that was not widely accepted as a legitimate part of practice. This indicator was the requirement that family practitioners who were seeing a patient with depression should record the severity of depression by means of a standard scale (e.g., the Patient Health Questionnaire 9). The rationale for this requirement was based on research evidence that family practitioners often underestimate

Table 1. Summary of Key Features of the Quality and Outcomes Framework, 2004–2014.

Original Program (2004)	Changes between 2005 and 2013	Changes for 2014–2015
Clinical indicators		
Clinical indicators covered coronary heart disease, heart failure, stroke and transient ischemic attack, hypertension, diabetes, chronic obstructive pulmonary disease, epilepsy, cancer, mental health, hypothyroidism, and asthma. Payments were on a sliding scale between upper and lower performance thresholds.	Clinical indicators were updated in 2006, 2009, and 2011, with a progressive increase in both upper and lower performance thresholds for established indicators. Included was the addition of new conditions, such as atrial fibrillation, peripheral artery disease, dementia, depression, chronic kidney disease, learning disability, osteoporosis, rheumatoid arthritis, and palliative care.	The proportion of family practitioners' income accounted for by the pay-for-performance program would be reduced by a third.
Organizational indicators		
Organizational indicators covered records and data, information for patients, education and training, practice management, and management of medications.	In 2012, organizational indicators were dropped and replaced in part by public health indicators (e.g., obesity, smoking, and sexual health).	Indicators requiring practices to review data on specialist referrals and unscheduled hospital admissions would continue to be developed, along with proactive case management for vulnerable elderly patients.
Patient experience indicators		
The reporting of patient experience was rewarded by encouraging family practitioners to engage with patients to review survey results.	A short-lived attempt to pay family practitioners according to survey results was replaced by incentives to establish formal patient groups to review survey results.	None

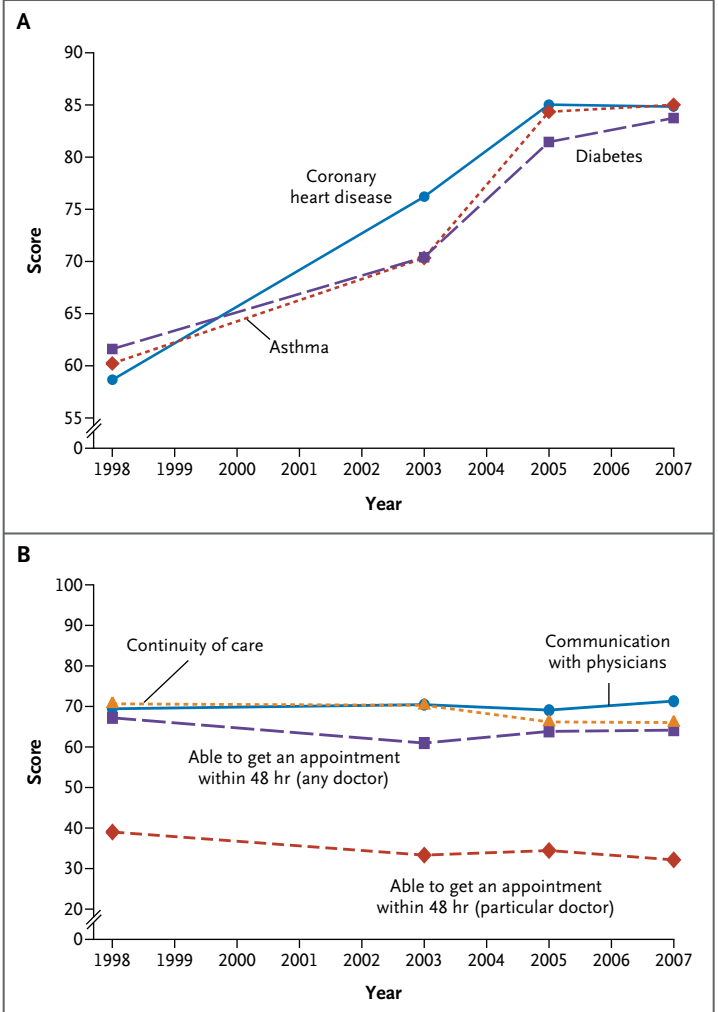

Figure 1. Mean Scores for Quality of Care in the United Kingdom, 1998–2007.

Panel A shows scores for the quality of care provided for coronary heart disease, asthma, and diabetes in a representative sample of practices (studied only in England) before and after the introduction of the Quality and Outcomes Framework, a national pay-for-performance program, in 2004. Quality scores range from 0% (no quality indicator was met for any patient) to 100% (all quality indicators were met for all patients). Panel B shows scores for patients' perceptions of communication with physicians, access to care, and continuity of care. Communication was assessed by asking seven questions, with the answers scored on a six-point scale ranging from "very poor" to "excellent." Continuity of care was assessed with the use of the same six-point scale and a single question: "How often do you see your usual doctor?" Access to care was scored as the percentage of patients who reported that they were able to get an appointment within 48 hours. All scores were rescaled to range from 0 to 100. Data are from Campbell et al.[13]

the severity of depression. This indicator was unpopular with family practitioners, partly because administering a questionnaire during a visit with a new patient with depression often seemed to intrude into difficult and sensitive consultations.[15] After some years, the indicator was dropped and was replaced by one that required physicians to record that they had completed a "biopsychosocial assessment," without any specification of what such an assessment should contain. This meaningless indicator has also been dropped in the most recent changes to the framework. The message from this experience is that not everything that is important can be measured, and indicators should not be forced onto aspects of practice that are not easily measured — for example, mental health care and care of the frail elderly with multiple chronic conditions.

ORGANIZATIONAL INDICATORS

Organizational indicators — which include records and data about patients, information for patients, education and training, practice management, and management of medications — were introduced in the original framework but were never a great success. Although these indicators contained some laudable incentives (e.g., annual appraisal of all staff and the requirement to document improvements in prescribing practices), they were easy to achieve. Since almost all family practitioners got the maximal incentive payments from the start, there was little room for improvement. Organizational indicators were removed in 2012. The focus of these indicators has now shifted to emphasize indicators that reward practices for reviewing data on specialist referrals and unscheduled hospital admissions and for proactive case management for vulnerable elderly patients. The benefits of these changes remain to be seen, but they are promising because they are designed to focus attention on frail patients with multiple coexisting conditions for whom quality indicators that are developed for single conditions often appear to be inappropriate and miss the real needs of this increasingly important group of patients.[16,17]

PATIENT EXPERIENCE INDICATORS

The initial introduction of patient experience into the Quality and Outcomes Framework gave family practitioners a substantial reward for doing little in this area — conducting surveys in their practices and engaging with patients in reviewing the survey results. At about that time,

government administrators became concerned that access to primary care was getting worse. Problems in access were in part an unintended consequence of a previous quality indicator that gave family practitioners an incentive to offer appointments to patients within 48 hours. Family practitioners had responded to this incentive by introducing "advanced access,"[18,19] which offered appointments that could be made on the same day but reduced the availability of future appointments, thus making it harder for patients to book ahead.[20] The result was that the government target had been met, but paradoxically patients found it harder to make appointments (Fig. 1B). The attempt to fix this problem went seriously wrong by introducing a large national patient survey in order to pay family practitioners on the basis of the reported ease of getting appointments.[21] Doctors were suspicious of the surveys, and many did not regard them as valid measures of quality.[22] To add to the problem, a misjudgment was made in the formula linking payments to survey scores, which resulted in the translation of large, random, year-to-year variation in scores into large, random, year-to-year variation in payments.[23] This meant that a practice that had made considerable efforts to improve access could find that payments for the indicator were reduced.

There are substantial problems with linking patient-experience scores directly to physicians' pay,[24] and this unpopular indicator was dropped in 2011. Despite the failure of this indicator, the attempt to provide incentives in practices to establish patient groups and to engage with them about improving care may have started to change relationships between practitioners and their patients.[25]

UNINTENDED CONSEQUENCES

Unintended consequences are high on the list of criticisms of pay for performance,[26,27] and the risks of perverse or unintended consequences associated with the publication of performance data are well known.[28] Pay-for-performance programs can reward only what can be measured and attributed, a limitation that can lead to less holistic care and inappropriate concentration of the doctor's gaze on what can be measured rather than on what is important.[11,29] There is some evidence that, as in a previous incentive program in the United Kingdom,[30] the Quality and Outcomes Framework has led to some adverse effects on the quality of care for medical conditions that are not included in the incentive program.[31] To some extent, this is an inevitable consequence of the diversion of clinical attention toward some conditions and, by implication, away from others. Although this change in focus is a common source of complaints, it would not necessarily be a bad thing if the focus of clinicians was diverted to conditions that had the greatest potential for health gain. However, such a balance is hard to strike.

From the start of the program, many observers had been concerned that physicians would "game" the system to maximize their incomes. It is likely that this fear has been realized to some extent,[32] since all the information from the Quality and Outcomes Framework is publicly available, physicians care about their reputations, and family practitioners expend substantial energy in getting very high scores. However, the abuse of exception reporting (which allows family practitioners to exclude patients from individual or groups of indicators) has not been as widespread as administrators feared.[33-35] Physicians remain convinced that it is important for them to exclude patients from indicators that they regard as inappropriate for individual patients. This flexibility is a key part of maintaining professional buy-in and removing incentives for family practitioners to disenroll problematic patients ("cherry picking") or to provide treatment that the practitioner judges is not in the patient's interest. The latter can be particularly damaging to professional motivation and morale.

PROFESSIONALISM AND INTERNAL MOTIVATION

Throughout the 10 years that the pay-for-performance program has been in operation, doctors have been ambivalent about it. They welcomed some aspects of the program in which the framework reinforced commonly accepted medical standards, but they were concerned about a loss of autonomy and professionalism and becoming less skilled in dealing with certain conditions, such as diabetes, since nurses were becoming more involved in chronic disease management.[36,37] Although most family practitioners welcomed the initial pay increase, many

Table 2. Lessons from the United Kingdom on Pay for Performance.

Pay for performance can be used to improve the quality of care, but it is not a "magic bullet" and needs to be combined with other quality-improvement initiatives to produce sustained improvements.
Aligning financial incentives with professional values may reduce the risk of unintended consequences, including gaming.
Pay-for-performance administrators need to recognize that large parts of clinical practice cannot currently be measured. It is better to recognize this than to force poorly designed indicators into a program.
Physicians care about their reputations. Public reporting of information on quality of care is often introduced at the same time as pay-for-performance programs and may be an important driver of behavior change.
Single-condition indicators do not adequately meet the needs of elderly patients with multiple coexisting medical conditions. Newer indicators attempt to address the quality of care for this increasingly important population.
Attaching 25% of income to pay for performance resulted in a major focus of family practitioners' attention on limited areas of clinical practice. A proposed redistribution of income that reduces this percentage has been widely welcomed.

of them began to resent the program as successive governments clawed back the initial large increases with a succession of below-inflation raises. The program was also resented by many nurses and salaried doctors who saw themselves contributing to income generation but not sharing in the benefits.[29,38]

There is clearly a problem in trying to include more and more conditions into a pay-for-performance program, especially when the overall money that is available remains the same. Progressively, the burden of the recording of data mounts, with consultations becoming increasingly disrupted by the need to respond to requests or prompts for information. Improvements in care that follow the introduction of an incentive may not be long lasting, and there is some evidence that quality of care may decline when an incentive is removed.[39] The most recent changes that have been adopted for 2014–2015 indicate some reining in of the program, with a reduction by a third in the proportion of family practitioners' income that is associated with pay for performance. This change has been generally welcomed.

Some of the lessons from 10 years' experience of pay for performance in the United Kingdom are summarized in Table 2. There is a case for including pay for performance as part of physicians' payment, not least because of the problems associated with all other payment systems. Thus, the choice should not be whether to adopt a pay-for-performance program but rather should be which type of program to use in combination with which other quality-improvement interventions.

Dr. Roland reports providing advice to the U.K. government and British Medical Association on the design and content of the original Quality and Outcomes Framework in 2002 and 2003; and Dr. Campbell, setting up and leading the external contractor that advised the National Institute for Health and Care Excellence (NICE) on the development and pilot testing of indicators for the Quality and Outcomes Framework from 2009 through 2012.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

From the University of Cambridge, Cambridge (M.R.), and the University of Manchester, Manchester (S.C.) — both in the United Kingdom.

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.
2. Eijkenaar F, Emmert M, Scheppach M, Schöffski O. Effects of pay for performance in health care: a systematic review of systematic reviews. Health Policy 2013;110:115-30.
3. Roland M. Pay-for-performance: not a magic bullet. Ann Intern Med 2012;157:912-3.
4. Dudley RA. Pay-for-performance research: how to learn what clinicians and policy makers need to know. JAMA 2005; 294:1821-3.
5. Van Herck P, De Smedt D, Annemans L, Remmen R, Rosenthal MB, Sermeus W. Systematic review: effects, design choices, and context of pay-for-performance in health care. BMC Health Serv Res 2010;10:247.
6. Campbell SM, Scott A, Parker RM, et al. Implementing pay-for-performance in Australian primary care: lessons from the United Kingdom and the United States. Med J Aust 2010;193:408-11.
7. Glasziou PP, Buchan H, Del Mar C, et al. When financial incentives do more good than harm: a checklist. BMJ 2012;345: e5047.
8. NHS Employers. Changes to QOF 2014/15 (http://www.nhsemployers.org/PayAndContracts/GeneralMedicalServicesContract/QOF/Pages/ChangestoQOF201415.aspx).
9. Campbell SM, Braspenning J, Hutchinson A, Marshall MN. Research methods used in developing and applying quality indicators in primary care. BMJ 2003;326:816-9.
10. Roland M. Incentives must be closely aligned to professional values. BMJ 2012;345:e5982.
11. Maisey S, Steel N, Marsh R, Gillam S, Fleetcroft R, Howe A. Effects of payment for performance in primary care: qualitative interview study. J Health Serv Res Policy 2008;13:133-9.
12. Checkland K, Harrison S, McDonald R, Grant S, Campbell S, Guthrie B. Biomedicine, holism and general medical practice: responses to the 2004 General Practitioner contract. Sociol Health Illn 2008;30:788-803.
13. Campbell SM, Reeves D, Kontopantelis E, Sibbald B, Roland M. Effects of pay for performance on the quality of primary care in England. N Engl J Med 2009;361:368-78.
14. Grimshaw JM, Shirran L, Thomas R, et al. Changing provider behavior: an overview of systematic reviews of interventions. Med Care 2001;39:Suppl 2:II-2–II-45.
15. Leydon GM, Dowrick CF, McBride AS, et al. Questionnaire severity measures for depression: a threat to the doctor-patient relationship? Br J Gen Pract 2011;61:117-23.
16. Boyd CM, Darer J, Boult C, Fried LP, Boult L, Wu AW. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. JAMA 2005;294:716-24.

17. Guthrie B, Payne K, Alderson P, McMurdo ME, Mercer SW. Adapting clinical guidelines to take account of multimorbidity. BMJ 2012;345:e6341.
18. Murray M, Berwick DM. Advanced access: reducing waiting and delays in primary care. JAMA 2003;289:1035-40.
19. Pope C, Banks J, Salisbury C, Lattimer V. Improving access to primary care: eight case studies of introducing Advanced Access in England. J Health Serv Res Policy 2008;13:33-9.
20. Salisbury C, Goodall S, Montgomery AA, et al. Does Advanced Access improve access to primary health care? Questionnaire survey of patients. Br J Gen Pract 2007;57:615-21.
21. Campbell J, Smith P, Nissen S, Bower P, Elliott M, Roland M. The GP Patient Survey for use in primary care in the National Health Service in the UK — development and psychometric characteristics. BMC Fam Pract 2009;10:57.
22. Asprey A, Campbell JL, Newbould J, et al. Challenges to the credibility of patient feedback in primary healthcare settings: a qualitative study. Br J Gen Pract 2013;63(608):e200-e208.
23. Roland M, Elliott M, Lyratzopoulos G, et al. Reliability of patient responses in pay for performance schemes: analysis of national General Practitioner Patient Survey data in England. BMJ 2009;339:b3851.
24. Salisbury C. Using patient experience within pay for performance programmes. BMJ 2009;339:b4224.
25. Box G. Patient participation groups: the national picture. Qual Prim Care 2009;17:291-7.
26. Lester HE, Hannon KL, Campbell SM. Identifying unintended consequences of quality indicators: a qualitative study. BMJ Qual Saf 2011;20:1057-61.
27. Campbell SM, Kontopantelis E, Hannon KL, Burke M, Barber A, Lester HE. Framework and indicator testing protocol for developing and piloting quality indicators for the UK quality and outcomes framework. BMC Fam Pract 2011;12:85.
28. Smith P. On the unintended consequences of publishing performance data in the public sector. Int J Public Adm 1995; 18:277-310.
29. Lester HE, Hannon KL, Campbell SM. Identifying unintended consequences of quality indicators: a qualitative study. BMJ Qual Saf 2011;20:1057-61.
30. Howie JG, Heaney DJ, Maxwell M. Care of patients with selected health problems in fundholding practices in Scotland in 1990 and 1992: needs, process and outcome. Br J Gen Pract 1995;45:121-6.
31. Doran T, Kontopantelis E, Valderas J, et al. Effect of financial incentives on incentivized and non-incentivized clinical activities: longitudinal analysis of data from the UK Quality and Outcomes Framework. BMJ 2011;342:d3590.
32. Gravelle H, Sutton M, Ma A. Doctor behaviour under a pay for performance contract: treating, cheating and case finding. Econ J 2010;120:F129-F156.
33. Campbell SM, Hannon K, Lester HE. Exception reporting in the Quality and Outcomes Framework: views of practice staff — a qualitative study. Br J Gen Pract 2011;61:183-9.
34. 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.
35. Doran T, Kontopantelis E, Fullwood C, Lester H, Valderas JM, Campbell S. Exempting dissenting patients from pay for performance schemes: retrospective analysis of exception reporting in the UK Quality and Outcomes Framework. BMJ 2012;344:e2405.
36. Campbell SM, McDonald R, Lester H. The experience of pay for performance in English family practice: a qualitative study. Ann Fam Med 2008;6:228-34.
37. Lester H, Matharu T, Mohammed MA, Lester D, Foskett-Tharby R. Implementation of pay for performance in primary care: a qualitative study 8 years after introduction. Br J Gen Pract 2013;63(611):e408-e415.
38. McGregor W, Jabareen H, O'Donnell CA, Mercer SW, Watt GC. Impact of the 2004 GMS contract on practice nurses: a qualitative study. Br J Gen Pract 2008;58:711-9.
39. Lester H, Schmittdiel 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: 10.1056/NEJMhpr1316051