View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Open Research Exeter

BMJ

BMJ 2011;342:d3590 doi: 10.1136/bmj.d3590

RESEARCH

Effect of financial incentives on incentivised and non-incentivised clinical activities: longitudinal analysis of data from the UK Quality and Outcomes Framework

Tim Doran *clinical research fellow*[1], Evangelos Kontopantelis *research associate*[1], Jose M Valderas *clinical lecturer*[2], Stephen Campbell *senior research fellow*[1], Martin Roland *professor of health services research*[3], Chris Salisbury *professor of primary healthcare*[4], David Reeves *senior research fellow*[1]

[1]National Primary Care Research and Development Centre, University of Manchester, Manchester M13 9PL, UK; [2]NIHR School for Primary Care Research, Department of Primary Health Care, University of Oxford, Oxford OX3 7LF; [3]General Practice and Primary Care Research Unit, University of Cambridge, Cambridge CB2 0SR; [4]Academic Unit of Primary Health Care, University of Bristol, Bristol BS8 2AA

Abstract

Objective To investigate whether the incentive scheme for UK general practitioners led them to neglect activities not included in the scheme.

Design Longitudinal analysis of achievement rates for 42 activities (23 included in incentive scheme, 19 not included) selected from 428 identified indicators of quality of care.

Setting 148 general practices in England (653 500 patients).

Main outcome measures Achievement rates projected from trends in the pre-incentive period (2000-1 to 2002-3) and actual rates in the first three years of the scheme (2004-5 to 2006-7).

Results Achievement rates improved for most indicators in the pre-incentive period. There were significant increases in the rate of improvement in the first year of the incentive scheme (2004-5) for 22 of the 23 incentivised indicators. Achievement for these indicators reached a plateau after 2004-5, but quality of care in 2006-7 remained higher than that predicted by pre-incentive trends for 14 incentivised indicators. There was no overall effect on the rate of improvement for non-incentivised indicators in the first year of the scheme, but by 2006-7 achievement rates were significantly below those predicted by pre-incentive trends.

Conclusions There were substantial improvements in quality for all indicators between 2001 and 2007. Improvements associated with financial incentives seem to have been achieved at the expense of small detrimental effects on aspects of care that were not incentivised.

Introduction

Over the past two decades funders and policy makers worldwide have experimented with initiatives to change physicians' behaviour and improve the quality and efficiency of medical care.[1] Success has been mixed, and attention has recently turned to payment mechanism reform, in particular offering direct financial incentives to providers for delivering high quality care.[2] In 2004 in the UK the Quality and Outcomes Framework (QOF) was introduced—a mechanism intended to improve quality by linking up to 25% of general practitioners' income to achievement of publicly reported quality targets for several chronic conditions.[3]

Should these incentives succeed, the potential benefits for patients with the relevant conditions are considerable.[4] Incentives might also improve general organisation of care, benefiting processes and conditions beyond those covered by the incentives.[5] Financial incentives have several potential unintended consequences, however. For example, they might result in diminished provider professionalism, neglect of patients for whom quality targets are perceived to be more difficult to achieve, and widening of health inequalities.[6 7] Doctors might also focus on the conditions linked to incentives and neglect other conditions[8] or, where certain activities are incentivised within the management of a particular condition, might neglect other activities for patients with that condition.

Practices in England generally performed well on incentivised activities in the first year of the UK incentive scheme, and overall performance improved over the next two years.[9-11] It is

Corresponding to: Tim Doran tim.doran@manchester.ac.uk

Extra material supplied by the author: Methodological appendix (see http://www.bmj.com/content/342/bmj.d3590/suppl/DC1)

not known, however, how much of this improvement is attributable to the incentive scheme and how much to underlying trends in quality improvement. There is also little evidence on the impact of the incentives on activities lying outside the incentive scheme.

Investigating these issues is problematic because performance data were not routinely collected before the scheme's implementation, and afterwards data were collected only at the practice level for activities included in the framework. Evidence from small patient groups suggests that achievement of incentivised activities did accelerate on the introduction of the scheme, with some positive spillover to non-incentivised activities for incentivised conditions in the first year[12] but not for non-incentivised conditions.[13]

The aim of our study is to use a longitudinal dataset at the patient level to examine changes in performance after the introduction of the incentive scheme for processes that became part of the incentive scheme and for processes that did not, and to compare the two groups.

Methods

The incentive scheme

The Quality and Outcomes Framework, introduced in 2004, links up to 25% of UK family practitioner income to performance on 76 clinical quality indicators and 70 indicators relating to organisation of care and patient experience.[3] Of the clinical indicators, 10 relate to maintaining disease registers, 56 to processes of care (such as measuring disease parameters and giving treatments), and 10 to intermediate outcomes (such as controlling blood pressure). Indicators are periodically reviewed, and can be adjusted or dropped from the scheme altogether, with new indicators being introduced. Physicians are permitted to use their clinical judgment to exclude inappropriate patients from achievement calculations ("exception report"). Practices are awarded points based on the proportion of patients for whom targets are achieved, between a lower achievement threshold of 40% for most indicators (that is, practices must achieve the targets for over 40% of patients to receive any points) and an upper threshold that varies according to the indicator. In 2007 each point earned the practice £125 (€141; $202), adjusted for patient population size and disease prevalence. A maximum of 1000 points was available, equating to £31 000 per physician.

Data

Patient level data were extracted from the General Practice Research Database (GPRD), which contains anonymised, patient based data on morbidity, prescribing, treatment, and referral collected from over 500 general practices, covering about 7% of the UK population (4.4 million patients).[14] Data are in Read code format—a hierarchical system used to code clinical data. Additional data on prescriptions and test results are available as free-text entries. We selected a sample of 148 practices that provided data to the GPRD continuously between January 2000 and December 2007, structured to include a range of list (patient panel) sizes. Selected practices were nationally representative in terms of patient sex and age distribution and area socioeconomic deprivation but had a relatively large average list size, reflecting a bias towards larger practices in the GPRD (table 1). Overall, the selected practices performed marginally better than national practices on the clinical indicators in the Quality and Outcomes Framework (table 1). A random selection of 4500 patients registered for at least one day between 1 January 2000 and 31 December 2007 was drawn from each practice. For practices with fewer than 4501 patients, all patients were selected. The final sample consisted of 653 500 patients. Patients with relevant conditions were identified from their diagnostic Read codes (see appendix on bmj.com). Patients for whom targets were met were identified from the relevant Read codes and free-text terms.

Construction of indicators

TD, SC, JMV, CS, and MR initially identified 428 quality of care indicators—combinations of processes of care and patient groups—from published indicators of quality of care that had been developed by a recognised method or which had broad professional consensus (for example, derived from the *British National Formulary*, published national guidelines, or authoritative statements from specialist societies).[15-26]

Indicators were rejected if the activity did not fall under the remit of all general practices or if data to assess performance were not available in practices' electronic patient records. In order to exclude, as far as possible, the effects on indicator achievement of factors other than inclusion in the incentive scheme, we also rejected indicators for which there was, during the period of observation, a substantial change in the evidence base, in coding procedures, or in availability of data from electronic laboratory reports. Our analysis measured the effects of incentives over both the short term (one year) and longer term (three years), and we therefore also rejected indicators that were incorporated into, or dropped from, the incentive scheme after the first year. Examples of rejected indicators are given in the appendix.

After this process, 42 indicators were selected for the study. The selected indicators were divided into two categories, incentivised indicators (that is, those included in the Quality and Outcomes Framework, where the particular process was incentivised for the particular patient group) or non-incentivised indicators (that is, those not included in the Quality and Outcomes Framework). Using an a priori schema, we further classified the indicators into two types of clinical activity—measurement (such as monitoring blood pressure) or prescribing (such prescribing β blockers). Research has shown that practices approach these types of activities differently under the incentive scheme: for example, rates of exception reporting (identifying patients as unsuitable for a clinical indicator) averaged 3% for measurement activities and 13% for treatment or prescribing activities in 2005-6.[27] The levels of achievement that are practically attainable also differ between the activity types, with achievement rates generally lower for prescribing compared with measurement activities. We therefore analysed indicators of the two activity types separately. Indicators of a third activity type—intermediate outcomes—were excluded as no examples of non-incentivised outcome indicators could be identified.

The four combinations of indicator category and activity type (that is, incentivised measurement, incentivised prescribing, non-incentivised measurement, and non-incentivised prescribing) were designated "indicator groups" for our analysis. Table 2 describes the individual quality indicators.

Analysis

We divided the study period into financial years (1 April to 31 March) to correspond to the assessment periods for the financial incentives. We designated 2000-1, 2001-2, and 2002-3 as pre-intervention time points; 2003-4 as a preparatory phase (when details of the forthcoming quality targets were in the public domain but incentives were not yet available); and

2004-5, 2005-6, and 2006-7 as post-intervention time points. Patients were included in the sample for a year provided they were registered with their practice for the entire year. All available events were then examined, and patients with relevant diagnostic or activity codes were included in the denominator (D_i) for a given quality indicator. From these patients, those for whom the indicator was met within the required time frame were included in the numerator (N_i). We excluded patients from non-incentivised indicators where the activity was incentivised for a co-existing condition—for example, patients with coronary heart disease were excluded from indicator C8 (cholesterol measurement for patients with hypertension). For each practice, annual achievement rates on each indicator were calculated as N_i/D_i. We applied logit transformations to achievement rates to reduce floor and ceiling effects.[28]

We analysed and compared indicator groups with respect to the difference between achievement predicted from the trend in the pre-incentive period (2000-1 to 2002-3) and actual achievement in both 2004-5 (the first year of the incentive scheme) and 2006-7 (the third year). Data were analysed as a two level, mixed, multivariate regression with indicator crossed with practice. Covariates included predicted achievement rates and control variables for differences in disease registers (indicator denominators, patient age, and patient sex). We derived estimates of means and standard errors for each indicator group, first controlling for differences in covariates across time points and practices within indicators (model 1), then adding control for differences between indicators (model 2). Full details of each model appear in the appendix on bmj.com. We first examined the impact of the incentive scheme on each of our four indicator groups separately, and then used post-estimation tests to compare between incentivised and non-incentivised indicators for each of the two activity types. We also analysed individual indicators. All statistical comparisons were made at an α level of 5%. Analyses were performed using Stata (version 11). For presentation purposes, we report means back transformed from the logit scale to percentages.

Results

Baseline performance before the introduction of the financial incentives

Over the pre-intervention period (2000-1 to 2002-3) achievement increased significantly for 32 of the 42 indicators, decreased significantly for two, and did not change for eight. In 2002-3 achievement varied from 1.4% for indicator C3 (women aged >50 with depression who have had thyroid function tests) to 98.7% for indicator D7 (patients treated with sumatriptan who do not have angina) (see table 3 for full list). There were systematic differences between indicator groups for the rate of improvement (P<0.001), with incentivised measurement indicators (that is, measurement indicators that became incentivised in 2004-5) having the fastest overall rate of improvement, and non-incentivised prescribing indicators the slowest overall rate (figure). At the end of the pre-incentive period (2002-3) indicator groups also differed in mean performance scores (p<0.001), with the highest overall score for non-incentivised prescribing indicators and lowest for non-incentivised measurement indicators.

Achievement rates in the first year of the financial incentives

In 2004-5, achievement rates for incentivised indicators were significantly higher than rates projected from pre-incentive trends for all 17 measurement indicators, and five of the six prescribing indicators (table 3). The increase in achievement above that predicted from pre-incentive trends varied from 1.2% to 37.7%, with four indicators (all relating to measuring smoking status) having increases of over 30%. Both groups of incentivised indicators showed overall increases in achievement significantly above the predicted rates (table 4).

For the non-incentivised indicators, there were significant increases in achievement above the predicted rates for two of the nine measurement indicators, and a significant decrease for one (that is, achievement was below that predicted from pre-incentive trends). There was no significant overall effect for this group of indicators. Two of the 10 non-incentivised prescribing indicators had a significant increase above predicted achievement rates, three had significant decreases, and there was no effect for the remaining five. Overall, there was no significant effect for this group of indicators (table 4). There was, however, substantial heterogeneity of effect within this group of indicators.

For both types of activity—measurement and prescribing—increases in achievement above predicted rates were significantly larger for incentivised indicators than non-incentivised indicators under both model 1 and model 2 (table 4).

Achievement rates in the second and third years of the financial incentives

There were significant increases in achievement rates for all four indicator groups between 2004-5 (the first intervention year) and 2006-7 (the third intervention year). With the exception of non-incentivised measurement indicators, these increases, although significant, were small (<3%, see table 4).

For the incentivised indicators, achievement remained significantly above projected rates in 2006-7 for 10 of the 17 measurement indicators, and four of the six prescribing indicators (table 3). However, rates were significantly below projections for five measurement indicators and one prescribing indicator. Even so, both incentivised groups continued to have overall achievement rates significantly higher than predicted.

For the non-incentivised indicators, achievement was significantly above projected rates in 2006-7 for one of the nine measurement indicators, but significantly below for seven. The overall achievement rate for this group was significantly lower than predicted (by 5.6%). Four of the 10 prescribing indicators had achievement rates significantly lower than predicted, and one had a significantly higher rate. The overall achievement rate for this group of indicators was also significantly lower than predicted (table 4), although heterogeneity of effect was high.

Relative to predicted values, overall achievement rates in 2006-7 for non-incentivised indicators were significantly below those for incentivised indicators, for both measurement and prescribing activities (table 4).

Discussion

The success of quality improvement initiatives depends not only on whether they achieve their intended outcomes, but on their unintended consequences. With “pay for performance” schemes there is a risk that rewarding performance of certain clinical activities will divert attention from other, unrewarded activities. The original UK Quality and Outcomes Framework was developed with a wide range of indicators, 76 clinical and 70 non-clinical, in part to prevent doctors focusing on too narrow a range of activities and to encourage a broader systematic

approach to improving quality of care, although quality improvement initiatives implemented in response to the incentives tended to focus specifically on the incentivised conditions and activities.[29] We used patient-level data from general practices using electronic patient records from 1999 onwards to examine trends in achievement for incentivised and non-incentivised activities.

Strengths and limitations of the study

Our study is subject to limitations. First, any changes in the consistency and accuracy of data recording over time would affect our findings, particularly for incentivised indicators. Such changes are more likely to affect indicators assessed by practices (such as body mass index) than indicators assessed by third parties such as laboratories (for example, serum creatinine concentration). However, we found no consistent relation between changes in achievement rates over time and the agent responsible for measuring indicator parameters.

Second, changes in case mix over time might have affected achievement rates, particularly if there were changes in case finding activity under the incentive scheme. To account for this, we controlled for changes over time in indicator denominators and age and sex profiles. We also found no substantial differences in patterns of achievement for different cohorts of patients (such as diabetic patients diagnosed in 2000-1 compared with patients diagnosed in later years).

Third, although the sampled practices were nationally representative in terms of patient demographics, they might have been atypical in terms of organisation and quality of care. Collectively, they had marginally higher achievement scores on incentivised indicators than the national average in each intervention year,[30] which might be attributable to higher baseline performance in the pre-incentive period or a greater response to the incentives. However, with respect to their general pattern of achievement for the incentivised indicators—high in the first year followed by diminishing improvements in the second and third years—the sampled practices followed the national trend.[8]

Fourth, we were highly conservative in selecting indicators, focusing on processes not subject to other forms of incentive and for which guidelines remained unchanged throughout the study period. This limits the generalisability of our findings, although comparison with trends in achievement rates for non-selected incentivised indicators suggests that the selected indicators were not atypical.

In addition to examining trends in achievement for incentivised and non-incentivised aspects of care, we made direct comparisons between these. To strengthen this comparison we grouped indicators, a priori, by activity type. We also controlled for different baseline achievement rates and trends, and for changes in disease registers over time. Nevertheless, the subgroups remained non-equivalent in important respects, both in terms of patient groups and care processes. For example, in contrast to other subgroups, most non-incentivised prescribing indicators related to "negative" (patient safety) activities, and the general focus of the incentives on promoting recommended processes of care rather than on avoiding proscribed activities may have diverted attention from patient safety issues. Activities that were incorporated into the incentive scheme may also have been accorded greater importance by clinicians even before the scheme was implemented.

Non-equivalence does not invalidate our findings on the effect of the incentives on incentivised and non-incentivised indicators, but it raises the question of whether such non-equivalence can account for the observed post-intervention differences between the groups. To investigate further, we used data from the first two pre-intervention years to predict rates of achievement in the last pre-intervention year (2002-3) and found only small deviations from expectation for all four indicator groups. This implies that uncontrolled sources of non-equivalence did not produce group differences in the immediate pre-intervention period, and that the post-intervention differences we observed developed after that point (see appendix on bmj.com). However, as the financial incentives were introduced simultaneously nationwide, with no control practices, we could not test our assumption that in the absence of the incentives the pre-intervention achievement trends would have continued.

The effect of the incentives

Of the 23 incentivised indicators we analysed, measurement indicators improved at the fastest rate in the pre-intervention period, from a relatively low baseline in 2000-1. After the introduction of incentives, achievement rates increased substantially for all measurement indicators, with increases above projected rates in 2004-5 of up to 38%. Of the six indicators with the greatest increases, five entailed recording smoking status—a technically straightforward activity that had a low baseline (either because practices were not asking most of their patients before the incentives or were not recording that they were doing so), required minimal patient compliance, and was monitored exclusively by the practices. Collectively, prescribing activities had higher baseline achievement rates than measurement activities in 2000-1, slower increases in achievement rates in the pre-intervention period, and smaller increases above projected rates in the first intervention year (2004-5) of between 1.2% and 8.3%. Achievement rates for all incentivised indicators reached a plateau in the second and third years of the scheme, and as a result only 14 of the 23 incentivised indicators had achievement rates significantly higher than rates projected from pre-intervention trends after three years.

For non-incentivised indicators the introduction of financial incentives seemed to have little overall impact in the first year, with quality continuing to improve at around the pre-intervention rate. In the second and third years, although quality continued to improve, the rate of improvement slowed. By 2006-7 quality was significantly worse than projected from pre-incentive trends, most notably for measurement activities, and was also significantly below—relative to projection—the quality for incentivised activities.

The general improvement in quality before the introduction of the financial incentives suggests that the quality initiatives implemented in the UK at that time—including clinical audits, development of information technology, and creation of quality oriented statutory bodies—were having an effect in improving quality of care. The Quality and Outcomes Framework built on this infrastructure and introduced a national set of quality targets (clinical guidelines) supported by computer prompts and feedback and backed by financial and reputational incentives. The scheme was associated with additional increases in performance in its first year across a broad range of incentivised activities, but the rate of improvement was not sustained. There are several possible explanations for this. Practices might have improved their recording procedures in the first year of the scheme in response to the incentives, effectively "correcting" achievement rates, particularly for some measurement indicators. After this, the rate of quality improvement seemed to plateau. Alternatively, practices might have reached their achievement limit by 2006-7, with the incentive scheme hastening their arrival

at this ceiling. A third explanation is that practices relaxed their efforts after the first year of the scheme. Most practices attained achievement rates above the maximum achievement thresholds (the level of achievement required to secure maximum remuneration under the scheme) for most incentivised indicators in the first year, and thereafter had no financial incentive for further improvement.

Conclusions and policy implications

Quality of primary care was generally improving in England in the early 2000s. The introduction of an incentive scheme seemed to accelerate this trend for incentivised activities, but quality quickly reached a plateau. Incentives had little apparent impact on non-incentivised activities in the short term, but seem to have had some detrimental effects in the longer term, possibly because of practices focusing on patients for whom rewards applied. Some aspects of the UK Quality and Outcomes Framework and its setting may have limited its impact on non-incentivised activities (such as the lack of financial penalties, the wide range of indicators, and the existence of other quality initiatives) and other aspects may have exacerbated its impact (such as the large size of the incentives and the appearance of incentive related prompts on clinical computing systems during consultations). Findings may be different under schemes with different incentive structures operating in different settings. Nevertheless, these findings show some important limitations of financial incentive schemes in health care, and the importance of monitoring, as far as possible, activities that are not incentivised in addition to those that are when determining the effects of such schemes.

We are grateful for comments provided by Arnold Epstein and Meredith Rosenthal of the Harvard School of Public Health during the drafting of the paper.

Contributors: TD participated in the planning of the study, analysis and interpretation of data, and drafting and editing the manuscript, and is guarantor for the integrity of the data and accuracy of the data analysis. EK participated in the planning of the study, analysis and interpretation of data, and drafting and editing the manuscript. JMV, SC, CR, and DR participated in the planning of the study, analysis and interpretation of data, and editing the final manuscript. All authors saw and approved the final version of the manuscript and had full access to all of the data in the study

Funding: There was no direct funding for this study, but the National Primary Care Research and Development Centre receives core funding from the UK Department of Health. The views expressed are those of the authors and not necessarily those of the Department of Health. At the time of writing, TD was supported by a grant from the Commonwealth Fund as a Harkness Fellow in Health Care Policy and Practice.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; MR served as an academic advisor to the government and British Medical Association negotiating teams during the development of the UK pay for performance scheme during 2001 and 2002.

Ethical approval: Not required: study based on publicly available data.

Data sharing: Technical appendix and statistical code available from the corresponding author (tim.doran@manchester.ac.uk). The dataset was derived from the General Practice Research Database and is not available from the authors, but it can be derived on application to GPRD.

1 Tsai A, Morton S, Mangione C, Keeler E. A meta-analysis of interventions to improve care for chronic illnesses. *Am J Manag Care* 2005;11:478-88.
2 Petersen L, Woodard L, Urech T, Daw C, Sookanan S. Does pay-for-performance improve the quality of health care? *Ann Intern Med* 2006;145:265-72.
3 Roland M. Linking physician pay to quality of care: a major experiment in the UK. *N Engl J Med* 2004;351:1448-54.
4 McElduff P, Lyratzopoulos G, Edwards R, Heller R, Shekelle P, Roland M. Will changes in primary care improve health outcomes? Modelling the impact of financial incentives introduced to improve quality of care in the UK. *Qual Saf Health Care* 2004;13(3):191-7.
5 Lester H, Sharp D, Hobbs F, Lakhani M. The Quality and Outcomes Framework of the GMS contract: a quiet evolution for 2006. *Br J Gen Pract* 2006;56:244-6.
6 McDonald R, Harrison S, Checkland K, Campbell S, Roland M. Impact of financial incentives on clinical autonomy and internal motivation in primary care: ethnographic study. *BMJ* 2007;334:1357-62.
7 Casalino L. Will pay-for-performance and quality reporting affect health care disparities? *Health Affairs* 2007;26:w405.
8 Heath I. The cawing of the crow … Cassandra-like, prognosticating woe. *Br J Gen Pract* 2004;54:320-1.
9 Millett C, Gray J, Saxena S, Netuveli N, Khunti K, Majeed A. Ethnic disparities in diabetes management and pay-for-performance in the UK: the Wandsworth Prospective Diabetes Study. *PLoS Med* 2007;4:e191.
10 Doran T, Fullwood K, Kontopantelis E, Reeves D. The effect of financial incentives on inequalities in the delivery of primary care in England. *Lancet* 2008;372:728-36.
11 Vaghela P, Ashworth M, Schofield P, Gulliford M. Population intermediate outcomes of diabetes under pay-for-performance incentives in England from 2004 to 2008. *Diabetes Care* 2009;32:427-9.
12 Campbell S, Reeves D, Kontopantelis E, Sibbald B, Roland M. Quality of primary care in England with the introduction of pay for performance. *N Engl J Med* 2007;351:181-90.
13 Steel N, Maisey S, Clark A, Fleetcroft R, Howe A. Quality of clinical primary care and targeted incentive payments: an observational study. *Br J Gen Pract* 2007;57:449-54.
14 GPRD. The GPRD practice and patient populations. 2011. www.gprd.com.
15 Agency for Healthcare Research and Quality. AHRQ quality indicators—guide to prevention quality indicators: hospital admission for ambulatory care: sensitive conditions. http://www.qualityindicators.ahrq.gov/.
16 Campbell S, Roland M, Shekelle P, Cantrill J, Buetow S, Cragg D. The development of review criteria for assessing the quality of management of stable angina, adult asthma and non insulin dependent diabetes mellitus in general practice. *Qual Health Care* 1999;8:6-15.
17 Department of Health. *National service framework for coronary heart disease.* DH, 2000.
18 Department of Health. *National service framework for diabetes.* DH, 2001.
19 Department of Health. *National service framework for mental health.* DH, 1999.
20 Department of Health. *National service framework for renal services.* DH, 2004.
21 Institute for Clinical Systems Improvement. *Health care guidelines.* 2007 www.icsi.org/guidelines_and_more/.
22 Joint Formulary Committee. *British national formulary.* 54th ed. British Medical Association and Royal Pharmaceutical Society of Great Britain, 2007.
23 Marshall M, Campbell S, Hacker J, Roland M, eds. *Quality indicators for general practice: a practical guide to clinical quality indicators for primary care professionals and managers* . Royal Society of Medicine Press, 2002.
24 Marshall M, Leatherman S, Mattke S: OECD Health Promotion, Prevention and Primary Care Panel. *Health technical paper 16: selecting indicators for the quality of health promotion, prevention and primary care at the health system level in OECD countries.* OECD, 2004.
25 RAND. *Quality of care assessment tools.* 2000. www.rand.org/health/surveys_tools/qatools.html.
26 Scottish Intercollegiate Guidelines Network. *Published guidelines.* 2007. www.sign.ac.uk/guidelines/index.html.
27 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.
28 Winer B. *Statistical principles in experimental design.* McGraw-Hill, 1962.
29 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.
30 The Information Centre. *Online GP practice results database* . 2011. www.qof.ic.nhs.uk/ .

Accepted: 18 April 2011

Cite this as: *BMJ* 2011;342:d3590

What is already known on this topic

Carefully constructed financial incentive schemes in healthcare can lead to improvements in quality of care for incentivised activities

Concerns remain about the potential unintended consequences of financial incentives, in particular that they could lead to neglect of non-incentivised aspects of patient care

What this study adds

We examined trends in quality of care for 42 activities (23 incentivised under the UK Quality and Outcomes Framework incentive scheme, and 19 not incentivised) selected from 428 identified quality of care indicators

Quality of care initially improved for incentivised activities but quickly reached a plateau. Incentives had little impact on non-incentivised activities in the short term, but after three years quality of care for some fell significantly below levels predicted from pre-incentive trends

Improvements associated with financial incentives seem to have been achieved at the expense of small detrimental effects on aspects of care that were not incentivised

Tables

Table 1| **Comparison of characteristics of general practices included in study and national practices (2006-7). Values are means (standard deviations)**

Characteristic	Study practices (n=148)	English practices (n=8226)
Patient characteristics		
Total number of patients	7117 (3441)	6478 (3967)
Percentage female	50.1 (1.7)	49.6 (2.8)
Percentage by age (years):		
15–44	41.3 (6.2)	43.2 (8.3)
45–64	25.6 (4.2)	24.3 (4.7)
≥65	15.8 (4.8)	15.0 (5.3)
Area characteristics		
Area deprivation*	26.7 (18.0)	26.4 (17.5)
Performance on Quality and Outcomes Framework indicators		
Overall reported achievement†	90.8 (4.6)	89.7 (5.0)
Overall exception reporting rate‡	8.4 (2.9)	8.3 (2.8)

*Measured by the Index of Deprivation 2007.

†Mean of reported achievement rates for all clinical indicators. Reported achievement is the proportion of eligible patients for whom the indicators have been met, excluding patients exception reported by the practice.

‡Mean of exception reporting rates for all clinical indicators. Practices are permitted to exclude (exception report) patients for whom the indicators are deemed inappropriate (for example, because of terminal illness or a contraindication to an incentivised treatment).

Table 2| **Quality of care indicators analysed in study of 148 general practices to compare performance of activities incentivised in the QOF scheme with activities not covered by the scheme**

Study reference (QOF reference*)	Indicator	Period†
A) Incentivised measurement indicators		
A1 (AST4)	Patients aged ≥20 with asthma whose notes record smoking status (except those who have never smoked, when smoking status should be recorded at least once)	15 months
A2 (BP2)	Patients with hypertension whose notes record smoking status (except those who have never smoked, when smoking status need be recorded only once)	15 months
A3 (BP4)	Patients with hypertension in whom there is a record of blood pressure	9 months
A4 (CHD3)	Patients with coronary heart disease whose notes record smoking status (except those who have never smoked, when smoking status need be recorded only once)	15 months
A5 (CHD5)	Patients with coronary heart disease whose notes have a record of blood pressure	15 months
A6 (CHD7)	Patients with coronary heart disease whose notes have a record of total cholesterol	15 months
A7 (COPD4)	Patients with chronic obstructive pulmonary disease in whom there is a record of smoking status	15 months
A8 (DM5)	Patients with diabetes who have a record of HbA_{1c} or equivalent	15 months
A9 (DM11)	Patients with diabetes who have a record of blood pressure	15 months
A10 (DM14)	Patients with diabetes who have a record of serum creatinine testing	15 months
A11 (DM16)	Patients with diabetes who have a record of total cholesterol	15 months
A12 (MH3)	Patients receiving lithium therapy with a record of lithium levels checked	6 months
A13 (MH4)	Patients receiving lithium therapy with a record of serum creatinine and TSH	15 months
A14 (STR3)	Patients with transient ischaemic attack or stroke whose notes record smoking status (except those who have never smoked, when smoking status need be recorded only once)	15 months
A15 (STR5)	Patients with transient ischaemic attack or stroke who have a record of blood pressure	15 months
A16 (STR7)	Patients with transient ischaemic attack or stroke who have a record of total cholesterol	15 months
A17 (THY2)	Patients with hypothyroidism with thyroid function tests recorded	15 months
B) Incentivised prescribing indicators		
B1 (CHD10)	Patients with coronary heart disease who are treated with a β blocker (unless contraindication or side effects recorded)	Current
B2 (CHD12)	Patients with coronary heart disease who have a record of influenza immunisation	7 months‡
B3 (CHD/LVD3)	Patients with coronary heart disease and left ventricular dysfunction who are treated with ACE inhibitors (or A2 antagonists)	Current
B4 (COPD8)	Patients with chronic obstructive pulmonary disease who have a record of influenza immunisation	7 months‡
B5 (DM18)	Patients with diabetes who have a record of influenza immunisation	7 months‡
B6 (STR10)	Patients with transient ischaemic attack or stroke who have a record of influenza immunisation	7 months‡
C) Non-incentivised measurement indicators		
C1	General population age 25–44 who have a record of blood pressure in their medical records	5 years
C2	Patients with peripheral arterial disease who have a record of total cholesterol	15 months
C3	Women aged >50 with depression who also have a record of thyroid function tests	6 weeks
C4	Patients with dementia who also have a record of thyroid function tests	6 weeks
C5	Patients with osteoarthritis who have a record of weight or body mass index	5 years
C6	Patients with coronary heart disease who have a record of blood sugar	Since diagnosis
C7	Patients with hypertension who have a record of blood glucose	Since diagnosis
C8	Patients with hypertension who have a record of total cholesterol	Since diagnosis
C9	Patients with hypertension who have a record of creatinine	Since diagnosis
D) Non-incentivised prescribing indicators		
D1	Patients with congestive heart disease who have a record of influenza immunisation	7 months‡
D2	Patients with chronic renal disease who have a record of influenza immunisation	7 months‡
D3	Patients with asthma who have been issued with at least one prescription for a β blocker§	15 months
D4	Patients with diabetes aged ≥65 treated with chlorpropamide or glibenclamide§	12 months

Table 2 (continued)

Study reference (QOF reference*)	Indicator	Period†
D5	Patients with osteoporosis treated with biphosphonates, selective oestrogen receptor modulator, HRT, parathyroid hormone, or calcitonin	15 months
D6	Patients with otitis media treated with oral decongestant§	15 months
D7	Patients treated with sumatriptan who also have angina§	15 months
D8	Patients with back pain treated with strong analgesics (co-dydramol upwards)§	15 months
D9	Patients with upper respiratory tract infection, otitis media, cough, or bronchitis treated with an antibiotic (excluding patients with diagnoses of chronic obstructive pulmonary disease or cystic fibrosis)§	Same consultation
D10	Patients who have had a splenectomy who have received pneumococcal vaccine	5 years

QOF=Quality and Outcomes Framework, TSH=thyroid stimulating hormone, ACE= angiotensin converting enzyme, A2=angiotensin II, HRT=hormone replacement therapy.

*Reference code used to identify indicator in the QOF.

†Period before financial year end (31 March) in which activity must have been performed.

‡For influenza immunisation indicators the activity must have been performed within the influenza season (1 September to 31 March).

§Reverse indicators—a lower proportion represents higher quality care.

Table 3| **Achievement rates for quality of care indicators among 148 general practices to compare performance of activities incentivised in the QOF scheme with activities not covered by the scheme**

Indicator (No of eligible patients)*	Actual achievement rates†							Projected achievement rates‡		Difference of actual from projected (95% CI)§	
	Pre-intervention period			Preparatory period	Post-intervention period						
	2000-1	2001-2	2002-3	2003-4	2004-5	2005-6	2006-7	2004-5	2006-7	2004-5	2006-7
A) Incentivised measurement indicators											
A1 (23 737)	25.9	30.9	37.4	64.4	84.6	81.6	84.9	51.4	67.4	35.6 (33.9 to 37.1)	19.7 (18.0 to 21.3)
A2 (60 466)	24.0	29.9	36.8	62.9	80.8	74.6	81.4	53.0	70.8	30.1 (28.0 to 32.0)	14.0 (11.9 to 15.9)
A3 (60 466)	79.3	82.9	86.0	89.9	95.1	96.1	96.2	92.2	95.1	3.7 (3.2 to 4.2)	1.7 (1.3 to 2.0)
A4 (18 979)	27.8	40.5	48.0	73.8	87.5	84.4	86.7	74.1	89.6	16.4 (14.9 to 17.7)	0.7 (−0.9 to 2.1)
A5 (18 979)	73.9	80.0	84.7	90.3	95.2	95.6	96.0	93.1	96.5	3.1 (2.6 to 3.6)	0.5 (0.0 to 0.9)
A6 (18 979)	41.0	53.3	62.2	74.4	86.6	88.2	89.1	83.2	93.2	5.8 (4.5 to 6.9)	−2.1 (−3.2 to −1.0)
A7 (6 723)	29.7	37.1	44.3	74.3	93.2	91.8	94.4	62.8	78.9	31.7 (30.9 to 32.4)	16.4 (15.7 to 17.1)
A8 (18 411)	70.5	77.3	82.2	88.6	92.6	93.4	93.5	92.3	96.1	1.6 (1.0 to 2.2)	−1.5 (−2.2 to −0.9)
A9 (18 411)	81.6	85.4	88.4	92.7	95.6	96.0	96.2	94.8	97.1	1.6 (1.2 to 1.9)	−0.1 (−0.5 to 0.3)
A10 (18 411)	51.8	61.9	70.8	85.2	92.0	93.6	93.8	90.6	97	2.7 (1.9 to 3.4)	−2.1 (−2.7 to −1.5)
A11 (18 411)	57.2	66.7	75.5	84.6	91.5	92.7	92.8	90.7	96.3	2.0 (1.3 to 2.6)	−2.5 (−3.2 to −1.9)
A12 (530)	38.1	48.9	51.9	68.1	86.4	86.6	89.3	67.1	78.6	13.9 (11.7 to 16.0)	5.3 (3.5 to 7.0)
A13 (530)	23.4	34.7	41.8	69.9	89.3	94.2	95.7	63.5	79.4	21.1 (18.9 to 23.0)	6.9 (5.4 to 8.3)
A14 (8 569)	20.9	27.4	34.7	64.3	86.0	80.4	82.9	51.4	71	37.7 (36.1 to 39.1)	15.2 (13.3 to 16.9)
A15 (8 569)	68.0	73.0	78.3	85.6	93.7	94.5	94.7	88.1	93	6.8 (6.2 to 7.5)	2.6 (2.1 to 3.1)
A16 (8 569)	25.3	35.2	44.5	62.4	81.2	84.3	86.3	69.1	86.2	14.5 (12.7 to 16.1)	2.0 (0.8 to 3.1)
A17 (14 084)	61.1	68.9	76.9	86.9	93.6	93.7	93.9	91.9	96.8	3.4 (2.6 to 4.0)	−1.3 (−2.0 to −0.7)
B) Incentivised prescribing indicators											
B1 (15 490)	39.2	42.5	45.0	49.1	53.5	54.9	56.0	51.9	58.6	2.0 (0.4 to 3.5)	−2.0 (−3.5 to −0.5)
B2 (18 979)	64.6	67.7	68.8	72.7	77.1	80.1	79.7	74.4	78.3	3.9 (2.7 to 5.1)	2.6 (1.4 to 3.8)
B3 (1 312)	70.3	72.0	75.0	79.4	82.0	83.0	84.1	80.5	84.5	1.2 (−0.9 to 3.1)	−0.4 (−2.4 to 1.3)
B4 (6 723)	71.7	71.4	71.7	75.5	80.3	82.6	82.1	73.7	73.7	8.3 (6.7 to 9.8)	10.4 (8.9 to 11.8)
B5 (18 441)	64.4	66.8	66.9	71.1	75.5	79.0	77.8	70.5	72.8	6.3 (5.0 to 7.7)	6.4 (4.9 to 7.8)
B6 (8 569)	63.3	66.4	66.8	69.5	74.2	76.6	76.8	71.6	74.9	3.8 (2.2 to 5.4)	3.2 (1.6 to 4.7)
C) Non-incentivised measurement indicators											
C1 (120 802)	59.9	61.3	63.9	66.9	69.8	71.4	72.3	68.4	72.2	1.5 (0.3 to 2.8)	0.2 (−1.7 to 2.0)
C2 (1 134)	16.3	22.7	30.7	36.6	48.1	50.5	53.7	46.2	63.8	1.6 (−2.8 to 6.1)	−9.5 (−13.7 to −5.3)
C3 (14 722)	1.4	1.5	1.4	1.6	1.5	1.1	1.1	1.4	1.4	−0.1 (−0.2 to 0.1)	−0.3 (−0.4 to −0.1)
C4 (1 643)	7.9	9.4	13.8	18.3	19.0	15.6	19.4	18.4	23	−0.9 (−3.0 to 1.3)	−6.4 (−8.6 to −3.9)
C5 (31 877)	51.6	53.5	57.2	62.3	67.0	71.1	76.7	63.9	69.5	6.0 (3.6 to 8.4)	11.6 (9.0 to 13.9)
C6 (15 123)	27.5	39.9	52.6	67.2	77.3	82.2	85.9	82.8	95.5	−1.3 (−3.5 to 0.7)	−5.0 (−6.7 to −3.6)

Table 3 (continued)

Indicator (No of eligible patients)*	Actual achievement rates†							Projected achievement rates‡		Difference of actual from projected (95% CI)§	
	Pre-intervention period			Preparatory period	Post-intervention period						
	2000-1	2001-2	2002-3	2003-4	2004-5	2005-6	2006-7	2004-5	2006-7	2004-5	2006-7
C7 (50 085)	25.4	34.8	46.0	58.6	67.7	73.8	78.4	74.2	91.4	−3.3 (−5.6 to −1.1)	−9.9 (−11.7 to −8.1)
C8 (39 290)	36.7	44.8	53.9	63.4	71.4	77.3	81.7	73	86.3	1.4 (−0.3 to 3.0)	−2.1 (−3.6 to −0.8)
C9 (49 999)	33.2	44.4	55.2	71.1	80.9	86.2	89.8	83.5	95.9	0.3 (−1.3 to 1.8)	−4.3 (−5.2 to −3.5)
D) Non-incentivised prescribing indicators											
D1 (1 109)	64.0	67.4	67.0	68.0	66.9	69.7	69.8	72.4	75.3	−3.1 (−6.7 to 0.3)	−5.2 (8.9 to −1.8)
D2 (5 622)	56.4	59.4	60.5	66.7	62.9	69.1	67.0	60.7	63.7	1.7 (−3.5 to 6.6)	−1.9 (−7.1 to 3.1)
D3 (30 809)¶	98.4	98.3	98.2	97.9	97.8	97.5	97.5	98.4	98.1	−0.1 (−0.3 to 0.1)	−0.08 (−0.3 to 0.2)
D4 (9 479)¶	90.9	93.1	94.6	95.8	96.7	97.5	98.1	97.8	98.7	−0.3 (−0.5 to −0.1)	−0.3 (−0.6 to −0.2)
D5 (6 281)	31.5	37.9	44.7	49.2	52.4	56.5	59.4	59.1	72.7	−6.8 (−8.8 to −4.7)	−12.8 (−14.9 to −10.9)
D6 (9 219)¶	96.0	96.0	96.7	97.3	97.6	97.1	97.2	97.8	98	0.2 (−0.1 to 0.5)	−0.1 (−0.5 to 0.2)
D7 (1 513)¶	98.7	98.1	98.7	98.4	98.6	98.4	98.5	93.5	93.6	−0.01 (−0.5 to 0.5)	0.3 (−0.3 to 0.8)
D8 (26 764)¶	58.9	58.9	57.9	58.3	58.2	57.5	56.7	57.2	56	1.3 (0.0 to 2.6)	0.7 (−0.8 to 2.2)
D9 (93 962)¶	49.5	48.2	46.7	46.6	47.1	48.3	47.6	42.8	39.2	4.0 (2.4 to 5.7)	8.2 (6.1 to 10.3)
D10 (509)	26.8	25.4	28.9	26.9	26.4	30.1	26.9	31.6	33.3	−3.8 (−6.9 to −0.5)	−5.8 (−9.9 to −1.3)

QOF=Quality and Outcomes Framework.

*Based on 2006-7 data.

†Proportion of eligible patients across all 150 practices for whom the indicator was achieved. Periods 2000-1 to 2002-3 were before QOF came into force, 2003-4 when details of the forthcoming quality targets were known but incentives not yet available, and 2004-5 to 2006-7 were after QOF came into force.

‡Projected rate based on achievement trend 2000-1 to 2002-3.

§Values are derived from logit transformed data and may not equate to the difference between the observed rate and projected rate.

¶Reverse indicators, for which a lower proportion represents higher quality care. Reported rates given as (100–achievement rate). These reported rates, rather than achievement rates, were used in the models.

Table 4| Increase in achievement rates above projected rates for quality of care indicators among 148 general practices under the QOF incentive scheme in year 1 (2004-5) and year 3 (2006-7) of scheme, and change in mean achievement rate between 2004-5 and 2006-7. Comparison of performance of measurement and prescribing indicators incentivised under QOF against indicators not covered by the scheme.* Values are means† (95% confidence interval) unless stated otherwise

	Measurement indicators			Prescribing indicators		
	Incentivised	Non-incentivised	P value of difference§	Incentivised	Non-incentivised	P value of difference§
Increase in mean achievement above projected trend†:						
Year 1 (2004-5)	14.5 (14.0 to 15.0)	0.3 (−0.8 to 1.3)	<0.001	4.3 (3.3 to 5.3)	−0.9 (−1.9 to 0.2)	<0.001
Year 3 (2006-7)	3.9 (3.2 to 4.5)	−5.6 (−6.6 to −4.6)	<0.001	2.9 (2.0 to 3.7)	−1.7 (−2.7 to −0.0)	<0.001
Change in mean achievement, 2004-5 to 2006-7	1.9 (1.4 to 2.5)	6.3 (5.7 to 6.8)	<0.001	2.6 (1.8 to 3.3)	1.2 (0.6 to 2.0)	0.002¶

*Results are for model 1 (which controlled for differences in covariates across time points and practices within indicators). Results that differed for model 2 (which controlled for differences in covariates between indicators) are indicated by ¶.

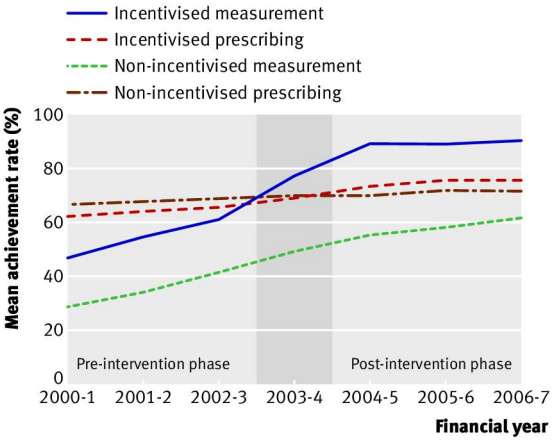
†Group means based on logit transformed data, back transformed to percentage scores.

‡Difference between observed achievement rate and rate projected from pre-incentive trend.

§Difference in increase/change between incentivised and non-incentivised indicators.

¶Not significant under model 2.

Figure



Mean achievement rate of 148 general practices for quality of care indicators from 2000-1 to 2006-7. Performance indicators grouped by activity and whether they were incentivised under the QOF scheme, which came into force from 2004-5. (The mean rate is the mean of the adjusted means for the individual indicators within each group)