

Quality of clinical primary care and targeted incentive payments: an observational study

Nicholas Steel, Susan Maisey, Allan Clark, Robert Fleetcroft and Amanda Howe

ABSTRACT

Background

Payments for recorded evidence of quality of clinical care in UK general practices were introduced in 2004.

Aim

To examine the relationship between changes in recorded quality of care for four common chronic conditions from 2003 to 2005, and the payment of incentives.

Design of study

Retrospective observational study comparing incentivised and non-incentivised indicators of quality of care.

Setting

Eighteen general practices in England.

Method

Medical records were examined for 1156 patients. The percentage of eligible quality indicators achieved for each patient was assessed in 2003 and 2005. Twenty-one quality indicators referred to asthma and hypertension: six subject to and 15 not subject to incentive payments. Another 15 indicators referred to depression and osteoarthritis which were not subject to incentive payments.

Results

A significant increase occurred for the six indicators linked to incentive payments: from 75% achieved in 2003 to 91% in 2005 (change = 16%, 95% confidence interval [CI] = 10 to 22%, $P < 0.01$). A significant increase also occurred for 15 other indicators linked to 'incentivised conditions': 53 to 64% (change = 11%, 95% CI = 6 to 15%, $P < 0.01$). The 'non-incentivised conditions' started at a lower achievement level, and did not increase significantly: 35 to 36% (change = 2%, 95% CI = -1 to 4%, $P = 0.19$).

Conclusion

The introduction of financial incentives was associated with substantial apparent quality improvement for incentivised conditions. For non-incentivised conditions, quality did not appear to improve. Patients with non-incentivised conditions may be at risk of poorer quality care.

Keywords

asthma; depression; hypertension; osteoarthritis; quality of health care; reimbursement.

INTRODUCTION

Many quality-improvement initiatives have been tried over the past 20 years, from education to clinical audit and guidelines. In the UK, recent policy interest in 'payment for performance' coincided with a renegotiation of GPs' contracts, resulting in a radical experiment in the use of financial incentives to improve quality. The new General Medical Services contract for general practices was introduced on 1 April 2004.¹ The Quality and Outcomes Framework (QOF) was part of this new contract, and provided practices with substantial financial rewards for achievement of quality indicators in 10 chronic conditions. The median reported achievement of clinical indicators by general practices was 97% in 2004–2005,² an achievement which could result in significant population health gain.³ The high standards of care reported are good news, but opinion has been divided over the likely effect of the incentives on conditions not included in the QOF.

The effect of financial incentives on 'non-incentivised' conditions is central to understanding the broader effects of payment for performance on patients; there is little independent evidence currently available.⁴ It has been argued that care for non-incentivised conditions may be relatively neglected,^{5–7} and that it could improve with development of a 'force that pulls all in its wake'.⁸ Quality of care for at least some of the conditions in

N Steel, MSc, MRCP, FFPH, senior lecturer; S Maisey, MSc, research associate; A Clark, PhD, statistician; R Fleetcroft, MSc, FRCGP, MRCP, ILTM, honorary senior lecturer; A Howe, MA, MD, MEd, FRCGP, professor, Primary Care Group, School of Medicine, University of East Anglia, Norwich.

Address for correspondence

Nicholas Steel, Primary Care Group, School of Medicine, Health Policy and Practice, University of East Anglia, Norwich, NR4 7TJ. E-mail: n.steel@uea.ac.uk

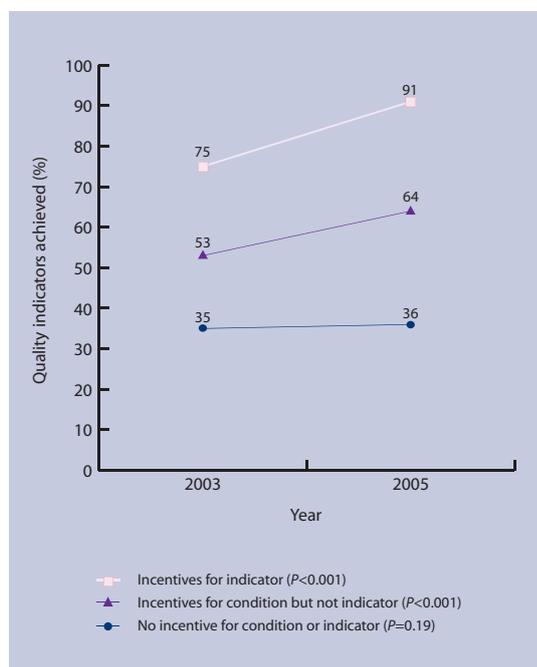
Submitted: 8 June 2006 Editor's response: 23 September 2006; final acceptance: 18 December 2006.

©British Journal of General Practice 2007; 57: 449–454.

How this fits in

Few studies have looked at patient data on quality of care actually provided for conditions incentivised and not incentivised in the Quality and Outcomes Framework, introduced in 2004 as part of the revised contract for general practices in the UK. This research suggests that quality of care improved substantially from 2003 to 2005 for care processes in clinical conditions that are subject to incentive payments. Quality of care did not appear to improve or deteriorate for care processes in clinical conditions that were not incentivised. The broader consequences of this focus on incentivised conditions remain to be seen.

Figure 1. Quality indicators achieved in 2003 and 2005 by incentive category. (See Table 3 for data)



the QOF was already at a high level and had improved significantly prior to 2004.⁹ It has been suggested that the extra payments have been for work done previously, and might not have stimulated further activity.^{10,11}

This study measured the recorded quality of care for four common chronic conditions, two of which were subject to financial incentives through the

QOF, and two of which were not. The study period was 1 year before (2003) and 1 year after (2005) the introduction of incentives.

METHOD

Quality of care for asthma, hypertension, osteoarthritis, and depression was measured in a sample of 18 general practices. These four conditions were selected because they are among the most common presenting to UK general practice,¹² they are predominantly managed in primary care, have lifelong implications, and have a range of published quality indicators available to define appropriate care. All QOF indicators for clinical management of hypertension and asthma were included, with the exceptions of smoking status for asthma, as this was already included in hypertension, and influenza immunisation, as this is only indicated in specific patient groups.

All quality indicators used were evidence based and peer reviewed, and came from at least one of four sources: the 2004 QOF,¹ the National Institute for Health and Clinical Excellence (NICE),¹³ RAND Health indicators adapted for the UK,¹⁴ and Quality Indicators for General Practice (QIGP) developed at the National Primary Care Research and Development Centre.¹⁵ There is a strong evidence base for all these indicators; indicators not in the QOF incentive scheme have been validated by independent expert panels including British GPs.

All quality indicators used in the study and their sources are listed in Supplementary Appendix 1. An example of an incentivised indicator is: 'the percentage of patients with hypertension in whom the last blood pressure (measured in last 9 months) is 150/90 [mmHg] or less'. An example of a non-incentivised indicator is: 'the percentage of patients receiving treatment for a new depression episode who were offered a follow-up appointment within 4 weeks'.

Practices in Norfolk primary care trusts were selected to give equal numbers in each of three groups stratified by national deprivation score, to ensure a spread of deprivation scores similar to the English profile. Practice deprivation scores were calculated as weighted means based on deprivation scores of each practice's registered patients.¹⁶ Between 20 and 40 randomly selected eligible patients with each condition at each practice were contacted for permission to examine their full patient records. The exact time period examined in each patient's notes varied according to the specification in each quality indicator. The sample size of 100 patients per time period and condition was estimated from previous work.¹⁷ Sampling from 2003 and 2005 was conducted

Table 1. Deprivation and size of practices in study compared with national practices.

Practice population	Practices, by thirds of national deprivation rank, %			Patients, by size of their practice, %		
	Most deprived	Middle	Least deprived	Small ^a	Medium ^b	Large ^c
England	33	33	33	24	43	33
Study	39	33	28	15	50	35

Data sources: General Medical Services databases 2003 and English Indices of Deprivation 2004. ^a501–5000 patients. ^b5001–10 000 patients. ^c>10 000 patients.

independently. Criteria for inclusion of patients in the study were specific for each clinical condition, and are given in Supplementary Appendix 2.

Data collection

Data were extracted from electronic and paper patient records to assess quality of care provided for asthma (seven indicators, two in the QOF, five not); hypertension (14 indicators, four in QOF, 10 not); osteoarthritis (nine indicators, none in QOF); and depression (six indicators, none in QOF). Inter-rater reliability between the research nurse and research associate extracting data was assessed by comparing data on all relevant indicators ($n = 225$) extracted by both researchers from 23 patient records. Seven patients had a diagnosis of depression, eight had asthma, and eight had hypertension (8% of all records examined by the research nurse).

Data analysis

Quality of health care was assessed as achievement of quality indicators. Each quality indicator referred to one of four health conditions, and was either subject to an incentive payment or not. For asthma and hypertension, both of which are health conditions included in the QOF, six quality indicators were subject to incentives, and 15 were not. For depression and osteoarthritis, neither of which was included in the QOF, none of 15 quality indicators was subject to incentives. Indicators were analysed in three groups: six indicators subject to an incentive; 15 indicators not subject to an incentive yet which referred to the care of a condition included in the QOF; and 15 indicators that referred to conditions excluded from the QOF.

The primary outcome measure was the number of indicators that were achieved for a patient. This number was expressed as a percentage of the

Table 2. Sample and response rates.

Condition	Patients, n		Mean age, years		Women, %	
	Study		Study		Study	
	Invited	population (%)	Invited	population	Invited	population
Asthma	402	253 (63)	45	50	53	56
Hypertension	399	304 (76)	67	67	58	59
Depression	529	279 (53)	45	50	69	72
Osteoarthritis	393	320 (81)	70	70	65	65
All	1723	1156 (67)	56	60	62	63

number of indicators for which that patient was eligible, assessed separately in different patients in 2003 and 2005. Analysis was based on a linear regression model of this percentage as the outcome variable; the indicator category and time point were independent variables; and changes over time (relative to osteoarthritis in 2003) were measured using a time-by-indicator interaction term in the model. Appropriate adjustments were made for clustering at individual patient and practice levels, using a hierarchical random effects model. All analyses were conducted using Stata (version 9.1 SE, Texas, US). The improvement as a percentage of maximum possible improvement was calculated as the percentage change multiplied by 100, divided by the percentage achieved in 2003 subtracted from 100. For example, if achievement in 2003 was 75%, and improved to 91% in 2005, the improvement as a percentage of maximum possible improvement is $[(91-75)/(100-75)] \times 100$, or 64%.

RESULTS

Data were collected from 1156 patient records: 586 from 2003 and 570 from 2005. The practices were nationally representative with respect to deprivation (scores ranged from 7 to 83% on a national practice deprivation scale), but slightly fewer patients were

Table 3. Quality indicators achieved in 2003 and 2005 by indicator category.

Presence of incentives for indicator and/or condition	Number of indicators	Number of patients		Mean indicators achieved, %		Difference, % (95% CI)	Change P -value		Improvement (% of max possible)
		2003	2005	2003	2005		2003–2005 ^a	Relative to category 3 ^b	
1. Incentives for indicator	6	279	278	75	91	16 (10 to 22)	<0.001	<0.001	64
2. Incentives for condition but not indicator	15	279	278	53	64	11 (6 to 15)	<0.001	<0.001	23
3. No incentive for condition or indicator	15	307	292	35	36	2 (–1 to 4)	0.19	—	3

^aFrom linear regression using patient level data, adjusting for clustering within practices. ^bFrom linear regression using patient level data, adjusting for clustering within practices and categories.

Table 4. Quality indicators achieved in 2003 and 2005 by condition.

Condition	Number of indicators	Number of patients		Mean indicators achieved, % (SD)		Difference, % (95% CI)	Change P-value ^a		Improvement (% of max possible)
		2003	2005	2003	2005	2003–2005	Relative to osteoarthritis		
Asthma	7	130	123	59 (24)	73 (23)	14 (8 to 20)	<0.001	0.004	34
Hypertension	14	149	155	58 (17)	70 (16)	12 (8 to 16)	<0.001	0.001	29
Depression	6	145	134	37 (14)	38 (14)	2 (-1 to 4)	0.22	0.92	3
Osteoarthritis	9	162	158	36 (19)	38 (22)	2 (-3 to 6)	0.43	–	3

^aFrom linear regression using patient level data, adjusting for clustering within practices.

registered with small practices in the study than nationally (Table 1). Patient response rate was 67% overall, and varied from 53% in depression to 81% in osteoarthritis (Table 2). The age and sex of non-responders were very similar to those of responders. Mean time taken for data extraction was 1.1 minutes per indicator, and there was no significant variation by condition or time point (2003 or 2005). Inter-rater reliability was high (K = 0.8, 95% confidence interval [CI] = 0.7 to 0.9).

Achievement of the quality indicators in the QOF showed the greatest improvement, rising from 75% to 91% (Table 3). Less improvement (53 to 64%) was seen for those indicators not subject to incentives, but referring to a condition in the QOF. These changes are highly significant in a regression model, allowing for the effect of clustering at patient and practice level and the different baseline quality

achievement. Indicators that referred to the conditions outside the QOF showed no significant improvement (35 to 36%). The number of patients eligible for, and number achieving each indicator at each time point, together with significance tests for changes are presented in Supplementary Appendix 2.

When the same indicators were grouped by clinical condition, achievement of asthma indicators improved significantly from 59% in 2003 to 73% in 2005, and hypertension improved from 58 to 70% (Table 4). In contrast, there was no significant improvement in osteoarthritis (36 to 38%) or depression (37 to 38%). Quality of care for the two conditions in the QOF (asthma and hypertension) started at a higher level than in the other two conditions, but still showed greater improvements when improvement was measured as a percentage

Table 5. Achievement of the Quality and Outcomes Framework quality indicators in the study and in nationally reported contract data.

Quality indicator	QMAS data 2004–2005 ^a		Study data 2005
	All 8577 practices in England (%)	18 practices in study (%)	18 practices in study (%)
Percentage of patients whose notes record smoking status at least once	94.2	94.8	99
Percentage of patients with hypertension who smoke, whose notes contain a record that smoking cessation advice or referral to a specialist service, if available, has been offered at least once	93.8	93.4	96
Percentage of patients with hypertension in whom there is a record of the blood pressure in the last 9 months	90.1	91.5	95
Percentage of patients with hypertension in whom the last blood pressure (measured in the last 9 months) is ≤150/90 mmHg	71.3	74.1	80
Percentage of patients aged ≥8 years diagnosed as having asthma from 1 April 2003 where the diagnosis has been confirmed by spirometry or peak flow measurement	90.1	91.5	95
Percentage of patients with asthma who have had an asthma review in the last 15 months	71.3	74.1	80
Mean of all 6 quality indicators	85.1	86.6	91

^aQMAS = Quality Management Analysis System 2004–2005, reporting Quality and Outcomes Framework data for contract payments.¹⁸

of the maximum possible improvement (Table 4). For asthma, the percentage of documented asthma reviews, which included all six of the elements recommended in the QOF, increased from 11% in 2003 to 25% in 2005.

Achievement on the quality indicators in the QOF was slightly higher in this study than the achievement reported nationally by the 18 study practices for payment under the revised contract (mean 91 compared to 87%) (Table 5). The nationally reported achievement for the study practices was similar to the figures for all practices in England (mean 87 compared to 85%).

DISCUSSION

Summary of main findings

Quality of care for incentivised conditions appeared to improve substantially from 2003 to 2005, with most improvement seen in the specific quality indicators directly subject to incentives through the QOF. A smaller but still significant improvement occurred in non-incentivised quality indicators linked to conditions included in the QOF. This study found a lower baseline level of quality and no significant improvement in quality for non-incentivised conditions.

Strengths and limitations of the study

All selected indicators measure aspects of clinical effectiveness, which is only one dimension of quality. Quality indicators can only measure specific aspects of care, and the relatively small number of indicators used in this study covers a small part of general practice, so no conclusions can be drawn about other aspects of care. For example, no measure of interpersonal care was included, and there is no way of knowing from this study whether the revised contract has altered interpersonal care.

Quality of care was assessed as higher at baseline for conditions in the incentive scheme, and it is possible that improvement in these conditions would have happened anyway. However, it could also be argued that further improvement for these conditions would be more difficult than for conditions with lower baseline quality. The practices were nationally representative in terms of deprivation and performance on the QOF, but patients registered with small practices were slightly under-represented. Larger clinical teams may deliver higher quality clinical care;¹⁹ therefore, the regression model was repeated including practice size as a sensitivity analysis, and the results did not change. It can be concluded that practice size is not a source of bias. Patient response varied by condition, but any bias in the sample for low response conditions (depression

and asthma) is unlikely to be responsible for the results, because one condition with a lower response (asthma) was included in QOF and one (depression) was not.

The study cannot show whether recorded data reflect actual performance. It is highly likely that incentives stimulated improved accuracy of electronic coding for QOF conditions between 2003 and 2005, and it is possible that these results simply reflect differential recording of care, rather than true improvements in performance. However, this is unlikely to explain the large changes seen for several reasons. Firstly, recording care is an essential component of effective team-based care for chronic diseases.²⁰ Secondly, both electronic and paper records were searched, and credit was given for any mention of the care, even if not fully documented. For example, credit for advising patients on non-steroidal anti-inflammatory drugs about gastrointestinal and renal risks was given for 'UWG' ('usual warning given') in the notes at any time, among other brief records. Thirdly, patients were included regardless of whether the practice had excluded them from their QOF returns. Practices were allowed to exempt patients under certain conditions, and there has been concern that exempting patients with poorer quality care could have led to artificially high scores. However, comparison of the present data with QOF returns (Table 5) shows that exception reporting is not a major reason for the high QOF scores.

One possible reason for high achievement levels in the QOF is that the indicators were consistent with GPs' professional values. All the non-incentivised indicators in this study were evidence based and validated by independent panels including British GPs, and so differences in professional values are unlikely to be an explanation for the differences seen.

There is currently no method for measuring the 'difficulty' of achieving an indicator, but the indicators here all went through a similarly rigorous development process, and so it is not likely that the non-incentivised indicators are overall inherently 'harder' to achieve. Even if it is argued that they are more difficult to achieve, which could explain the lower baseline level, it is still remarkable that quality did not improve at all for these conditions.

Comparison with existing literature

The quality score recorded for asthma in 2003 (59%) is lower than the 70% found by Campbell *et al* in 2003.⁹ This is most likely due to differences in the quality indicators used. Where the indicator is identical, the difference in scores between the two studies is negligible (60 compared with 59%).

Campbell *et al* found a 10% improvement in quality of asthma care over 5 years, compared with the 14% improvement over 2 years reported here. This suggests that, although care for asthma improved prior to the introduction of the new contract, the rate of improvement increased afterwards.

Implications for clinical practice

If these results were to be replicated, they suggest that payments to general practices were associated with improvements in quality of care for incentivised conditions only. The QOF demands a significant amount of work and effort, and so it may be expected that care for non-incentivised conditions would be marginalised.²¹ Expanding the QOF to cover more conditions may lead to improved quality of care in those conditions, although the inclusion of more conditions in the scheme might lead to dilution of the improvement effect. Alternatively a larger number of conditions could be included, with a smaller random number selected for payment at the end of the year. If a condition were to be rotated out of the framework, quality may deteriorate again.

This observational study of four conditions should be regarded as a preliminary assessment of the effects of financial incentives in the revised contract. The findings need to be validated in a larger study including more conditions and quality indicators, as well as measures of other important aspects of general practice, such as interpersonal care. Other observational methods such as video would help to validate the link between recorded and actual activity.

The strong association found in this study between financial incentives and quality of clinical care presents a simple analysis of a complex policy intervention. The effects of improved computer support, clinical prompts, and public reporting of quality information, which are important innovations in the new contract, are likely to be at least as important as the financial incentives.

Supplementary information

Additional information accompanies this article at <http://www.rcgp.org.uk/bjgp-suppinform>

Funding body

Nicholas Steel was funded by a Primary Care Researcher Development Award from the UK National Coordinating Centre for Research Capacity Development (RDA03/21). Norwich Primary Care Trust Research Office contributed to practice expenses for this project. Robert Fleetcroft receives protected time for continuing professional development from Great Yarmouth Teaching Primary Care Trust

Ethics committee

The Norwich Research Ethics Committee provided ethical approval of the study (REC 05/Q0101/37)

Competing interests

The authors have stated that there are none

Acknowledgements

We thank all the practice staff who helped with the study. Jonathan Cox of Norfolk NHS Public Health Network provided practice age and sex data and calculated practice deprivation scores. Max Bachmann, Richard Cookson, and Richard Holland commented on an earlier version of this paper.

REFERENCES

1. General Practitioners Committee BMA, The NHS Confederation. *Investing in general practice. The new General Medical Services contract*. London: The NHS Confederation, 2003.
2. Doran T, Fullwood C, Gravelle H, *et al*. Pay-for-performance programs in family practices in the United Kingdom. *New Engl J Med* 2006; **355**: 375–384.
3. McElduff P, Lyratzopoulos G, Edwards R, *et al*. 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**: 191–197.
4. Fleetcroft R, Cookson R. Do the incentive payments in the new NHS contract for primary care reflect likely population health gains? *J Health Serv Res Policy* 2006; **11**: 27–31.
5. Roland M. Linking physicians' pay to the quality of care — a major experiment in the United Kingdom. *New Engl J Med* 2004; **351**: 1448–1454.
6. Heath I. The cawing of the crow ... Cassandra-like, prognosticating woe. *Br J Gen Pract* 2004; **54**: 320–321.
7. Roland M, Campbell S, Bailey N, *et al*. Financial incentives to improve the quality of primary care in the UK: predicting the consequences of change. *Prim Health Care Res Dev* 2006; **7**: 18–26.
8. Lester H, Sharp DJ, Hobbs FD, Lakhani M. The quality and outcomes framework of the GMS contract: a quiet evolution for 2006. *Br J Gen Pract* 2006; **56**: 244–246.
9. Campbell SM, Roland MO, Middleton E, Reeves D. Improvements in quality of clinical care in English general practice 1998–2003: longitudinal observational study. *BMJ* 2005; **331**: 1121.
10. Maynard A, Street A. Seven years of feast, seven years of famine: boom to bust in the NHS? *BMJ* 2006; **332**: 906–908.
11. Timmins N. Do GPs deserve their recent pay rise? *BMJ* 2005; **331**: 800.
12. McCormick A, Fleming D, Charlton J. *Morbidity statistics from general practice. 4th national study 1991–1992*. London: HMSO; 1995.
13. National Institute for Health and Clinical Excellence. www.nice.org.uk (accessed 29 Mar 2007).
14. Steel N, Melzer D, Shekelle PG, *et al*. Developing quality indicators for older adults: transfer from the USA to the UK is feasible. *Qual Saf Health Care* 2004; **13**: 260–264.
15. Marshall M, Campbell S, Hacker J, Roland M. *Quality indicators for general practice. A practical guide for health professionals and managers*. London: Royal Society of Medicine Press, 2002.
16. Noble M, Wright G, Dibben C, *et al*. *The English indices of deprivation 2004*. Wetherby: ODPM Publications, 2004.
17. Steel N, Huppert FA, McWilliams B, Melzer D. Physical and cognitive function. In: Marmot M, Banks J, Blundell R, *et al* (eds). *Health, wealth and lifestyles of the older population in England: the 2002 English Longitudinal Study of Ageing*. London: The Institute for Fiscal Studies, 2003; 249–300.
18. Quality and Outcomes Framework 2004/05. www.ic.nhs.uk/services/qof/data/ (accessed 29 Mar 2007).
19. Sutton M, McLean G. Determinants of primary medical care quality measured under the new UK contract: cross sectional study. *BMJ* 2006; **332**: 389–390.
20. Hamilton WT, Round AP, Sharp D, Peters TJ. The quality of record keeping in primary care: a comparison of computerised, paper and hybrid systems. *Br J Gen Pract* 2003; **53**: 929–933.
21. Maynard A, Bloor K. Do those who pay the piper call the tune? *Health Policy Matters* 2003; **8**: 1–8.