Developing performance indicators for primary care: Walsall's experience

Narinder Sahota, Andrew Hood, Anandagiri Shankar, Barbara Watt and Sam Ramaiah

ABSTRACT

Background

There has been increasing interest in the development of performance indicators in primary care, especially since the introduction of the Quality and Outcomes Framework (QOF). Public health and primary care trusts collect a range of data from routine or non-routine sources that may be useful for this purpose.

Aim

To assess whether performance against the QOF is a robust measure of practice performance when compared with health-inequality indicators and to contribute to the development of a tool to monitor and improve primary care services.

Design of study

A retrospective cross-sectional study.

Setting

Sixty-three GP practices contracted with Walsall Teaching Primary Care Trust.

Method

Correlation analysis and scatter plots were used to identify possible significant relationships between QOF scores and health-inequality data. The study also utilised confidence limit theory and control chart methodology as tools to identify possible performance outliers.

Results

Little correlation was found between overall QOF score and deprivation score. Uptake of flu immunisation (r^2 = 0.22) and cervical screening (r^2 = 0.11) both showed a slight increase with increased QOF score. Benzodiazepine (r^2 = 0.06) and antibiotic prescribing levels (r^2 = 0.02) decreased slightly with increased QOF scores, although not significantly. An increase in practice-population deprivation score was correlated with a reduction in cervical screening uptake (r^2 = 0.27) and an increase in benzodiazepine prescribing (r^2 = 0.25). Statistically significant relationships were found between the patient: GP ratio and flu immunisation uptake (r^2 = 0.1) and antibiotic prescribing (r^2 = 0.1). The majority of GPs found it acceptable to use performance indicator data as part of their annual appraisal.

Conclusion

QOF and health-inequality data can be used together to measure practice performance and to develop tools to help identify areas for performance development and the sharing of best practice.

Keywords

general practice; quality indicators; performance indicators; primary health care; public health.

INTRODUCTION

There has been increasing interest in performance indicators in primary care, especially since the introduction of the Quality and Outcomes Framework (QOF) for general practice in the UK in 2004.[1] The QOF measures achievement against a scorecard of clinical and non-clinical indicators (146 indicators were included in 2004/2005), but it presents only a partial picture of practice performance. Primary care trusts (PCTs) hold data for a number of other useful indicators, such as prescribing rates, screening uptake, and patient surveys. Walsall Teaching PCT monitors 123 performance indicators, including Accident & Emergency attendance, smoking in pregnancy, and access to primary care professionals. Such indicators have not previously been brought together in a systematic way to provide a wider view of primary care than can be provided by QOF data alone.

This exploratory study aimed to assess whether performance against QOF criteria alone is a robust measure of practice performance, or whether QOF data combined with a variety of other health-inequality indicators provides a better measure. The influence of practice-population deprivation status on practice achievement was also

***N Sahota**, DCH, MRCGP, associate medical director of primary care and GP; **A Hood**, BA, senior public health intelligence analyst; **A Shankar**, MD, MPH, specialist registrar in public health medicine; **B Watt**, MSc, MPH, PhD, MFPH, locum consultant in public health; **S Ramaiah**, FFPHM, FRCP, medical director and director of public health, Department of Public Health, Walsall Teaching Primary Care Trust, Walsall.*

Address for correspondence

Dr Narinder Sahota, Department of Public Health, Walsall Teaching Primary Care Trust, Jubilee House, Bloxwich Lane, Walsall WS2 7JL. E-mail: narinder.s.sahota@Walsall.nhs.uk

Submitted: 6 December 2007; **Editor's response:** 4 March 2008; **final acceptance:** 11 July 2008.

This article was originally online first. Cite this article as: *Br J Gen Pract* 2008; **58:** 856–861. Advance online publication. DOI: 10.3399/bjgp08X376096

examined. The study further aimed to provide practices, practice-based commissioning clusters,[2] and PCTs with performance benchmarking measures and tools to help to identify areas for performance development, as well as examples of best practice. It is intended that these indicators and tools will be used to complement the annual contract review visits and QOF visits already carried out by the teaching PCT.

How this fits in

The Quality and Outcomes Framework (QOF) is a used as a performance tool for general practice. Questions have been raised as to how good QOF is as a measure of performance. QOF achievement alone is not necessarily a good indicator of performance; however, when used in combination, QOF and health inequality data can provide a wider measure of performance. Little correlation was found between overall QOF score and deprivation score.

METHOD

Selection of indicators

A wide range of possible indicators is available. For this exploratory study, it was decided to focus on indicators that were largely based on GP-initiated care and amenable to changes, and for which data were readily available; for example, the prescribing of antibiotics or benzodiazepines is largely dependent on the GP. The following indicators were chosen:

- QOF score;
- antibiotic prescribing;
- benzodiazepine prescribing;
- cervical screening uptake;
- flu immunisation uptake; and
- MMR (measles, mumps and rubella) immunisation uptake.

A number of other indicators were considered but rejected; for example, smoking cessation (as there are a number of service providers other than GPs in Walsall), and patient:nurse ratio (as data were incomplete).

Data sources

Data for the financial year 2004/2005 were obtained for all 63 GP practices in Walsall from the Quality Management and Analysis System (which supports the QOF), the Office for National Statistics' Indices of Multiple Deprivation (IMD 2004),[3] Child Health Systems (immunisation data), the Prescription Prescribing Authority (prescribing data), and the Exeter system (patient registration and cervical screening).

QOF data for three practices were excluded from the analysis; this was due to one GP from one practice retiring part-way through the financial year and incomplete QOF returns from the other two practices.

Analysis of data

The analysis in this study is based on only the four principal QOF domains (clinical, organisational, additional services, and patient experience). When added together to obtain a composite score, these domains accounted for a maximum 870 of the total 1050 points achievable under the QOF.[1] Practice-level scores for the other domains were unavailable at the time that this analysis was undertaken.

Correlation analysis

The relationship between practice performance (as measured by practice QOF scores for each of the four domains and overall composite QOF score) and practice achievement in the five selected 'health inequalities' indicators was explored. Data were analysed for significant relationships, using scatter plots and correlation analysis.

The relationship between the various indicators against practice-population deprivation status and patient:GP ratio was also investigated. The deprivation index used in this study was calculated at practice level, based on the percentage of registered patients living in the most deprived lower super output areas (SOAs) in Walsall, according to IMD 2004. The use of SOAs avoids the problems caused by the inconsistent and unstable electoral ward geography. Lower SOAs have a minimum size of 1000 residents and 400 households, but average 1500 residents.

Traffic-light system

For each indicator, 95% confidence intervals (CIs) were calculated around the Walsall mean value to provide suitable thresholds for the identification of outliers. A traffic-light system was developed, whereby any practice falling outside the Walsall 95% CI was appropriately colour coded: red to indicate 'under'-performance and green for 'over'-performance. The remaining practices were coded as amber.

Given the demographics and socioeconomic profile of Walsall, and the fact that three-quarters of practices are small (fewer than three GPs) in size, it was felt that GPs would find it more acceptable to be compared with other practices in the borough. The Walsall figure was, therefore, used as the benchmark, rather than regional or national figures.

To provide an alternative means of presenting the performance data, control charts were also plotted (using 99.8% CIs) for each indicator versus practice list size.

Table 1. Correlation coefficients ('*r*' values) between study variables for 2004/2005 performance data.

	IMD 2004 (% in group 1)	QOF score (sum of four domains)[a]	Patient:GP ratio	Cervical screening uptake	Flu immunisation uptake[b]	Benzodiazepine prescribing	Antibiotic prescribing	MMR immunisation[c]
IMD 2004 (% in group 1)	–							
QOF score (sum of four domains	–0.160	–						
Patient:GP ratio	0.091	–0.176	–					
Cervical screening uptake	–0.515[e]	0.336[d]	–0.115	–				
Flu immunisation uptake[a]	–0.160	0.468[e]	–0.315[d]	0.494[e]	–			
Benzodiazepine prescribing	0.496[e]	–0.237	0.021	–0.287[d]	–0.234	–		
Antibiotic prescribing	0.139	–0.141	0.316[d]	0.052	–0.235	0.299[d]	–	
MMR immunisation uptake[b]	–0.053	0.214	0.190	0.096	–0.132	0.062	0.182	–

[a]Clinical, organisational, management, and patient experience. [b]In patients aged over 65 years. [c]At 2 years of age. [d]P<0.05 for correlation coefficient. [e]P<0.01 for correlation coefficient. IMD = Indices of Multiple Deprivation. Group 1 = the proportion of practice population living in the 20% most deprived lower super output areas nationally.

RESULTS

There are 63 practices in Walsall, three of which were excluded from the analysis. The results for the remaining 60 practices are presented here.

Correlation analysis

Little evidence was found of an association between practice achievement in the four individual QOF domains (clinical, organisational, management, and patient experience); however, they were all positively correlated with both cervical screening and flu immunisation uptake. There was a significant association between total QOF score (combined score for all four domains) and both cervical screening uptake ($r^2 = 0.11$) and flu immunisation uptake ($r^2 = 0.22$). MMR immunisation uptake had a positive weak correlation with QOF score, but this was not statistically significant. Cervical screening uptake was strongly associated with increased flu

Figure 1. Indicator scores and red/amber/green ratings for the best and worst of QOF-achieving practices in Walsall (2004/2005).

Practice ID	QOF score (sum of four domains[a]), %	Patient:GP ratio	Benzodiazepine prescribing	Antibiotics prescribing	Cervical screening uptake (%)	Flu immunisation uptake (%)	MMR immunisation uptake (%)
1	27.8	2690	7.496	1.223	76.5	72.6	58.6
2	49.3	2421	10.491	0.828	78.4	40.6	87.0
3	52.7	2845	15.582	1.401	76.4	62.0	87.3
4	56.8	2820	9.479	0.955	70.7	54.1	85.2
5	63.9	6686	4.45	0.921	80.3	62.0	78.0
16	82.5	3585	5.464	1.319	78.3	71.9	81.0
23	88.7	1364	14.263	0.936	80.0	64.0	76.2
30	91.1	1877	7.731	1.491	87.6	75.8	88.9
35	93.0	2219	5.740	1.310	85.9	75.9	81.9
41	94.5	1402	10.747	0.997	86.6	76.2	88.7
58	99.3	1992	4.031	0.929	79.9	77.4	78.6
59	99.7	3017	5.326	1.083	89.1	70.2	81.5
60	99.9	1512	5.496	1.081	81.4	75.2	84.1
61	99.9	1314	7.690	0.663	81.8	72.0	78.9
62	99.9	1814	7.955	0.812	84.4	73.8	73.8

Red: worse than Walsall average 95% confidence intervals.

Amber: within the range of Walsall average 95% confidence intervals.

Green: better than Walsall average 95% confidence intervals.

[a]Clinical, organisational, management, and patient experience. MMR = measles, mumps and rubella. QOF = Quality and Outcomes Framework.

immunisation uptake (r^2 = 0.24). Benzodiazepine (r^2 = 0.06) and antibiotic prescribing levels (r^2 = 0.02) decreased slightly with increased QOF scores, although not significantly.

No strong association was found between patient:GP ratio and QOF scores; however, eight out of the 10 practices with the lowest QOF scores had a patient:GP ratio of >2400:1. Two of the top 10 QOF-scoring practices had a patient:GP ratio of <2400:1. A higher patient:GP ratio was associated with decreased flu immunisation uptake (r^2 = 0.1) and increased antibiotic prescribing (r^2 = 0.1).

Deprivation of the practice population was significantly correlated with a reduction in cervical screening uptake (r^2 = 0.27) and was also linked to an increase in benzodiazepine prescribing (r^2 = 0.25). However, deprivation scores were not significantly correlated with QOF score (Table 1).

Traffic-light system

The results obtained using the 'traffic-light' approach are shown in Figure 1. Significant 'over'-performance compared with the Walsall average was designated as green, and 'under'-performance as red. Practices are shown in ascending order of QOF achievement; the highest five QOF scores (green), lowest five (red), and five intermediate scores (amber) have been included.

In general, those practices with lower QOF scores achieved lower scores across the other indicators, and those with high QOF achievement scored higher in the other indicators (Figure 1). Scores across the intermediate QOF achievers were more variable for the other indicators.

Control charts

Control charts (also known as funnel plots or Shewhart charts) are an alternative means of highlighting variation in performance and can provide a useful visual comparison tool.[4,5] The configuration of a control chart consists of a centre line that represents the mean of the data set, and control limits that are calculated to represent ± three standard deviations (99.8% CI). Control limits are usually set at three standard deviations of the mean to minimise the number of outliers, so that resources can be concentrated on investigating only the more extreme outliers.

For many of the charts, the control limits were narrow, with the result that the majority of practices were outliers for most indicators, that is, 'under'- or 'over'-performing. Antibiotic prescribing rates produced the most interesting control charts, due to lower levels of variability in the data. They showed three practices that were prescribing at rates significantly above the Walsall average and three practices that were prescribing at significantly lower rates (Figure 2).

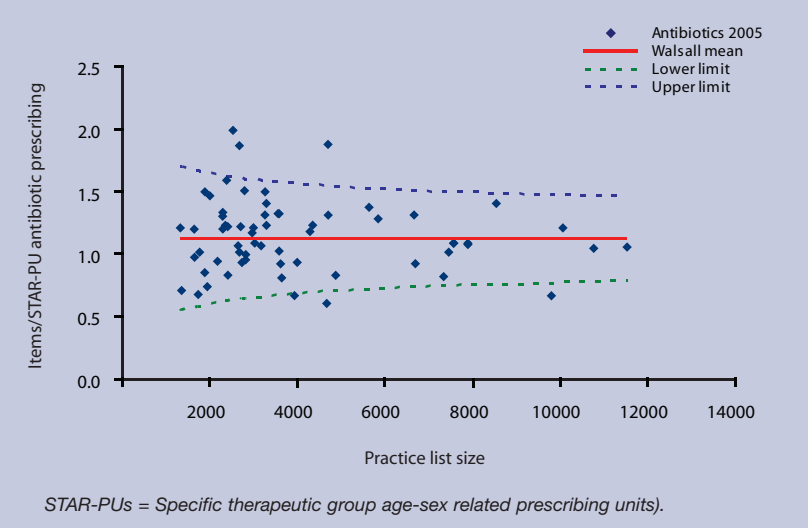

Figure 2. Control chart of antibiotic prescribing in 2005 against list size with 99.8% control limits around the Walsall mean.

DISCUSSION

Summary of main findings

This appears to be the first study to use QOF and health-inequality data together with a view to developing primary care performance-improvement tools. This article shows that QOF achievement alone is not necessarily a good indicator of performance in some health domains but that QOF and health-inequality data can be used in combination to provide a wider measure of performance. The data can be used to develop screening tools (for example, traffic-light tables and control charts) to identify outlier practices and help to identify areas for personal development through GP appraisals. Such screening tools also offer the opportunity to share excellence, as 'over'-performing practices can be readily identified and encouraged to share good practice with GP colleagues.

Strengths and limitations of the study

It can be argued that the bringing together of a number of different indicators can provide a wider view of primary care performance than QOF data used in isolation. In particular, the inclusion of indicators that have no financial incentive attached adds an interesting dimension to practice assessment. However, more work is needed to ensure that consideration of important 'human' aspects of general practice, such as empathy compassion, trust, and attitude, are incorporated into any developed tool, for example by the inclusion of the results of patient surveys.

Cervical screening uptake and some elements of flu immunisation uptake were components of the

QOF in 2004/2005 (as well as in subsequent years). However, these contributed only a maximum of 2.3% and 3.3% respectively to the total possible QOF scores used in this study. As the potential contribution was small, it was concluded that it was reasonable to include these indicators in the correlation analysis.

The sample size in this study was 63 practices. Although this is not large, it is above the average for PCTs in England (28 in 2004/2005 prior to reorganisation, and 55 post reorganisation).

One of the possible limitations with the data sources used could be their timeliness, as most of the data were >12 months old when they became available. It is hoped that, in future, data will be available from routine and other data sets within 6 months of the end of the financial year.

Comparison with existing literature

In the present study, no link was found between deprivation levels and QOF scores. Recent studies investigating the relationship between deprivation and QOF achievement have produced mixed findings.[6–8] A study of practices in England showed higher QOF scores for practices in less-deprived areas,[9] while a study of practices in Scotland showed that deprivation was associated with higher QOF scores.[10] A study of 1024 general practices in Scotland showed little systematic association between QOF achievement and deprivation.[11]

A further study has shown that smaller practices tended to obtain fewer QOF points than larger practices.[11] The present data did not support this; one reason could be that the majority (72%) of practices studied were small and had fewer than three GPs.

Another study, which analysed data at PCT level, found little evidence that current indicators (with the exception of screening indicators) have sufficient validity to measure 'the underlying concept of quality'.[12] They also showed a negative linear correlation between the IMD 2004 and QOF unlike the current study.

Implications for future research and clinical practice

The data were presented to the teaching PCT's performance committee as well as groups of GPs, including the local medical committee. In general, the results were seen as useful. GP colleagues were not disinclined to consider the continuation of this approach and it was agreed that the data may be more revealing when set out as a trend over time.

These data were also used during GP appraisals, with a view to identifying areas for GP performance development and learning needs. For example, if they were high antibiotic prescribers this might be considered in their personal development plan.

This work should be seen as an initial exploratory effort to utilise routinely available data for benchmarking GP performance. It is acknowledged that some practices, for a variety of reasons (both systematic and due to external influences), fall some way below average achievement. The results of this study will complement the existing processes that exist within the PCT to identify and work with those practices that require development and support.

There is a plan to repeat this exercise annually to identify trends. The researchers will also look at what other performance data can be included, such as mortality, referral patterns, hospital admissions, and patient surveys. The indicators used will therefore change and develop in future years and it is likely that new indicators will replace ones that have become less relevant. The strategic health authority is continuing to develop a framework for the implementation of a 'balanced scorecard' to assess the quality of primary care; it is possible that future versions of this study will merge with that process.

This present study shows that assessing performance is complex and that performance cannot be viewed as a unitary entity that is 'good practice' or 'bad practice'. No clear correlation was found when looking at national (QOF) performance compared with the researchers' 'home-grown' multiple indicator approach. The authors believe that there is an inevitable trend to develop a composite tool for assessing performance. If general practice does not contribute to the methodology of performance improvement then other bodies will, including entrants to primary care from the commercial sector. This study goes some way to developing a rationale for a composite tool that takes into account QOF and other data. More research needs to be done to identify those indicators that are good at measuring performance in primary care.

Ethical approval

Not applicable

Competing interests

Dr Narinder S Sahota is a practising GP in Walsall. All other authors have stated that there are none

Acknowledgements

The authors would like to acknowledge the help of Professor Nick Pugh and Dr Susan Laverty in reviewing the manuscript and suggesting improvements. The administration support received from Mrs Sally Andrews is also gratefully acknowledged.

Discuss this article

Contribute and read comments about this article on the Discussion Forum: http://www.rcgp.org.uk/bjgp-discuss

REFERENCES

1. Department of Health. Rewarding quality and outcomes. *In:* Department of Health. *The new GMS contract.* London: The Stationery Office, 2003: 17–24.
2. Department of Health. *Practice based commissioning: promoting clinical engagement.* London: Department of Health, 2004. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4098564 (accessed 28 Aug 2008).
3. Office of the Deputy Prime Minister. *The English Indices of Deprivation 2004: summary (revised).* http://www.communities.gov.uk/publications/communities/englishindices (accessed 30 Sep 2008).
4. Mohammed MA, Booth K, Marshall D, *et al.* A practical method for monitoring general practice mortality in the UK: findings from a pilot study in a health board of Northern Ireland. *Br J Gen Pract* 2005; **55(518):** 670–676.
5. Marshall T, Mohammed MA, Lim HT. Understanding variation for clinical governance: an illustration using the diagnosis and treatment of sore throat. *Br J Gen Pract* 2002; **52(477):** 277–283.
6. Ashworth M, Seed P, Armstrong D, *et al.* The relationship between social deprivation and the quality of primary care: a national survey using indicators from the UK Quality and Outcomes Framework. *Br J Gen Pract* 2007; **57(539):** 441–448.
7. Saxena S, Car J, Eldred D, *et al.* Practice size, caseload, deprivation and quality of care of patients with coronary heart disease, hypertension and stroke in primary care: national cross-sectional study. *BMC Health Serv Res* 2007; **7:** 96.
8. Strong M, Maheswaran R, Radford J. Socioeconomic deprivation, coronary heart disease prevalence and quality of care: a practice-level analysis in Rotherham using data from the new UK general practitioner Quality and Outcomes Framework. *J Public Health (Oxf)* 2006; **28(1):** 39–42.
9. Ashworth M, Armstrong D. The relationship between general practice characteristics and quality of care: a national survey of quality indicators used in the UK Quality and Outcomes Framework, 2004–5. *BMC Fam Pract* 2006; **7:** 68.
10. Sutton M, McLean G. Determinants of primary medical care quality measured under the new UK contract: cross sectional study. *BMJ* 2006; **332(7538):** 389–390.
11. Wang Y, O'Donnell CA, Mackay DF, Watt CG. Practice size and quality attainment under the new GMS contract: a cross-sectional analysis. *Br J Gen Pract* 2006; **56(532):** 830–835.
12. Brown C, Lilford R. Cross sectional study of performance indicators for English Primary Care Trusts: testing construct validity and identifying explanatory variables. *BMC Health Serv Res* 2006; **6:** 81.