

Mark Ashworth, Peter Schofield, Tim Doran, Richard Cookson, Matthew Sutton, Paul T Seed, Amanda Howe and Robert Fleetcroft

The Public Health Impact score:

a new measure of public health effectiveness for general practices in England

Abstract

Background

Health policy in the UK is increasingly focused on the measurement of outcomes rather than structures and processes of health care.

Aim

To develop a measure of the effectiveness of primary care in terms of population health outcomes.

Design and setting

A cross-sectional study of general practices in England.

Method

Twenty clinical quality of care indicators for which there was evidence of mortality reduction were identified from the national Quality and Outcomes Framework (QOF) pay-for-performance scheme. The number of lives saved by 8136 English practices (97.97% of all practices) in 2009/2010 was estimated, based on their performance on these measures, and a public health impact measure, the PHI score, was constructed. Multilevel regression models were used to identify practice and population predictors of PHI scores.

Results

The mean estimated PHI score was 258.9 (standard deviation [SD] = 73.3) lives saved per 100 000 registered patients, per annum. This represents 75.7% of the maximum potential PHI score of 340.9 (SD = 91.8). PHI and QOF scores were weakly correlated (Pearson $r = 0.28$). The most powerful predictors of PHI score were the prevalence of the relevant clinical conditions ($\beta = 0.77$) and the proportion of patients aged ≥ 65 years ($\beta = 0.22$). General practices that were less successful at achieving their maximum potential PHI score were those with a lower prevalence of relevant conditions ($\beta = 0.29$), larger list sizes ($\beta = -0.16$), greater area deprivation ($\beta = -0.15$), and a larger proportion of patients aged ≥ 65 years ($\beta = -0.13$).

Conclusion

The PHI score is a potential alternative metric of practice performance, measuring the estimated mortality reduction in the registered population. Rewards under the QOF pay-for-performance scheme are not closely aligned to the public health impact of practices.

Keywords

health outcomes; population mortality reduction; primary health care.

INTRODUCTION

Pay-for-performance schemes have been introduced in several healthcare systems, with the aim of improving quality of care and patient outcomes.¹⁻³ While some schemes appear to have improved the performance of processes of care,³ evidence for improvements in health outcomes and population health is lacking, leading to a call for the development of a 'pay-for-population-health performance system'.⁴

In the UK, the 'Quality and Outcomes Framework' (QOF) pay-for-performance scheme was introduced in 2004, providing financial incentives to family practices for achieving targets for over 100 quality of care indicators.³ Payments for these indicators are currently weighted on the basis of predicted GP workload and not on health outcomes. Health policy in the UK is shifting towards an emphasis on the measurement of outcomes rather than structures and processes of health care.^{5,6} This research therefore aimed to develop a measure of primary care effectiveness in terms of population health outcomes. This measure would be based on practice quality indicators selected from the current indicators included in the QOF, weighted according to their potential for mortality

reduction, and applicable to all practices in England. This measure was termed the 'Public Health Impact' ('PHI') score. The use of the PHI score to define variations in the delivery of health outcome, and their relationship with practice and population characteristics, were also explored.

METHOD

Data

QOF data were obtained for all general practices in England, covering the year 2009/2010, from the NHS Information Centre.⁷

A detailed summary of practice characteristics was obtained from the general medical services database, including practice list size, age and sex of the registered population, number of full-time equivalent (fte) GPs, and GP training practice status.⁷

Demographic data covering ethnicity (from the 2001 national UK census) and social deprivation (Index of Multiple Deprivation 2010) were obtained for lower layer super output areas, which cover a mean population of 1500 residents.^{8,9} Pooled demographic data from these localities were used as a proxy for the characteristics of the registered population of each practice.¹⁰

M Ashworth, DM, FRCGP, MRCP, clinical senior lecturer; **P Schofield**, PhD, senior research fellow, Department of Primary Care and Public Health Sciences, King's College London School of Medicine. **T Doran**, MD, MFPH, clinical research fellow, School of Health Sciences, University of Manchester. **R Cookson**, PhD, reader, Centre for Health Economics, University of York. **M Sutton**, PhD, professor of health economics, Health Sciences Research Group, University of Manchester. **PT Seed**, MSc, CStat, senior lecturer in medical statistics, Division of Women's Health, King's College, London. **A Howe**, MD, FRCGP, FAcadMed, professor of primary care; **R Fleetcroft**, MD, FRCGP, MRCP, lecturer in general practice, Department of Population

Health and Primary Care, Norwich Medical School, University of East Anglia.

Address for correspondence

Mark Ashworth, Department of Primary Care and Public Health Sciences, King's College London School of Medicine, Capital House, 42 Weston Street, London SE1 3QD.

E-mail: mark.ashworth@kcl.ac.uk

Submitted: 22 July 2012; **Editor's response:**

12 September 2012; **final acceptance:** 2 November 2012.

This is the full-length article (published online 25 Mar 2013) of an abridged version published in print. Cite this article as: **Br J Gen Pract 2013; DOI: 10.3399/bjgp13X665260**

How this fits in

This study has taken a series of 20 QOF indicators, converted achievement of these into estimates of mortality reduction based on published evidence, and derived a composite score for the sum total of mortality reduction attributable to these QOF indicators. For the first time, general practices will have data to describe their effectiveness in terms of an estimate of mortality reduction. At an individual level and based on 2009/2010 figures, the average GP in England saved 4.7 lives per year through disease prevention activity.

Study design

A retrospective cross-sectional study was carried out.

Participants

QOF data were available for all 8305 general practices in England. Of these, 159 were excluded from the analysis on the basis of their list size (<750 registered patients) or list size per GP (<500 patients per fte GP), since practices with such small list sizes are likely to be highly atypical. A further 10 practices were omitted on the basis of incomplete data. The analysis was conducted on the remaining 8136 general practices (97.97% of all practices).

Construction of the PHI score

Selection of indicators. Twenty clinical QOF indicators for which there was evidence of mortality reduction were identified. These were based on 25 QOF indicators identified by Fleetcroft *et al*, which, with subsequent changes in the QOF, have now been subsumed into 19 QOF indicators.^{11,12} Estimates of mortality reduction for each of these indicators are displayed in Table 1. QOF indicators not included in Fleetcroft's original analysis were re-examined, and it was considered that there was sufficient evidence to include one further indicator, the cervical screening indicator (Table 1).¹³ The term 'QOF(20)' was used to describe this collection of indicators. QOF indicators that did not feature in the selection of QOF(20) indicators lacked a sufficient evidence base to warrant their inclusion. Although there are other measures of public health activity in primary care, only QOF indicators were considered for this analysis, to ensure consistency of data collection.

Mortality-reduction estimates for selected indicators. Estimates of mortality reduction were obtained, derived from the available

literature, identifying the highest level of evidence for risk reduction in all-cause mortality.^{11–13} Risk reduction estimates, defined as absolute risk reduction (ARR), relative risk reduction (RRR), or as odds ratios (ORs), were converted into estimated mortality reduction rates per 100 000 population, per annum (Table 1).

In deriving the estimates of mortality reduction, a conservative interpretation was used and the original published estimates were retained, even though in subsequent years the targets for two of the targets for diabetes were tightened (DM23 and DM25) and the smoking cessation indicator (Smoking4) was broadened to cover additional conditions.¹⁴

Comorbidity correction factor. The total estimated mortality reduction achieved by each practice will be less than the sum of the 20 individual indicators, since many patients have comorbidity. In the absence of patient-level data in the QOF database, a further QOF indicator (Smoking3) was used to calculate a proxy for comorbidity. The Smoking3 denominator is the sum of all patients with any one of eight chronic conditions regardless of their smoking status: all six included in the present study plus asthma and psychosis.¹⁴ The sum of individual prevalences at national level for these eight conditions was 34.7%, whereas the sum of combined prevalences (the denominator for Smoking3) was 21.5%. Therefore, the overall mortality reduction corrected for comorbidity by the national mean equals $21.5/34.7$, that is, a correction of 62.4%. For the analysis in this study, each practice was allocated its individual comorbidity correction factor.

Prevalence calculation for selected indicators. Practices with higher prevalence of each condition have a greater potential for mortality reduction, and therefore this factor was included in the calculation.

A composite practice-level value was derived for the overall prevalence of all conditions included in the mortality estimates. This value was based on the mean ratio of practice prevalence to national prevalence for each of the QOF(20) indicators, weighted according to the estimated mortality reduction of each indicator and according to practice-level comorbidity. Thus a practice with double the national prevalence of one of the conditions associated with greater mortality reduction would have a higher overall composite prevalence value than a practice with double the prevalence

Table 1. Summary of clinical indicators included in the Public Health Impact (PHI) score and estimated mortality reduction for each indicator

QOF indicator	Summary description of indicator	Crude prevalence per 100 000 registered patients, mean (SD)	Annual mortality reduction, per 100 000 registered patients
DM18	Diabetes: influenza vaccination	4420 (1881)	63.7
CHD12	CHD: influenza vaccination	3448 (1487)	61.6
BP5 ^a	Hypertension: BP ≤150/90 mmHg	13 548 (5117)	48.2
CHD10 ^a	CHD: beta-blocker treatment	3448 (1487)	45.9
STROKE10	Stroke/TIA: influenza vaccination	1649 (967)	28.1
DM23 ^a	Diabetes: HbA1c ≤7.0%	4420 (1881)	26.5
CKD3	CKD: BP ≤140/85 mmHg	3229 (2026)	26.2
COPD8	COPD: influenza vaccination	1626 (958)	24.9
CHD9 ^a	CHD: aspirin or other antithrombotic therapy	3448 (1487)	24.8
AF3 ^a	AF: warfarin or other antithrombotic therapy	1324 (717)	22.7
CHD8 ^a	CHD: cholesterol ≤5.0 mmol/l	3448 (1487)	15.8
STROKE12 ^a	Stroke (non-haemorrhagic): aspirin or other antithrombotic therapy	1080 (649)	15.8
DM12	Diabetes: BP ≤145/85 mmHg	4420 (1881)	13.5
HF3 ^a	Heart failure: ACEI or ARB therapy	388 (283)	13.0
CHD6 ^a	CHD: BP ≤150/90 mmHg	3448 (1487)	11.3
SMOKING4	CHD, stroke/TIA, hypertension, DM, CKD, COPD, asthma, psychosis: smoking cessation advice	3903 (2525)	10.9
DM25	Diabetes: HbA1c ≤9.0%	4420 (1881)	7.4
DM15 ^a	Diabetes with proteinuria or microalbuminuria: ACEI or ARB therapy	505 (513)	3.4
CHD11 ^a	CHD (myocardial infarction): ACEI or ARB therapy	572 (291)	1.5
CS1	Women aged 25–64 years with cervical smear in last 5 years ^b	25 253 (17849)	0.8

ACEI = angiotensin-converting enzyme inhibitor. AF = atrial fibrillation. ARB = angiotensin receptor blocker. BP = blood pressure. CHD = coronary heart disease. CKD = chronic kidney disease. COPD = chronic obstructive pulmonary disease. DM = diabetes mellitus. SD = standard deviation. TIA = transient ischaemic attack. ^aIndicators included in PH(11) score. ^bThe value stated for 'prevalence' refers to the total number of eligible women in the 25–64 years age group, per 100 000 registered patients.

of a condition lower down the rankings of mortality reduction. This composite prevalence score was termed the 'prevalence factor'. The national mean prevalence value was set at 1.0.

Calculation of practice-level mortality-reduction estimates. Estimates of overall mortality reduction in the registered population of each practice were calculated as a function of the clinical indicator achievement percentage, ARR, comorbidity, and prevalence (for details, see Appendix 1).

Measures of public health impact. Three measures of PHI were derived for each practice:

- the PHI score: the estimated mortality reduction, per 100 000 registered patients, per annum;
- the maximum potential PHI score: the estimated mortality reduction assuming 100% achievement of each of the 20 indicators included in the PHI score; and
- the PHI% performance score: the PHI score achievement for each practice as a percentage of the maximum potential PHI score.

Multipredictor analysis

Univariate and multivariate analyses were conducted to assess the contribution of practice and population predictor variables as determinants of PHI and PHI% performance scores. A two-level multilevel regression model was used, which allowed the researchers to take into account the likelihood that practice characteristics

Table 2. Quality and Outcomes Framework (QOF) and Public Health Impact (PHI) scores for general practices in England

	QOF points, %	QOF clinical points, %	QOF(20) points, %	QOF(20) true achievement, %	PHI(20) ^a achievement: estimated mortality reduction, per 100 000 patients	PHI(20) ^a : maximum potential mortality reduction, per 100 000 patients	PHI(20)% ^a : achievement, %	PHI(11) ^a achievement: estimated mortality reduction, per 100 000 patients	PHI(11) ^a : maximum potential mortality reduction, per 100 000 patients	PHI(11)% ^a : achievement, %
Mean	94.1	96.3	97.5	79.0	258.9	340.9	75.7	118.4	161.2	72.9
SD	5.2	5.3	4.4	3.7	73.3	91.8	4.3	36.7	46.4	4.7
Percentiles										
1	73.4	72.7	80.3	68.0	75.5	104.7	63.9	31.6	49.4	58.7
25	92.5	95.1	96.8	77.0	213.4	286.0	73.4	94.6	132.4	70.5
50	95.2	98.0	99.3	79.1	262.1	345.3	75.9	120.4	163.8	73.3
75	97.3	99.6	100.0	81.2	305.7	400.1	78.3	142.8	191.7	75.9
99	100.0	100.0	100.0	87.4	429.9	555.1	85.4	203.6	269.1	82.5

SD = standard deviation. ^aPHI(20) refers to values for the PHI score calculated using 20 QOF indicators; PHI(11) refers to values calculated using 11 QOF indicators.

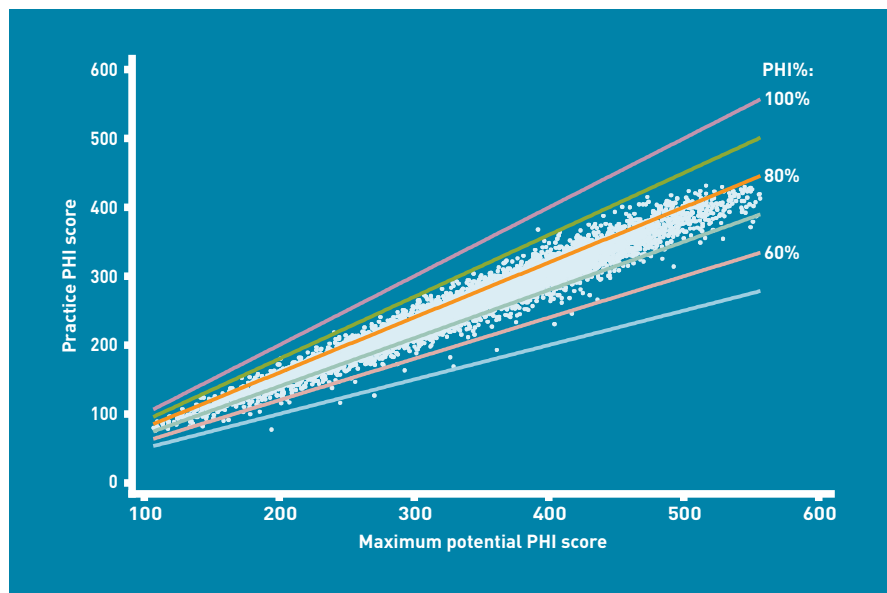


Figure 1. The relationship between practice achievement (Public Health Impact [PHI] score) and maximum potential achievement (PHI%). PHI% = PHI score as a percentage of the maximum possible score. All values represent estimates of mortality reduction, per 100 000 registered patients, per annum.

were clustered at primary care trust (PCT) level, the local managerial level of health service organisation. All multivariate analyses excluded the highest and lowest 1% of the dependent variable, in order to avoid distortion of the regression model by outlier values.

Sensitivity analysis

For a sensitivity analysis, the analysis was repeated on a subset of 11 of the original QOF(20) indicators for which there was at least randomised controlled trial (RCT) with research evidence of mortality reduction

(Table 1).^{11,12} The remaining indicators excluded from the sensitivity analysis were derived from non-RCT studies.

RESULTS

Disease prevalence

Summaries of disease prevalence, as represented by the selected indicators, are presented in Table 1. Composite practice prevalence values varied widely: practices in the lowest percentile had a mean prevalence value of 0.36 or less, and practices in the highest percentile had a mean prevalence of 1.66 or more; the 5th and 95th centile values were 0.49 and 1.47, respectively.

Estimated mortality reduction; the PHI score

The mean estimated reduction in mortality achieved by general practices in England, based on their performance on the QOF(20) indicators in 2009/2010, was 258.9 lives per 100 000 registered patients, per annum, which represented 75.7% of their theoretical maximum potential to reduce mortality (Table 2).

From the perspective of individual GPs, the estimate of mortality reduction was 4.7 lives saved per fte GP, per annum. This value was derived from the national QOF database, which showed an average registered population of 1820 patients in 2009/2010.

Given the variance in PHI% achievement, an estimate of the additional lives that could have been saved through improvements

Table 3. Association between practice characteristics and estimated mortality reduction: PHI scores

Practice characteristic	Unadjusted coefficient, <i>B</i> (95% CI)	Adjusted coefficient, <i>B</i> (95% CI)	Standardised adjusted coefficient, <i>β</i> (<i>P</i> value)
Patients			
Patients aged ≥65 years, %	9.37 [9.19 to 9.55]	2.79 [2.64 to 2.95]	0.22 (<0.001)
South Asian population, %	-0.54 [-0.64 to -0.43]	0.21 [0.17 to 0.25]	0.04 (<0.001)
Black/black British population, %	-1.79 [-2.07 to -1.51]	-0.04 [-0.15 to 0.06]	0.00 [0.417]
Disease prevalence			
Prevalence factor	2.16 [2.14 to 2.18]	1.80 [1.77 to 1.83]	0.77 (<0.001)
Practice			
GP training practice	3.17 [0.48 to 5.85]	-2.27 [-3.44 to -1.10]	-0.01 (<0.001)
Practice list size ^a	-0.03 [-0.06 to -0.00]	-0.08 [-0.10 to -0.07]	-0.05 (<0.001)
List size per fte GP ^a	-0.48 [-0.62 to -0.35]	0.08 [0.03 to 0.14]	0.01 [0.002]
Area			
North versus South	49.46 [35.72 to 63.20]	9.43 [6.48 to 12.38]	0.07 (<0.001)
IMD-2010 score	-0.17 [-0.25 to -0.09]	0.1 [0.07 to 0.14]	0.03 (<0.001)
Unexplained (residual) variance (95% CI)			
	Variance overall (95% CI)	%	%
PCT level	2114 [1675 to 2668]	45	15
Practice level	2586 [2506 to 2667]	55	85

fte = full-time equivalent. IMD = Index of Multiple Deprivation. PCT = primary care trust. ^aMeasured in 100s of patients.

Table 4. Association between practice characteristics and achievement (%) of maximum potential mortality reduction: PHI% performance scores

Practice characteristic	Unadjusted coefficient, <i>B</i> (95% CI)	Adjusted coefficient, <i>B</i> (95% CI)	Standardised adjusted coefficient, β (<i>P</i> value)
Patients			
Patients aged ≥ 65 years, %	0.11 (0.09 to 0.12)	-0.09 (-0.12 to -0.06)	-0.13 (<0.001)
South Asian population, %	-0.01 (-0.02 to -0.01)	0.00 (-0.00 to 0.01)	0.01 (0.508)
Black/black British population, %	-0.09 (-0.11 to -0.07)	-0.04 (-0.06 to -0.02)	-0.07 (<0.001)
Disease prevalence			
Prevalence factor	0.03 (0.03 to 0.03)	0.04 (0.03 to 0.04)	0.29 (<0.001)
Practice			
GP training practice	-0.24 (-0.43 to -0.05)	0.01 (-0.20 to 0.22)	0.00 (0.903)
Practice list size ^a	-0.01 (-0.01 to -0.01)	-0.01 (-0.02 to -0.01)	-0.16 (<0.001)
List size per fte GP ^a	-0.02 (-0.03 to -0.01)	-0.02 (-0.03 to -0.01)	-0.04 (0.001)
Area			
North versus South	0.99 (0.58 to 1.39)	0.41 (0.03 to 0.79)	0.05 (0.034)
IMD-2010 score	-0.03 (-0.03 to -0.02)	-0.03 (-0.04 to -0.03)	-0.15 (<0.001)
	Variance overall (95% CI)	%	Unexplained (residual) variance (95% CI)
PCT level	1.3 (1.0 to 1.7)	9	0.8 (0.6 to 1.1)
Practice level	12.8 (12.4 to 13.2)	91	11.6 (11.2 to 12.0)
			%
			7
			93

fte = full-time equivalent. *IMD* = Index of Multiple Deprivation. *PCT* = primary care trust. ^aMeasured in 100s of patients.

in practice performance was calculated. The target was arbitrarily set at the 75th centile level (upper quartile) of PHI% achievement. Practices in the highest-performing quartile were achieving a minimum of 78.3% of their theoretical maximum mortality-reduction potential. If all lower-performing practices in England were brought up to the minimum level of the upper quartile, this would result in an

additional 5361 lives saved per annum.

The relationship between practices' actual PHI scores, their maximum potential achievement, and their PHI% performance scores is illustrated in Figure 1.

Univariate correlates

The PHI score correlated weakly with the total QOF score (Pearson's $r = 0.28$) and clinical QOF score ($r = 0.25$). The PHI%

Table 5. Association between practice characteristics and estimated mortality reduction: PHI(11) scores

Practice characteristic	Unadjusted coefficient, <i>B</i> (95% CI)	Adjusted coefficient, <i>B</i> (95% CI)	Standardised adjusted coefficient, β (<i>P</i> value)
Patients			
Patients aged ≥ 65 years, %	5.10 (5.02 to 5.18)	1.89 (1.83 to 1.95)	0.29 (<0.001)
South Asian population, %	-0.42 (-0.47 to -0.37)	0.00 (-0.01 to 0.02)	0.00 (0.728)
Black/black British population, %	-1.14 (-1.28 to -1.00)	-0.09 (-0.13 to -0.04)	-0.02 (<0.001)
Disease prevalence			
Prevalence factor	1.08 (1.07 to 1.09)	0.82 (0.81 to 0.83)	0.70 (<0.001)
Practice			
GP training practice	3.00 (1.67 to 4.33)	-0.80 (-1.27 to -0.32)	-0.01 (0.001)
Practice list size ^a	0.01 (-0.00 to 0.02)	-0.02 (-0.02 to -0.01)	-0.02 (<0.001)
List size per fte GP ^a	-0.29 (-0.35 to -0.22)	0.03 (0.01 to 0.05)	0.01 (0.014)
Area			
North versus South	23.29 (16.01 to 30.57)	5.06 (3.63 to 6.49)	0.07 (<0.001)
IMD-2010 score	-0.25 (-0.29 to -0.21)	-0.07 (-0.08 to -0.05)	-0.03 (<0.001)
	Variance overall (95% CI)	%	Unexplained (residual) variance (95% CI)
PCT level	563.5 (446.8 to 710.7)	47	15.6 (12.2 to 20.1)
Practice level	637.9 (618.3 to 658.2)	53	59.3 (57.4 to 61.2)
			%
			21
			79

fte = full-time equivalent. *IMD* = Index of Multiple Deprivation. *PCT* = primary care trust. ^aMeasured in 100s of patients.

Table 6. Association between practice characteristics and achievement (%) of maximum potential mortality reduction: PHI(11)% performance scores

Practice characteristic	Unadjusted coefficient, <i>B</i> (95% CI)	Adjusted coefficient, <i>B</i> (95% CI)	Standardised adjusted coefficient, β (<i>P</i> value)
Patients			
Patients aged 65 years or over, %	0.21 (0.19 to 0.23)	-0.07 (-0.10 to -0.04)	-0.09 (<0.001)
South Asian population, %	-0.07 (-0.07 to -0.06)	-0.04 (-0.05 to -0.04)	-0.15 (<0.001)
Black/black British population, %	-0.15 (-0.17 to -0.13)	-0.07 (-0.08 to -0.05)	-0.11 (<0.001)
Disease prevalence			
Prevalence factor	0.05 (0.05 to 0.05)	0.05 (0.05 to 0.06)	0.37 (<0.001)
Practice			
GP training practice	0.22 (0.03 to 0.42)	-0.01 (-0.21 to 0.20)	0 (0.935)
Practice list size ^a	0.00 (-0.00 to -0.00)	0 (-0.01 to -0.00)	-0.05 (<0.001)
List size per fte GP ^a	-0.03 (-0.04 to -0.02)	-0.01 (-0.02 to -0.00)	-0.03 (0.009)
Area			
North versus South	1.39 (0.77 to 2.00)	0.38 (-0.06 to 0.82)	0.04 (0.093)
IMD-2010 score	-0.04 (-0.05 to -0.04)	-0.03 (-0.04 to -0.03)	-0.14 (<0.001)
	Variance overall (95% CI)	%	Unexplained (residual) variance (95% CI)
PCT level	3.4 (2.6 to 4.3)	20	1.3 (1.0 to 1.7)
Practice level	13.4 (13.0 to 13.8)	80	11.4 (11.0 to 11.7)
			%
			10
			90

fte = full-time equivalent. *IMD* = Index of Multiple Deprivation. *PCT* = primary care trust. ^aMeasured in 100s of patients.

performance score correlated moderately with the total QOF (Pearson's $r = 0.54$) and clinical QOF ($r = 0.51$) scores. All correlations were significant, $P < 0.001$.

Multipredictor analysis

The most powerful determinants of the PHI score were the composite prevalence of the QOF(20) conditions and the proportion of patients aged ≥ 65 years. The full model summarised in Table 3 explained 55% of the variance in PHI score at practice level, and 45% at PCT level.

A similar multivariate analysis of the determinants of the PHI% achievement score was conducted. Practices that maximised their potential for mortality reduction were those with higher prevalence of QOF(20) conditions, situated in less deprived areas, with smaller list sizes, and with fewer patients aged ≥ 65 years. The study model explained 91% of the variance in PHI% at practice level and 9% at PCT level (Table 4).

Sensitivity analysis

PHI(11) (based on the QOF(11) indicators) scores for achievement, maximum potential achievement, and percentage of maximum potential achievement correlated strongly with their PHI(20) counterparts (based on the QOF(20) indicators): $r = 0.97$, 0.97 , and 0.78 , respectively (all significant, $P < 0.001$).

PHI(11) values, and the determinants of PHI(11) and PHI(11)% achievement are displayed in Tables 2, 5, and 6.

DISCUSSION

Summary

This study has produced a measure of the public health impact of individual general practices. It estimates that achievement levels for a subset of 20 clinical QOF indicators in 2009/2010 translates into a mean mortality reduction of 258.9 lives per 100 000 registered patients, per annum. For a GP in England with the national average list size of 1820 patients, this equates to 4.7 lives saved per year. Nationally, this equates to 139 100 lives saved. This value represents 75.7% of the theoretical maximum potential for mortality reduction; just 1% of practices exceeded 85% of their potential maximum score. Increasing PHI% performance to the level of the top quartile of practices would save an additional estimated 5361 lives each year in England.

Overall mean achievement rates of the 20 clinical indicators were broadly similar, regardless of whether they were unweighted or weighted according to mortality reduction (79.0% and 75.7%, respectively). In contrast, QOF clinical point scores for these practices were high (mean = 96.3%), confirming that most practices have achievement rates above the QOF upper payment thresholds for most indicators, but falling well short of 100% achievement. Moreover, the correlations between QOF points scores and PHI scores were not strong, implying that for many practices the financial rewards of the QOF may not be closely aligned with the PHI of

Box 1. Worked examples of general practices and their PHI scores

Practice 1: PHI score: 260 per 100 000 registered patients, per annum; PHI% performance score: 85%

This practice has a PHI score that is just about average for the whole country. However, it has maximised its potential to reduce mortality and is in the top 1% of performers in the country, in terms of potential achievement.

This mismatch between average PHI score and high PHI% performance score occurs because Practice 1 has low disease prevalence compared to the average practice, thus imposing a ceiling on the maximum potential number of lives saved.

Practice 2: PHI score: 400 per 100 000 registered patients, per annum; PHI% performance score: 70%

This practice appears to be doing very well in terms of PHI score, achieving above the 90th centile. However, it has achieved below national average in terms of PHI% performance.

This mismatch between high PHI score and low PHI% performance score occurs because Practice 2 has high disease prevalence compared to average. Although disease-prevention activity has already saved a large estimated number of lives, this practice could save many more lives based on the high disease prevalence in the practice. This pattern might occur in an inner-city area or in an area with a large older population in which the practice is underachieving in terms of disease-prevention activity.

Practice 3: PHI score: 200 per 100 000 registered patients, per annum; PHI% performance score: 85%

This practice has a low PHI score and yet it has achieved highly in terms of potential. This might be typical of a practice serving a student population where morbidity is low; hence, however hard the practice tried and however successful it was at disease-prevention activity, it could never achieve the mortality reduction of practices in areas of higher disease prevalence.

the practice's activities.

Calculating practice prevalence for relevant conditions was an integral part of constructing PHI scores. Unlike clinical achievement scores (which had a narrow variance) it was found that prevalence varied considerably. The study plans further work on the factors that predict these large variations in the prevalence, and the reasons why some practices appear to carry a far higher burden of morbidity than others serving similar populations.

Practices that underachieved in terms of fulfilling their potential to save lives were those with lower burdens of morbidity (as defined by the prevalence factor), suggesting that successful implementation of disease-prevention activity is linked to higher prevalence. When findings were adjusted for prevalence, practices located in deprived areas and with an older population were less likely to maximise their potential to reduce mortality. These are the practices where there is further potential to reduce mortality and where interventions to improve public health outcomes might be most successful. Smaller practices were better at maximising their potential to save lives. Several studies have demonstrated the specific features of quality of care linked with small practices, particularly in continuity of care, and it is possible that some of these attributes are linked with effectiveness at disease prevention.¹⁵ Examples of practices with high and low

PHI scores with varying levels of PHI% achievement are given in Box 1.

Strengths and limitations

This study has produced a new metric for general practices in England, based on estimates constructed using the most conservative interpretation of trial data. Sensitivity analysis conducted using the subset of 11 clinical indicators with the strongest evidence base suggests that the PHI scores are robust. The final estimate for mean annual mortality reduction (258.9 per 100 000 registered patients) compares with the overall 2009 national (England and Wales) mortality rate of 896 per 100 000 registered patients.¹⁶ This reflects the contribution of the conditions included in the PHI score to overall mortality, particularly coronary heart disease and chronic obstructive pulmonary disease, and the importance of activities such as influenza vaccination.

The findings of the study are constrained by several limitations. There are other clinical interventions that have the potential to save lives, and primary-care-based public health interventions span a broader range of activity than the current list of indicators contained within the QOF. However, the subset of 20 indicators used in this study represents those interventions with the strongest evidence base for mortality reduction, even though RCT-level evidence of mortality reduction could only be found for 11 of these indicators. Ideally, indicators should be selected from a wider pool of indicators, all ranked according to mortality-reduction estimates derived from high-quality meta-analyses of RCT data. This study relied on a 2008 analysis of mortality-reduction estimates, although both the sensitivity analysis and the strong correlation between PHI% achievement scores and QOF(20) achievement scores weighted for prevalence suggest that the mortality weightings do not greatly alter the final scores. The approach in this study though, was not dependent on one dataset and could readily be updated as new estimates become available. By using proxy measures, it is also likely that the role of ethnicity and social deprivation in determining health inequalities has been underestimated.¹⁰

The study analysis was confined to mortality reduction. Arguably, the impact of primary care is likely to be greater in terms of improving quality-adjusted life years or reducing disability-adjusted life years, but this study, like others, found few data to conduct such an analysis.¹⁷

The PHI score was heavily influenced by mortality reduction for influenza immunisation indicators. Doubts have been raised about whether findings of vaccine effectiveness in younger age groups can be generalised to an older population, and few trials have included patients aged ≥ 70 years, the age group that accounts for three-quarters of influenza-related deaths.¹⁸ The overall estimate of mortality reduction may therefore have to be revised downwards. However, all four influenza vaccination indicators were excluded from the sensitivity analysis in this study, and convergence with the main analysis suggests that these indicators exerted only a modest effect on the regression models.

The overall estimate of mortality reduction was corrected for comorbidity. It could be argued that the analysis should not be corrected for comorbidity, since each of the 20 primary care interventions are effective, regardless of how many other illnesses are experienced by each patient. The present authors preferred the more cautious analysis, since the same intervention (for example influenza vaccination) would be unlikely to have an additive effect if given to a patient with several comorbidities for which influenza vaccination was recommended. Similarly, multiple different preventative interventions on the same patient would be unlikely to have an additive effect on mortality reduction. The lack of a strong evidence base on mortality reduction in multimorbidity encouraged use of the more cautious analysis.

Finally, it is important to place the present findings within the context of overall primary care activity. The PHI score is a measure of the impact of QOF-related activity within primary care, rather than necessarily demonstrating the impact of the QOF itself.

Comparison with existing literature

Linkage between the selection of performance indicators for primary care and health gain was first suggested in 1992.¹⁹ More recently, it has become possible to quantify predicted benefits. The study identified two studies that modelled financial incentives and health outcomes in primary care, both of which produced findings of a similar magnitude to those derived in the present study.^{11,20} One study explored the likely health outcomes of achieving five clinical targets,²⁰ and a further study derived weightings for eight health-promotion initiatives, although these were applied at a higher managerial level ('primary care group') rather than

at practice level.²¹ In the US, the current physician payment model provides incentives for increased volume of activity, but lack of computerisation has hindered efforts to evaluate the public health role of primary care.²² The present study is the first to report predicted health outcomes linked to primary care-level achievement of such a broad range of indicators, and the first to develop a score based on this measure.

Implications for practice and research

At a practice level, the study considers that the PHI scores will allow GPs and health service commissioners to focus on maximising the effectiveness of general practices in reducing all-cause mortality in their registered population and locality.

Under the QOF, general practices are financially incentivised to reach, but not exceed, achievement thresholds of less than 100% for clinical targets. These thresholds, which are arbitrarily set, are relatively low for many indicators, and the findings of this study suggest that public health effectiveness may be curtailed as a result. One solution might be to introduce higher targets with stepped increases in the incentive, so that the incentive reflects the difficulty of reaching higher levels of achievement.

This study has highlighted the lack of a readily available up-to-date evidence base on which to base calculations of public health effectiveness. Current NHS reforms²³ are likely to require newly formed commissioning groups to make decisions on resource allocation for initiatives to reduce mortality, and these decisions require access to health economic modelling of the cost-benefits of proposed investments or disinvestments.

This study takes a step towards filling the information gap between primary care processes and outcomes, by translating process quality performance into estimated health outcome performance in terms of mortality reduction. The resulting suite of practice-level measures of public health impact provides an alternative, more outcomes-oriented approach, which can be used to gauge the contribution of primary care to public health.

The PHI score is a potential new metric of practice performance. Although derived from 20 clinical QOF indicators, achievement of PHI scores did not correlate strongly with total or clinical QOF scores, implying that the QOF pay-for-performance scheme offers financial rewards that may not be well aligned to the task of reducing mortality from preventable disease.

Funding

The research was supported by the National Institute for Health Research (NIHR) Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust and King's College London. Tim Doran was supported by the National Institute for Health Research (NIHR-CDF-2011-04-016). The views expressed in this publication are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Ethical approval

Guy's Research Ethics Committee (Chairman's action, 8/2/06).

Provenance

Freely submitted; externally peer reviewed.

Competing interests

The authors have declared no competing interests.

Discuss this article

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

REFERENCES

- Starfield B, Shi L, Macinko J. Contribution of primary care to health systems and health. *Milbank Q* 2005; **83**(3): 457–502.
- Pearson SD, Schneider EC, Kleinman KP, *et al*. The impact of pay-for-performance on health care quality in Massachusetts, 2001–2003. *Health Aff* 2008; **27**(14): 1167–1176.
- Roland M. Linking physicians' pay to the quality of care — a major experiment in the United Kingdom. *N Engl J Med* 2004; **351**(14): 1448–1454.
- Kindig DA. A pay-for-population health performance system. *J Am Med Assoc* 2006; **296**(21): 2611–2613.
- Department of Health. *Transparency in outcomes — a framework for the NHS*. London: DoH, 2010.
- Department of Health. *Healthy lives, healthy people: our strategy for public health in England*. London: DoH, 2010.
- HM Government. *Quality and Outcomes Framework — 2009–10, England level*. <http://data.gov.uk/dataset/quality-and-outcomes-framework-achievement-data> [accessed 4 Mar 2013].
- Department for Communities and Local Government. *English indices of deprivation 2010*. London: Crown Copyright, 2011. <http://www.communities.gov.uk/publications/corporate/statistics/indices2010> [accessed 4 Mar 2013].
- Office for National Statistics. A beginners' guide to geography. <http://www.ons.gov.uk/ons/guide-method/geography/beginner-s-guide/index.html> [accessed 4 Mar 2013].
- 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* 2006; **28**(1): 39–42.
- Fleetcroft R, Parekh-Bhurke S, Howe A, *et al*. The UK pay-for-performance programme in primary care: estimation of population mortality reduction. *Br J Gen Pract* 2010; DOI: 10.3399/bjgp10X515359.
- Fleetcroft R, Parekh S, Steel N, *et al*. Joint executive summary: potential population health gain of the quality and Outcomes Framework. Report to the Department of Health. York: Centre for Health Economics, 2008. <http://www.york.ac.uk/che/pdf/jointexecutivesummaryUEA-York-%20270308final.pdf> [accessed 4 Mar 2013].
- Raffle AE, Alden B, Quinn M, *et al*. Outcomes of screening to prevent cancer: analysis of cumulative incidence of cervical abnormality and modelling of cases and deaths prevented. *BMJ* 2003; **326**(7395): 901–904.
- British Medical Association. *NHS Employers, Quality and Outcomes Framework guidance for GMS contract 2019/10*. London: NHS Confederation (Employers) Company Ltd, 2009. http://www.nhsemployers.org/Aboutus/Publications/Documents/QOF_Guidance_2009_final.pdf [accessed 4 Mar 2013].
- Freeman G, Hughes J. *Continuity of care and the patient experience*. London: The Kings Fund, 2010.
- Office for National Statistics. *Death registrations summary table, England and Wales, 2009, (final)*. <http://www.ons.gov.uk/ons/rel/vsob1/death-reg-sum-tables/2009--final-/index.html> [accessed 4 Mar 2013].
- Walker S, Mason AR, Claxton K, *et al*. Value for money and the Quality and Outcomes Framework in primary care in the UK NHS. *Br J Gen Pract* 2010; DOI: 10.3399/bjgp10X501859.
- Simonsen L, Taylor RJ, Viboud C, *et al*. Mortality benefits of influenza vaccination in elderly people: an ongoing controversy. *Lancet Infect Dis* 2007; **7**(10): 658–666.
- Siu AL, McGlynn EA, Morgenstern H, *et al*. Choosing quality of care measures based on the expected impact of improved care on health. *Health Serv Res* 1992; **27**(5): 619–650.
- 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**(3): 191–197.
- McColl A, Roderick P, Gabbay J, *et al*. Performance indicators for primary care groups: an evidence approach. *BMJ* 1998; **317**(7169): 1354–1360.
- Public Policy Committee of the American College of Physicians; Ginsburg JA, Doherty RB, Ralston JF Jr, *et al*. Achieving a high-performance health care system with universal access: what the United States can learn from other countries. *Ann Intern Med* 2008; **148**(1): 55–75.
- Department of Health. *Equity and excellence: liberating the NHS*. London: DoH, 2010.

Appendix 1. Formula used for the calculation of mortality reduction estimates, the 'Public Health Impact' (PHI) score

$$\sum_{i=1}^{n=20} \text{ARR}(i) * \frac{\text{disease prevalence}}{100\,000} (ip) * \text{true achievement} (ip) * \text{comorbidity correction factor} (p)$$

where *i* represents each indicator, and *p* represents each practice.

Note: 'disease prevalence' refers to the prevalence value for each of the selected indicators.