PAPER PRESENTED TO HESG, BRUNEL 2007

What quality improvement did the Quality and Outcomes Framework produce?*

Matt Sutton[+]*, Health Economics Research Unit, University of Aberdeen*
Ross Elder, Information Services Division, NHS National Services Scotland
Bruce Guthrie, Tayside Centre for General Practice, University of Dundee
Graham Watt, General Practice and Primary Care, University of Glasgow

Abstract

The Quality and Outcomes Framework (QOF) is an expensive, innovative pay for performance scheme for general practices introduced across the UK in 2004. Average performance in the first year was very high but, since comparable data were not routinely collected prior to introduction, little is known about the extent to which this is attributable to the new scheme. We aim to identify the impact of the QOF by comparing trends in performance on incentivised and non-incentivised indicators. We analyse annual rates of recording of blood pressure, smoking status, cholesterol, body mass index (BMI) and alcohol status based on individual patient records from 315 general practices over the period 2001/2 to 2005/6. The recording of each risk factor is designated as incentivised or unincentivised for each individual based on their diagnostic history. Rates of recording increased for all risk factors for all diagnostic groups. Nevertheless, rates of recording increased most rapidly where it was explicitly incentivised in the QOF. We have estimated the overall increase in rates of recording to be 10 percentage points. The estimated impacts of the QOF were larger when comparisons were made between incentivised and unincentivised indicators for individuals with incentivised conditions. The impacts were largest when recording of BMI was used as the comparator. Responses were greater on indicators that attracted more payment and required more stringent performance. Practices that were performing less well prior to the QOF showed the greatest response.

[+] Corresponding author: m.sutton@abdn.ac.uk

* This work was funded by NHS Quality Improvement Scotland as part of a larger project to examine clinical effectiveness using SPICE data (Elder et al, 2007). We are grateful to other members of the project team (Bill Ramsay, Michelle Fitzpatrick, Colin Fischbacher, Colin Simpson, Margaret MacLeod) for comments on previous versions of this work. HERU receives funding from the Scottish Executive Chief Scientist Office. The authors are responsible for the views expressed.

Introduction

The nGMS contract introduced in April 2004 included a new funding stream for quality and outcomes, rewarding achievements on a wide range of quality indicators for ten chronic diseases (NHS Confederation and British Medical Association, 2003). Introduction of this Quality and Outcomes Framework (QOF) represents a major initiative for improving quality in primary medical care (Roland, 2004).

The QOF is an expensive, innovative pay for performance scheme. It introduced a new data system for measuring quality and was introduced universally and simultaneously across the UK. Average performance in the first year was very high but, since comparable data were not routinely collected prior to introduction, little is known about the extent to which this is attributable to the new scheme. This is key to understanding whether QOF payments represent retrospective compensation for work already being done (Buckman, 2007) or an incentive scheme to produce quality improvement.

Gulliford et al (2007) extracted clinical data from the records of 2,099 patients at 26 general practices in South London. They concentrated on achievement of incentivised activities for patients with diabetes. Annual recording of risk factors increased between 2000 and 2003: for BMI from 66 to 82%; for blood pressure from 75 to 89%; and for cholesterol from 50 to 77%. Coverage on the respective QOF indicators had increased to 94%, 98% and 93% by 2005.

Tahrani et al (2007) also concentrate on incentivised aspects of diabetes care. They analyse data for 16,867 diabetes patients across 66 general practices in Shropshire between April 2004 and March 2006. The percentage of diabetics with records increased between April 2004 and March 2006: for BMI from 73 to 89%; for smoking from 44 to 95%; for blood pressure from 87 to 97%; and for cholesterol from 78 to 93%..

Steel et al (2007) examined the records of 1,156 patients in 18 general practices in Norfolk in 2003 and 2005. For two incentivised conditions (hypertension and asthma) they collected 21 indicators, 6 of which were incentivised in the QOF. They also collected 15 indicators for two unincentivised conditions (osteoarthritis and depression). There were significant increases for

the six incentivised indicators and the 15 unincentivised indicators for the incentivised conditions. The indicators for the unincentivised conditions did not increase significantly.

Campbell et al (2007) examined the quality of care in 42 practices in 1998, 2003 and 2005 for three incentivised conditions (CHD, asthma and diabetes). They compared trends in 30 incentivised indicators with 17 unincentivised indicators using information from the records of 2,300 patients in 1998, 1,495 patients in 2003 and 1,482 patients in 2005. Mean achievement percentages increased from 59 to 85 for CHD, 62 to 81 for diabetes and 60 to 84 for asthma. These were compared with predicted scores based on extrapolation of a fitted trend between 1998 and 2003. The 2003 to 2005 increases were significantly higher than the fitted trend for asthma and diabetes but only marginally so for CHD ($p=0.07$). The rate of improvement for the incentivised indicators did not differ significantly from the rate of improvement for the unincentivised indicators.

These previous studies have been based on small samples of practices. Only one study has considered unincentivised conditions. None have clearly exploited the presence of condition and risk factor combinations that are or are not incentivised and they do not appear to have allowed for comorbidity. As we show, this is important for assigning activities applied to individual patients to incentivised and unincentivised groups.

The Scottish Programme for Improving Clinical Effectiveness in Primary Care (SPICE-PC) is a quality initiative under the auspices of the Royal College of General Practitioners. A series of Care Management Screens was developed to offer guidance on good practice. Practice participation is voluntary and involves unpublished feedback on achievements relative to the national average. The nGMS contract was introduced during the period in which SPICE has operated, offering the opportunity for before and after comparison of performance. Since SPICE collects information on indicators not measured and rewarded in the QOF, it also offers opportunity to assess whether quality has improved faster on the incentivised activities.

In this paper, we analyse trends in quality over the period 2000/1 to 2005/6. We examine the recording of five risk factors for six patient groups, including a group of patients not diagnosed with five of the ten QOF conditions. We use a hierarchy of diagnoses to ensure that recording of some of the risk factors is incentivised for some patient groups and not incentivised for others. Unique among recent studies, we consider the entire population.

Among the risk factors we include body mass index, which is attracting growing attention as evidenced by inclusion of an obesity register in the post-2006 revision of the QOF, and alcohol status, which is emerging as a major cause of death (Leyland et al, 2007). In addition, we do not adopt the QOF definitions of achievement to allow us to examine aspects of the design properties of the QOF. We find that the existence and the design of incentives impact on performance.

Data

The SPICE data were provided by the Primary Care Clinical Informatics Unit, Department of General Practice, Aberdeen University. A full extract of the data collated during the Spring quarter 2006 containing data from 315 general practices was imported into the NHSScotland Information Services Division.

We used information on registration and de-registration dates, patients' age and sex, encounter dates, clinical events (including Read codes), measurements, care values, and SPICE Care Management Screen values. These latter two sets of data were introduced as the computer system (GPASS) evolved and we use all the data available from each source.

We considered five chronic conditions included in the QOF: Coronary Heart Disease (CHD); Chronic Obstructive Pulmonary Disease (COPD); Diabetes; Hypertension; and Stroke/TIA. Individuals were considered diagnosed if they had received a relevant Read code prior to the start of the financial year. Details of the Read codes used are in the full report (Elder et al, 2007). The validity of diagnoses was assessed by comparisons with QOF disease register sizes in 2005/6.

Participation in SPICE is voluntary and information may be only partially recorded. To be suitable for inclusion, the indicators should have reasonable coverage prior to the introduction of the QOF. Rates of recording of five risk factors were selected as they were consistently collected and were of relevance across a range of conditions. The five risk factors were: smoking status; alcohol status; blood pressure; cholesterol; and Body Mass Index. For blood pressure, only measurements generating values within plausible ranges were included - systolic blood pressure between 80 and 250 mmHg and diastolic blood pressure between 40

and 130 mmHg. For cholesterol, the range included was between 2 and 27. Information about alcohol consumption and smoking status is recorded using Read codes. The list of codes used is in the full report (Elder et al, 2007).

Individuals' records were included in the analysis if they were registered with the practice prior to the start of the financial year and remained registered until the end of the financial year. Individuals were counted as having a risk factor recorded if such information was collected during the financial year. The percentages therefore indicate the percentage of patients having the risk factor recorded during the financial year. The analysis only used permanent residents, who accounted for approximately 95% of the patients.

The age of each patient was calculated as at the 31st of March within each individual year. The analysis is restricted to patients aged 45 years and over.

Representativeness

A range of practice characteristics for the 315 practices in the sample were compared to the 721 other Scottish practices (Elder et al, 2007). On average, 16% of the registered population for SPICE practices is aged 65 years and over, with the respective figure for non-SPICE practices similar at 15%. Only a quarter of the practices in the most deprived quintile are SPICE practices compared to one-third in less deprived areas. The SPICE practices also show some non-random spatial concentration across Health Boards.

The average number of partners in a SPICE practice is 4.8, with 48% of these being female. The respective non-SPICE figures are slightly lower at 4.2 and 45%. SPICE practices, in particular, are less likely to be single-handed (7.9% versus 13% for non-SPICE practices).

Participation in voluntary activities, such as Practice Team Information (PTI), Personal Medical Services (PMS) and training, is more prevalent in SPICE practices. Five times as many SPICE practices participate in PTI (12.4% versus 2.5%), and almost double the proportion of SPICE practices hold PMS contracts (14% compared with 7.4%). A quarter of SPICE practices are training practices compared with approximately one-fifth of non-SPICE practices. Finally, SPICE and non-SPICE nGMS contract practices were compared using the Quality and Outcomes Framework data for 2005/6. On average, SPICE practices performed better in each of the domains with an extra 12.4 points overall.

These results suggest caution in extrapolating the results to all Scottish practices.

Existence and design of QOF incentives

Table 1 provides information on whether the recording of each of the five risk factors is included in the QOF. Where included in the QOF, the maximum points available for recording (and controlling) each risk factor is shown. In addition, practices can earn points for "records and information about patients", which includes up to 11 points for recording the smoking status of all patients aged 15–75 years and up to 15 points for recording in the preceding five years the blood pressure of all patients aged 45 years and over.

Quality points are awarded according to a linear function between a lower and an upper threshold. The lower threshold is 25% for all indicators. The upper threshold is 90% for most indicators with the exception of the two indicators for all patients, which reach a maximum at 75%. For most indicators, achievement for an individual patient requires that the risk factor is recorded within the last fifteen months. Blood pressure recording for hypertension patients is required more frequently - every nine months. The recording of smoking status for hypertension patients, and blood pressure and smoking status for all patients, is required less frequently – within the last five years.

In the SPICE Care Management Screens, recording of information is recommended for more risk factors than are rewarded on the QOF for several conditions. Table 1 indicates instances where risk factor recording is included in the SPICE criteria but not rewarded on the QOF. For diabetes and CHD, collection of information on all risk factors is recommended. For hypertension, all risk factors except cholesterol are recommended for collection. The COPD Screen did not exist before the QOF was introduced.

Mutually exclusive groups

Individuals can appear in more than one diagnostic group. Comparison of incentivised and unincentivised indicators requires mutually exclusive groups. We therefore created a hierarchy of diagnoses based on the number of risk factors incentivised in the QOF. Each individual patient appears in just one of the following six groups:

- *Diabetes*

- *CHD* (excluding those diagnosed with diabetes)
- *Stroke* (excluding diabetes and/or CHD)
- *Hypertension* (excluding diabetes, CHD and/or stroke)
- *COPD* (excluding diabetes, CHD, stroke and/or hypertension)
- *Other* (excluding diabetes, CHD, stroke, hypertension and/or COPD)

Figure 1 shows the numbers of individuals included in at least one year of the analysis for each diagnostic group. Also shown are the numbers excluded from each diagnostic group because they appear in a diagnostic group higher in the hierarchy. The largest group is for hypertension. Approximately 36% (46,000 of 127,000) of this group are assigned to another group because they have also been diagnosed with diabetes, CHD and/or stroke. COPD is the smallest group and a substantial proportion of these individuals is assigned to another group.

Methods

The unit of analysis is the recording or not of each risk factor for each patient in each year. Within a year therefore, there are five observations for each patient. The dataset is unbalanced because patients are included only if they have been registered throughout the year. Patients that are registered with a practice throughout all five years appear 25 (5 risk factors x 5 observation years) times. Alongside each risk factor record, we have an indicator of the condition group, the practice, the year and the patient's age and sex.

To this dataset, we match on indicators of whether the indicator is incentivised by the QOF at that time, the number of points available for this indicator, the upper threshold for incentivised indicators, and the width of the interval in which this risk factor requires recording to indicate achievement in the QOF. We also calculate the practice's prevalence rate for this condition, the practice's total list size, and the practice's pre-QOF (2002) rate of achievement on this indicator for this condition group.

The specification of the basic model is:

$$y^{*}_{ijkt} = x'\alpha + \beta_j + \gamma_k + \delta_t + \theta q_{jkt} + \varepsilon_{ijkt} \qquad (1)$$

in which y^*_{ijkt} is the latent quality index for individual *i* on risk factor *j* in condition group *k* at time *t*. *x* is a vector of patient-specific variables, which includes interactions between sex and a cubic function of age. The parameters β, γ and δ represent fixed effects for the risk factors, condition groups and years respectively. The binary variable q_{jkt} indicates whether the recording of risk factor *j* is incentivised by the QOF for condition *k* at time *t*. The estimated parameter θ measures the incremental effect of the QOF on risk factor recording.

The latent index y^*_{ijkt} is not observed. We assume a probit link function to the binary observed indicator of whether the risk factor has been recorded. We present marginal effects from this model. The standard errors are adjusted for clustering by practice.

Following estimation of the basic model, we consider three extensions to equation (1). Our first extension expands the specification of q_{jkt}. We add in:

- the number of QOF points available for maximum achievement on this indicator for this condition group,
- the upper threshold at which maximum achievement is awarded for this indicator for this condition group, and
- the width of the interval (in months) in which the QOF requires recording of this risk factor to be considered 'achieved'.

Each of these variables is included as a main effect and an interaction with q_{jkt}. The coefficient on the interaction term captures differential growth in recording following introduction of the QOF.

The second extension involves the addition of practice-specific variables. We add to equation (1) main effects and interaction terms with q_{jkt} for the following two variables:

- the prevalence rate of condition *j* in the practice at time *t*, and
- the square root of the list size of the practice at time *t*.

We use the square root of the list size to capture any effects caused by the disjuncture between work and reward caused by the Adjusted Disease Prevalence Factor (Guthrie et al, 2006).

Third, we consider whether quality improvement is related to past performance. We measure the proportion of patients in this condition group for which the risk factor was recorded in this

practice in 2002 (t=2). We include this variable, plus an interaction with q_{jkt}, and restrict the period to t>2. The sign of the coefficient on the interaction term indicates whether higher performing practices pre-QOF increased their performance by more or less than lower performing practices when the QOF is introduced.

Comparisons can be made across indicators for the same (disease group of) patients and for the same indicator across different groups of patients. There are many comparisons that can be made. The results of these models are presented (a) for the entire population and set of indicators, (b) the subset of indicators recommended for collection in the SPICE Care Management Screens, and then (c) a suite of specific comparisons.

Smoking status recording is incentivised for all of the five conditions we have considered. For one of these (COPD), however, there was no Care Management Screen within SPICE. The collection of data on smoking status for COPD patients might therefore be considered a completely 'new' indicator. Comparison of trends with the four other conditions offers insight into whether performance improvement is possible with payment alone.

Results

The dataset is very large. To ease computation, we include only a 50% random sample of individuals not diagnosed with any of the five QOF conditions. Summary statistics for this dataset of 5,548,264 observations are provided in Table 2.

Risk factors are recorded on one third of occasions. The mean age of observations is 63 years and 47% of the sample is male. The proportions of sample observations associated with the QOF conditions are: 13% for CHD, 3% for COPD, 10% for diabetes, 24% for hypertension and 4% for stroke. The sample increases through the period - the final year (2005/6) accounts for 23% of observations. In the dataset 47% of observations are for combinations of risk factors and conditions that are, or become, incentivised in the QOF. The average observation is drawn from a practice with a list size of 5,100 patients, varying between 30 and 11,000.

The regression results for all conditions and risk factors are given in Table 3. Risk factor recording varies by sex and age, and increases over the period 2001 to 2005. Risk factor

recording is higher for all the QOF conditions than the rest of the population. The rates are highest for diabetes and CHD. Rates of recording are higher than the reference risk factor (alcohol status) for BMI and blood pressure and lower for cholesterol and smoking.

Prior to the introduction of the QOF, recording was significantly higher for condition and risk factor combinations that become incentivised. Recording rates for these combinations increase further by 9.8 percentage points after the QOF is introduced (Model 1). Model 2 examines the significance of the design features of the QOF. Before the QOF is introduced, rates of recording are significantly higher for condition and risk factor combinations that will attract more points in the QOF, require higher rates to reach the upper threshold and where the QOF measurement interval is wider. After the QOF is introduced, recording rates increase significantly more for those combinations that require higher achievement to reach the upper threshold and significantly less for those where a wider measurement interval is allowed for qualifying for QOF success.

Model 3 considers the introduction of practice variables. Prior to the introduction of the QOF, recording rates were lower where prevalence and list size are higher but not significantly so. Higher prevalence rates are associated with smaller increases in recording once the QOF is introduced. Model 4 shows that a higher level of pre-QOF achievement is associated with a smaller increase in recording after the QOF is introduced.

Table 4 presents equivalent models for combinations of conditions and risk factors that were included in the SPICE Care Management Screens. The results are broadly the same. Introduction of the QOF is associated with a 16.5 percentage point increase in recording on incentivised activities. The reaction to the introduction of the QOF is greater where more points are available and where the required standards (upper thresholds) were more demanding. The width of the QOF measurement interval is negative as before. Unlike in Table 3 the response to the introduction of the QOF does not depend on the prevalence rate but, as in Table 3, the increase in achievement is negatively related to pre-QOF achievement.

Table 5 presents the first set of results for specific comparisons across risk factors. The first model examines differences between Cholesterol and BMI recording for CHD patients. The rate of BMI recording is significantly higher at baseline but the increase following the introduction of the QOF is 26 percentage points higher for the incentivised risk factor

(cholesterol). This can be seen clearly in Figure 2. The second model shows that for CHD and hypertension patients the increase in recording of blood pressure recording is 22 percentage points higher than BMI following the introduction of the QOF. The third model compares the recording of smoking status with alcohol status. Smoking status recording increases by 14.9 percentage points more than alcohol status recording when the QOF is introduced (see Figures 3 and 4).

In Table 6 we compare the recording of particular risk factors across condition groups. BMI recording is incentivised for diabetes but not for CHD and hypertension. When the QOF is introduced, BMI recording for diabetes increases by 7.7 percentage points more than for the other two conditions. Table 6 also compares changes in the recording of smoking status for COPD and other conditions. COPD was not included as a SPICE Care Management Screen. The recording of smoking status for COPD patients increases by 10.1 percentage points more than the other conditions for which it is also incentivised but for which there were Care Management Screens that recommended the recording of smoking status for these patients. This can also be seen in the raw trends in Figure 4.

Discussion

The availability of SPICE data from before the introduction of the QOF provides a rare opportunity to assess the contribution of the QOF to recent improvements in the quality of patient care in general practice. We focused on the recording of five risk factors across six mutually exclusive groups of patients. Five of these six patient groups were defined by QOF conditions and a hierarchy of diagnoses was defined to ensure that the recording of some risk factors was not incentivised in the QOF for some groups. The recording of one risk factor (alcohol status) was not incentivised for any of the groups despite being recommended on the SPICE Care Management Screens for three of the groups.

Rates of recording increased for all risk factors for all groups. Nevertheless, rates of recording increased most rapidly where it was explicitly incentivised in the QOF. We have estimated the overall increase in recording on the incentivised indicators to be 10 percentage points. The estimated impacts of the QOF were larger when comparison was made within incentivised conditions and largest when recording of BMI was the comparator. Responses were greater on

indicators that attracted more payment and required more stringent performance. Practices that were performing less well prior to the QOF showed the greatest response.

We also analysed the improvement on one new completely new indicator. The improvement in recording of smoking status for COPD was larger than for any of the other conditions. This suggests that a financial incentive not previously supported by informational intervention can also result in substantial improvements in performance.

We have focused on achievement in terms of recording of risk factors rather than achievement of improved outcomes. Improved recording is a pre-requisite for achievement of improved risk factor control. We have relied on the completeness of computerised records but have adjusted for general improvements over time. Our study identifies and isolates double-counting of performance through co-morbidity of conditions. We have concentrated on quantifying incremental changes in coverage.

Our difference-in-difference design relies on the appropriateness of the comparator groups. This may not be suitable if there are significant spillovers (Blundell and Costa Dias, 2002). This is certainly possible. We would expect the QOF to divert effort towards activities that attract payment. This has a positive side for incentivised activities ("now that I am being paid for it, I will do X"); and a negative side for unincentivised activities ("now that I am being paid to do X, I don't have the motivation or time to do Z"). However, these negative effects might be offset by positive 'spillovers' through a new focus on quality in general, such as for incentivised activities but for all patients ("now I have learned to collect this information systematically, I will do this for all my patients…") and for all activities for incentivised patients ("while you are here, I might as well….").

Looking forward, we have not considered the extent to which new patients are brought into regular monitoring and the extent to which monitored patients become more regularly monitored. Furthermore, though we have allowed for co-morbidity in our design, we have not considered the implications of this double-counting for differences between disease areas and practices in the rewards for each intervention. In addition, though we have shown a larger response from previous poorer performers, the dataset is big enough to allow further examination of whether the QOF narrowed or widened variations in performance between practices. Each of these would be a suitable avenue for future research.

References

Buckman L. Is doctors' self interest undermining the National Health Service? *British Medical Journal* 2007; 334: 235.

Campbell S, Reeves D, Kontopantelis E, Middleton E, Sibbald B, Roland M. Quality of primary care in England with the introduction of pay for performance. *New England Journal of Medicine* 2007; 357(2): 181-90.

Elder R, Kirkpatrick M, Ramsay W, Macleod M, Guthrie B, Sutton M, Watt G. *Measuring quality in primary medical services using data from SPICE*. Clinical Indicators Support Team, ISDScotland, Edinburgh, 2007. www.indicators.scot.nhs.uk

Gulliford MC, Ashworth M, Robotham D, Mohiddin A. Achievement of metabolic targets for diabetes by English primary care practices under a new system of incentives. *Diabetic Medicine* 2007; 24(5): 505-11.

Leyland AH, Dundas R, McLoone P, Boddy A. *Inequalities in mortality in Scotland 1981-2001*. MRC Social and Public Health Sciences Unit. Occasional Paper no.16. February 2007.

Guthrie B, McLean G, Sutton M. Workload and reward in the quality and outcomes framework of the 2004 general practice contract. *British Journal of General Practice* 2006; 56: 836-41.

NHS Confederation and British Medical Association. *New GMS contract: Investing in General Practice*. London 2003.

Roland M. Linking physician pay to quality of care: a major experiment in the United Kingdom. *New England Journal of Medicine* 2004;351:1448-54.

Steel N, Maisey S, Clark A, Fleetcroft R, Howe A. Quality of clinical primary care and targeted incentive payments: an observational study. *British Journal of General Practice* 2007; 57: 449-54.

Tahrani AA, McCarthy M, Godson J, Taylor S, Slater H, Capps N, Moulik P, Macleod AF. Diabetes care and the new GMS contract: the evidence for a whole county. *British Journal of General Practice* 2007; 57: 483-5.

Table 1: Maximum points available in QOF for recording (+ managing) risk factors

Risk factor	Diabetes	CHD	Stroke	Hypertension	COPD	Other
Blood pressure	3 (+17)	7 (+19)	2 (+5)	20 (+56)	-	15
Body Mass Index	3 (+0)	SPICE	-	SPICE	-	-
Cholesterol	3 (+6)	7 (+16)	2 (+5)	-	-	-
Alcohol status	SPICE	SPICE	-	SPICE	-	-
Smoking status	3 (+5)	7 (+4)	3 (+2)	10 (+10)	6 (+6)	11

Table 2 Summary statistics

Variable	Mean	S.D.	Min.	Max.
Risk factor recorded	0.331	0.470	0	1
Individual is male	0.473	0.499	0	1
Individual's age (years/100)	0.627	0.122	0.45	1
Individual diagnosed with Diabetes	0.099	0.298	0	1
Individual diagnosed with CHD (excludes above)	0.130	0.336	0	1
Individual diagnosed with Stroke (excludes above)	0.040	0.197	0	1
Individual diagnosed with Hypertension (excludes above)	0.235	0.424	0	1
Individual diagnosed with COPD (excludes above)	0.026	0.158	0	1
Year=2002	0.181	0.385	0	1
Year=2003	0.201	0.401	0	1
Year=2004	0.219	0.413	0	1
Year=2005	0.234	0.423	0	1
Recording is or becomes incentivised in the QOF	0.468	0.499	0	1
Maximum points facilitated by recording	9.5	16.7	0	76
Threshold at which maximum QOF points achieved (%)	39.3	42.2	0	90
Width of measurement interval allowed in QOF (months)	17.3	24.3	0	60
Practice list size	5075.0	2611.4	30	10849
Square root of practice list size	68.5	19.6	5.5	104.2
Recording rate at practice in 2002	0.200	0.241	0	1

Table 3 Probit regression - all conditions and risk factors

Model	(1)		(2)		(3)		(4)	
Variable	dF/dx	z	dF/dx	z	dF/dx	z	dF/dx	z
Age	-4.839	-8.80	-5.132	-9.25	-5.286	-9.72	-6.684	-10.45
Age2	8.802	10.52	9.318	11.02	9.582	11.57	12.184	12.49
Age3	-5.187	-12.53	-5.481	-13.07	-5.627	-13.73	-7.133	-14.73
Male	-0.248	-1.77	-0.225	-1.60	-0.220	-1.57	-0.116	-0.64
Male*age	0.234	0.35	0.125	0.19	0.108	0.16	-0.589	-0.70
Male*age2	0.545	0.54	0.700	0.68	0.710	0.70	1.927	1.51
Male*age3	-0.500	-1.00	-0.569	-1.12	-0.565	-1.11	-1.201	-1.90
Year=2002	0.075	7.44	0.075	7.52	0.076	7.65	-	-
Year=2003	0.258	19.94	0.256	20.05	0.257	19.69	-	-
Year=2004	0.354	21.40	0.348	21.37	0.364	12.47	0.133	4.57
Year=2005	0.355	20.99	0.349	20.98	0.365	12.43	0.136	4.66
CHD[1]	0.439	47.85	0.389	35.14	0.248	3.24	0.172	2.13
COPD[1]	0.216	37.78	0.166	28.89	0.018	0.26	-0.014	-0.18
Diabetes[1]	0.558	62.27	0.522	49.97	0.396	5.16	0.269	3.30
Hypertension[1]	0.359	42.42	0.308	29.73	0.174	2.42	0.139	1.77
Stroke[1]	0.306	47.68	0.256	31.43	0.106	1.40	0.067	0.83
BMI[2]	0.052	7.91	0.052	8.01	0.052	7.93	0.006	0.73
Blood pressure[2]	0.125	13.36	0.131	14.55	0.130	14.33	0.080	6.61
Cholesterol[2]	-0.054	-5.63	-0.058	-6.13	-0.058	-6.09	-0.041	-3.19
Smoking[2]	-0.017	-2.01	0.017	2.00	0.018	2.09	0.045	3.76
In QOF	0.169	30.48						
Facilitated points			0.001	6.95	0.001	6.63	0.001	2.74
Upper threshold			0.001	17.01	0.002	19.50	0.003	16.34
Measurement interval width			0.000	1.93	-0.001	-5.50	-0.001	-6.36
Prevalence rate					-0.092	-1.11	-0.163	-1.72
Square root of listsize					-0.001	-1.80	-0.001	-1.46
Pre-QOF achievement rate							0.465	8.65
Interactions with post-QOF period								
In QOF	0.098	11.67						
Facilitated points			0.000	0.59	0.000	1.36	0.001	4.36
Upper threshold			0.003	18.96	0.002	12.42	0.001	7.85
Measurement interval width			-0.003	-16.93	-0.001	-7.54	-0.001	-7.24
Prevalence rate					-0.121	-10.25	-0.068	-5.36
Square root of listsize					0.000	0.97	0.000	0.68
Pre-QOF achievement rate							-0.255	-5.68
N	5548859		5548859		5548859		3624492	
Pseudo-R2	0.2630		0.2675		0.2690		0.2665	

[1] Reference category = 'other' patients. [2] Reference category = alcohol status.

Table 4 Probit regression - conditions and risk factors included in SPICE

Model	(1)		(2)		(3)		(4)	
Variable	dF/dx	z	dF/dx	z	dF/dx	z	dF/dx	z
Age	-4.920	-6.55	-4.889	-6.49	-4.901	-6.45	-4.787	-6.88
Age2	9.957	8.95	9.913	8.89	9.946	8.86	9.875	9.66
Age3	-6.273	-11.59	-6.252	-11.53	-6.275	-11.53	-6.248	-12.68
Male	0.256	1.14	0.254	1.13	0.252	1.12	0.248	1.04
Male*age	-1.416	-1.36	-1.404	-1.34	-1.396	-1.33	-1.391	-1.24
Male*age2	2.252	1.44	2.230	1.43	2.215	1.41	2.198	1.31
Male*age3	-1.050	-1.37	-1.037	-1.36	-1.027	-1.34	-1.011	-1.24
Year=2002	0.105	10.80	0.105	10.79	0.106	10.62	-	-
Year=2003	0.310	22.41	0.310	22.46	0.311	21.12	-	-
Year=2004	0.396	22.38	0.396	22.36	0.373	10.45	0.123	3.82
Year=2005	0.407	21.84	0.408	21.83	0.384	10.74	0.134	4.25
CHD[1]	-0.152	-16.96	0.116	16.54	0.116	8.69	0.064	5.40
Hypertension[1]	-0.239	-24.27	0.265	34.61	0.265	22.55	0.150	12.74
Stroke[1]	-0.278	-34.56	0.019	1.76	0.019	0.62	0.002	0.08
BMI[2]	0.061	7.46	0.066	8.22	0.066	8.22	0.007	0.81
Blood pressure[2]	0.250	21.25	0.219	19.05	0.219	19.03	0.137	10.15
Cholesterol[2]	0.002	0.18	0.006	0.46	0.006	0.46	0.037	2.88
Smoking[2]	-0.001	-0.06	0.023	2.21	0.023	2.21	0.052	4.47
ln QOF	0.148	17.69						
Facilitated points			0.001	5.63	0.001	5.73	0.000	1.79
Upper threshold			0.001	11.07	0.001	9.76	0.002	9.48
Measurement interval width			0.000	-1.91	0.000	-1.68	0.000	-1.80
Prevalence rate					0.024	0.04	0.344	0.74
Square root of listsize					-0.001	-0.86	0.000	-0.95
Pre-QOF achievement rate							0.443	9.66
Interactions with post-QOF period								
ln QOF	0.165	14.72						
Facilitated points			0.001	2.33	0.001	1.75	0.001	3.80
Upper threshold			0.002	10.48	0.002	7.86	0.001	4.61
Measurement interval width			-0.001	-2.50	-0.001	-1.88	-0.001	-3.34
Prevalence rate					-0.006	-0.01	-0.307	-0.99
Square root of listsize					0.000	0.76	0.000	0.71
Pre-QOF achievement rate							-0.206	-4.63
N	2706091		2706091		2706091		1794497	
Pseudo-R2	0.2084		0.209		0.2092		0.1736	

[1] Reference category = diabetes patients. [2] Reference category = alcohol status.

Table 5 Probit regression – comparisons across risk factors

Diagnoses	CHD		Hypertension, CHD		Hypertension, CHD, Diabetes	
Risk factors	Chol, BMI		BP, BMI		Alcohol, Smoking	
Variable	dF/dx	z	dF/dx	z	dF/dx	z
Age	-2.867	-2.37	-7.691	-13.93	-4.422	-8.79
Age2	8.247	4.79	13.685	16.97	8.851	11.99
Age3	-6.172	-7.68	-7.867	-20.37	-5.570	-15.68
Male	0.370	1.04	0.040	0.21	0.097	0.57
Male*age	-1.407	-0.86	-0.580	-0.66	-0.775	-0.99
Male*age2	1.503	0.63	1.133	0.86	1.464	1.26
Male*age3	-0.350	-0.31	-0.537	-0.83	-0.737	-1.29
Year=2002	0.122	36.12	0.090	45.11	0.093	47.14
Year=2003	0.312	99.27	0.248	135.41	0.339	190.35
Year=2004	0.392	108.85	0.316	149.89	0.425	210.14
Year=2005	0.413	115.71	0.342	165.77	0.423	210.74
CHD	-	-	0.075	56.48	-0.102	-67.54
Hypertension	-	-	-	-	-0.220	-160.89
Blood pressure	-	-	-	-	-	-
BMI	0.018	6.86	-0.313	-202.56	-	-
Smoking	-	-	-	-	0.150	102.53
QOF	0.255	64.28	0.224	93.07	0.149	70.18
N	288044		809120		1028536	
Pseudo-R2	0.1568		0.1989		0.1623	

Table 6 Probit regression – comparisons across conditions

Diagnoses	Hypertension, CHD, Diabetes		Hypertension, CHD, Diabetes, Stroke, COPD	
Risk factor	BMI		Smoking	
Variable	dF/dx	z	dF/dx	z
Age	-7.627	-10.48	-3.541	-5.40
Age2	13.862	12.93	7.455	7.79
Age3	-8.174	-15.79	-4.872	-10.64
Male	0.314	1.31	0.132	0.59
Male*age	-1.783	-1.59	-1.023	-1.00
Male*age2	2.848	1.70	1.863	1.23
Male*age3	-1.322	-1.61	-0.928	-1.26
Year=2002	0.091	33.96	0.087	34.94
Year=2003	0.225	88.32	0.372	177.51
Year=2004	0.334	129.84	0.520	266.80
Year=2005	0.357	140.82	0.502	255.81
CHD	-0.296	-119.18	0.126	27.71
Hypertension	-0.441	-182.57	-0.002	-0.47
Diabetes	-	-	0.231	51.77
Stroke	-	-	0.010	2.00
QOF	0.077	19.50	-	-
Not previously in SPICE	-	-	0.101	15.47
N	514268		587586	
Pseudo-R2	0.1472		0.197	

Figure 1: Numbers in each disease group included in mutually exclusive classification

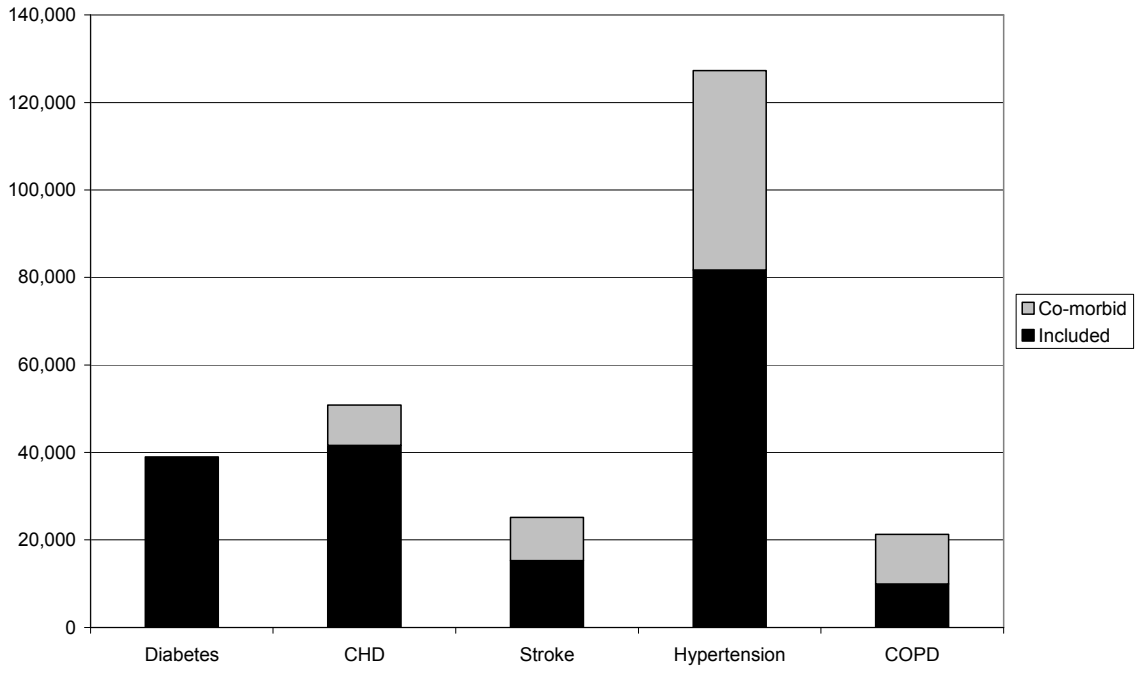
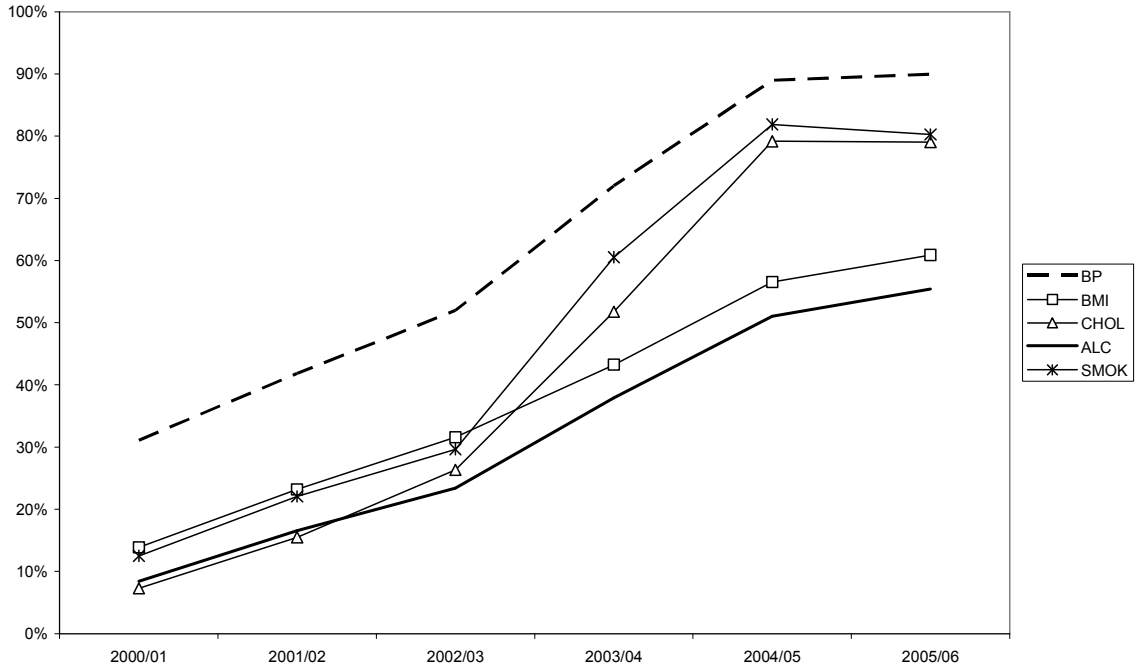


Figure 2 Recording of five risk factors for CHD patients



Excludes patients with diabetes.

Figure 3 Alcohol status recording by group and year

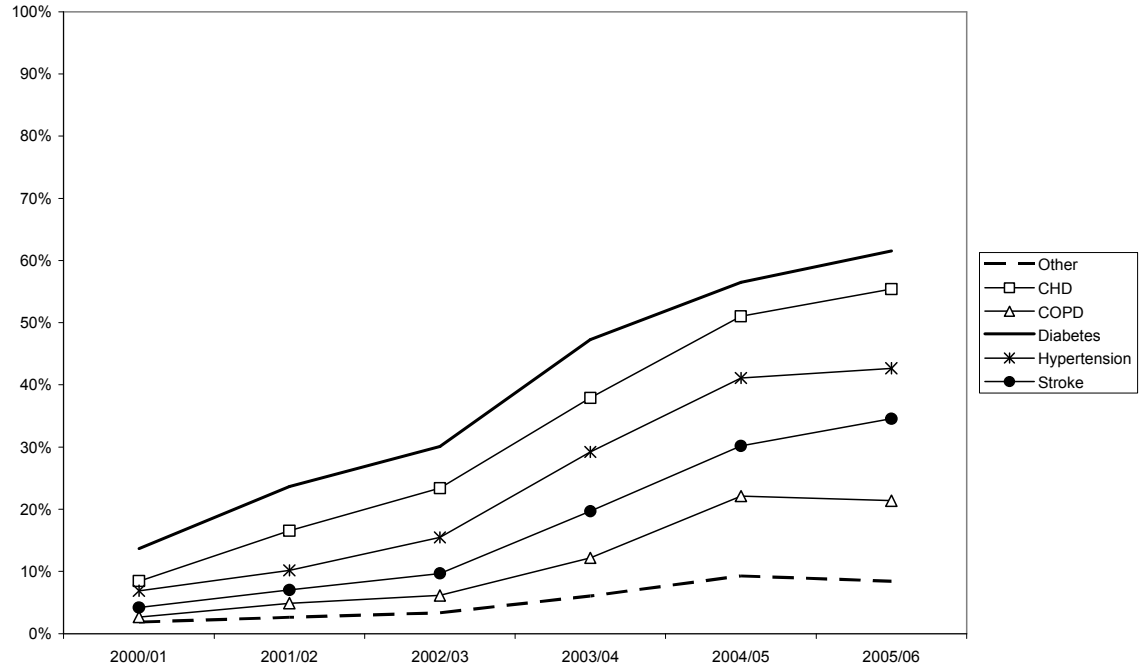


Figure 4 Smoking status recording by group and year