*Family Practice*, 2019, 607–613
doi:10.1093/fampra/cmy128
Advance Access publication 21 December 2018

OXFORD

Health Service Research

# An evaluation of a multifaceted, local Quality Improvement Framework for long-term conditions in UK primary care

**Frank Gabel[a,*], Ruth Chambers[b], Tracey Cox[c], Stefan Listl[a,d] and Neal Maskrey[e]**

[a]Translational Health Economics Group, Department of Conservative Dentistry, Heidelberg University, Heidelberg 69120, Germany, [b]Stoke-on-Trent Clinical Commissioning Group, Stoke-on-Trent, Staffs ST1 4FA, UK, [c]Stafford and Surrounds Clinical Commissioning Group and Cannock Chase Clinical Commissioning Group, Stafford ST16 2LP, UK, [d]Department of Quality and Safety of Oral Health Care, Radboud University, Nijmegen 6500, The Netherlands and [e]Keele University, Staffordshire ST5 5BG, UK.

*Correspondence to: Frank Gabel, Translational Health Economics Group, Heidelberg University, Heidelberg 69120, Germany; E-mail: frank.gabel@med.uni-heidelberg.de

## Abstract

**Background.** The evidence that large pay-for-performance schemes improve the health of populations is mixed—evidence regarding locally implemented schemes is limited.
**Objective.** This study evaluates the effects in Stoke-on-Trent of a local, multifaceted Quality Improvement Framework including pay for performance in general practice introduced in 2009 in the context of the national Quality and Outcomes Framework that operated from 2004.
**Methods.** We compared age-standardized mortality data from all 326 local authorities in England with the rates in Stoke-on-Trent using Difference-in-Differences, estimating a fixed-effects linear regression model with an interaction effect.
**Results.** In addition to the existing downward trend in cardiovascular deaths, we find an additional annual reduction of 36 deaths compared with the national mean for coronary heart disease and 13 deaths per 100 000 from stroke in Stoke-on-Trent. Compared with the national mean, there was an additional reduction of 9 deaths per 100 000 people per annum for coronary heart disease and 14 deaths per 100 000 people per annum for stroke following the introduction of the 2009 Stoke-on-Trent Quality Improvement Framework.
**Conclusion.** There are concerns about the unintended consequences of large pay-for-performance schemes in health care, but in a population with a high prevalence of disease, they may at least initially be beneficial. This study also provides evidence that a local, additional scheme may further improve the health of populations. Such schemes, whether national or local, require periodic review to evaluate the balance of their benefits and risks.

**Key words:** General practice remuneration, health services evaluation, long-term conditions, mortality data, pay for performance, quality improvement

## Introduction

Stoke-on-Trent is an industrial conurbation with a ceramics, mining and steel heritage and a registered population of ~285,000. Of the 326 local authorities in England, Stoke-on-Trent is ranked the 16th most deprived, with large areas in the city ranked among the top 10% most deprived in the whole of England.

Across a range of health and lifestyle indicators, outcomes in Stoke-on-Trent are poor. Male life expectancy at birth in 2012 was 76.5 years compared with 79.4 years in England; female life expectancy was 80.6 and 83.1 years in England (1).

As a response to poor health indicators, a local Quality Improvement Framework (QIF) commenced in primary care in 2009. The important context for QIF was that in 2004 as part of a new contract for GPs, the UK introduced a large, national P4P scheme—the Quality and Outcomes Framework (QOF).

This article describes the evaluation of the local QIF up to 2015, in the context of the continuing national QOF.

## Methods

For outcomes, we used mortality data in the Compendium of Population Health Indicators. Seven indicators relevant to the long-term conditions included within the QIF analysed. For each indicator, Stoke-on-Trent was compared with (i) national, (ii) regional (West Midlands) and (iii) a basket of localities with similar population demographics and other characteristics relevant to the determinants of health (peer localities).

### The national pay-for-performance Quality and Outcomes Framework

In 2004, as part of a new contract for GPs, the UK government introduced a pay-for-performance scheme with 136 indicators. The population included in the indicators is defined by practice-based disease registers [e.g. patients with coronary heart disease (CHD)] and the indicator measures the achievement of evidence-based targets (e.g. 'the percentage of patients with coronary heart disease in whom the last blood pressure reading measured in the preceding 12 months is 150/90 mmHg or less'). The indicators covered the management of chronic disease, practice organization and patients' experiences with respect to care. Electronic clinical records, which were already used in many practices, became universal because they were needed to support payment for work undertaken, though GPs employed more administrative staff to collect the required data, and there was an acceleration of existing trends to shift care for chronic physical conditions to nurse-led clinics. Practices required more intensive internal and external management support to ensure they achieved the targets. Periodic revisions to the scheme added or removed indicators and topics depending on local priorities. Payments make up ~25% of general practice income, and 99.6% of general practices participated in the scheme, which remains voluntary. The scheme continues in England but has been replaced in Scotland, Wales and Northern Ireland.

### The quality improvement approach used in the local Quality Improvement Framework

The team leading the local QIF programme in Stoke-on-Trent designed and delivered a wide-ranging approach to quality improvement in all practices. The QIF had a local implementation strategy, which is a close fit with the evidence on the best approaches to spread good practice (2).

The aims were to identify patients with long-term conditions currently undiagnosed, to improve the management and treatment of people with those conditions and to reduce health inequalities both within localities in the city and between the Stoke-on-Trent population and other areas in England. The QIF was much more than a pay-for-performance scheme; a multifaceted design included data feedback on achievement of locally agreed chronic disease management standards, and an educational programme comprising (i) individual support as bursaries, (ii) multidisciplinary learning events for primary care teams and (iii) QIF-focussed practice visits from clinical leaders and managers to encourage sharing of approaches between practices (3).

Pre-requisites for annual review of acceptance of each practice into the QIF programme included thresholds for numbers of registered patients per whole time equivalent practice clinicians, prevalence rates for specific long-term conditions versus those expected, minimum QOF attainment of clinical indicators, completion of clinical audits and progress with addressing clinical indicators of unwarranted clinical variation. All of these were designed to, and became, more challenging over time.

A panel of local stakeholders including patients was convened each year to review attainment of progress with existing QIF indicators. Quality improvement support was individualized to each practice with annual practice-related comparative reports covering ~50 key indicators. These included the practice attainment in addressing adverse lifestyle issues such as smoking cessation quit rates, conversion rates for urgent cancer referrals, location of diagnosis of cancer, as well as comparison with peer practice populations and England average rates. Each year the practices that generated most concerns about attainment of the quality indicators were visited by the QIF team who agreed a regularly monitored development plan.

Practice income derived from the QIF was supplementary to practice's funding derived from national contracts. Payments were set at ~£6 per patient (an additional 4.4% of average, gross practitioner income) if all standards were achieved and gradually less if only part of the standards were achieved. The patient population registered with the 55 general practices in Stoke-on-Trent was 265 000 in 2009. Over the first 7 years of the QIF scheme, there was 100% participation of all practices; this includes on average three practices each year that failed to match the pre-requisite criteria for participation at the start of the year—for example, being able to meet data-availability criteria. All such practices remained engaged via quality and performance development in order to achieve the criteria for participation the following year, but did not receive in-year direct funding for that year.

Further details of the QIF and QOF designs are reported using TIDieR (4) checklists in Supplementary Material 1 and 2; Supplementary Material 3 shows QIF targets in detail.

### Statistical approach to mortality data

We used directly standardized mortality rates from a total of 326 local authorities in England. Data were available as three-year-rolling averages. The time frame consisted of yearly observations from 1995 to 2013, totalling a balanced panel of 5542 observations for each condition without any missing data.

The following key conditions were analysed: CHD, stroke, diabetes, chronic obstructive pulmonary disease (COPD), asthma, chronic kidney disease (CKD) and epilepsy. Data from four age bands (all age groups, <65 years, <75 years and 65–74 years) were available for several of the above conditions, helping to increase the validity of analyses.

### Statistical methods

To determine any impact of the 2009 QIF and the 2004 QOF, we used a differences-in-differences setup, estimating a fixed-effects linear regression model with an interaction effect and a linear time trend (5):

$$y_{it} = \beta_0 + \beta_1 \text{year} + \delta(\text{place}_{\text{treatment}} \times \text{after}_{\text{treatment}})_{it} + \gamma_i + \varepsilon_{it}$$

The coefficient of interest is δ, representing the effect of being in the treatment group (Stoke-on-Trent) after the treatment went into place (2004/2009). Here, $\text{place}_{\text{treatment}}$ and $\text{after}_{\text{treatment}}$ are the corresponding dummy variables. $\gamma_i$ is a one-hot-encoded variable representing a regional fixed effect, and $\varepsilon_{it}$ is an error term. Note that this approach also enables to check for an effect of the 2004 national QOF in Stoke-on-Trent; if there is no effect in Stoke-on-Trent, the interaction effect is insignificant.

This obtained coefficient is only valid under the parallel slope assumption. In the absence of this assumption, the above model yields biased estimates. Since this was the case for several of our models, we circumvented this problem by merging four respective pre- and post-treatment years, looking at the mean number of yearly deaths pre- and post-treatment. This procedure is indicated in the results by an asterisk in the corresponding tables.

To test for significant changes in mortality rates following the 2004 national QOF, we used an interrupted time-series regression with a linear time trend:

$$y_t = \beta_0 + \beta_1 \text{year} + \delta(\text{after}_{\text{treatment}})_t + \varepsilon_{it}$$

The variables have similar meanings as described earlier; however, as we are dealing with a national reform with no differential local implementation, there is neither a possible comparison of treated and untreated authorities nor a possibility of applying authority-level fixed effects. The validity of this secondary approach rests on the assumption that the slope, had there been no reform, would have continued to follow the same slope.

In order to account for eventual correlation in the data, authority-level cluster-robust standard errors were used in all regressions. Statistical analyses were performed with R version 3.2.1.

## Results

Table 1 (left panel) displays the mean mortality rates per 100 000 people in England before any intervention took effect.

We found that from 2004 the downward trend in the national mean mortality rate for the conditions we analysed increased by an additional 3.72 fewer deaths for CHD and 5.53 fewer deaths per 100 000 general population per annum for stroke (see Table 1, Fig. 1). This came in addition to a yearly trend indicating a reduction of 11.07 deaths per 100 000 general population per annum for

**Table 1.** Interrupted time series analysis of mortality rates in England following the 2004 Quality and Outcomes Framework introduction based on data from 1998 to 2014

| Absolute mortality rate before 2004 | Linear (yearly) time trend coefficient | Intervention coefficient |
|---|---|---|
| | [95% confidence interval] | [95% confidence interval] |
| CHD | | |
| All age groups | **−11.07** | **−3.72** |
| 256.93 | [−11.32, −10.82] | [−5.46, −1.97] |
| <65 years | **−1.85** | **0.39** |
| 39.45 | [−1.92, −1.77] | [−0.11, 0.89] |
| <75 years | **−5.05** | **−0.03** |
| 97.68 | [−0.20, −4.91] | [−0.87, 0.81] |
| 65–74 years | **−29.63** | **−3.22** |
| 544.07 | [−30.46, −28.79] | [−8.76, 2.33] |
| Stroke | | |
| All age groups | **−4.37** | **−5.53** |
| 129.89 | [−4.51, −4.22] | [−6.59, −4.47] |
| <65 years | **−0.42** | **−0.44** |
| 11.19 | [−0.45, −0.39] | [−0.71, −0.17] |
| <75 years | **−1.18** | **−1.40** |
| 28.53 | [−1.23, −1.13] | [−1.78, −1.02] |
| 65–74 years | **−6.97** | **−8.75** |
| 161.43 | [−7.29, −6.64] | [−11.35, 6.16] |
| Diabetes | | |
| All age groups* | **−0.35** | **0.30** |
| 13.61 | [−0.38, −0.32] | [0.09, 0.50] |
| Epilepsy | | |
| All age groups* | **0.02** | **−0.11** |
| 1.80 | [0.01, 0.03] | [−0.21, −0.00] |
| COPD | | |
| All age groups* | **−0.60** | **−0.73** |
| 56.48 | [−1.29, 0.09] | [−0.82, −0.65] |
| Asthma | | |
| All age groups* | **−0.08** | **−0.09** |
| 2.98 | [−0.09, −0.06] | [−0.23, 0.05] |
| CKD | | |
| All age groups* | **−0.04** | **0.17** |
| 3.16 | [−0.06, −0.02] | [0.00, 0.33] |

The first column represents coefficients of a linear (per year) time trend. The second column represents intervention coefficients of an interrupted time-series analysis. Negative interaction coefficients represent lowered mortality when compared with the baseline (before the reform).

CHD and 4.37 deaths per 100 000 general population per annum for stroke.

In relation to the comparison groups 'national' and 'West Midlands', pre-2004 mean mortality rates in Stoke-on-Trent (Table 2, italic) for all the conditions and age bands were considerably higher. The mortality rates of the peer group of local authorities show a mixed picture across conditions and age groups, and are mostly similar to Stoke-on-Trent (see Fig. 1).

### Effects of the introduction of the 2004 Quality Outcome Framework

We found a statistically significant greater benefit in Stoke-on-Trent on CHD mortality associated in time with the 2004 introduction of the national QOF with an additional reduction of 36 deaths per 100 000 general population per annum in Stoke-on-Trent compared with the national mean (see Table 2). This effect occurred in all age groups and is especially relevant for the 65- to 74-year-old age group with an excess reduction of 166 deaths per 100 000 per annum in this population group (see Fig. 1, upper right). When compared with the West Midlands, this effect consistently becomes less (for all age bands, see Fig. 1, upper left) over time; when compared with peer localities, the reduction marginally fails to be statistically significant for all age groups but is significant for all three age subgroups.

For stroke, we found a significant benefit on mortality for the 65- to 74-year age groups and <75-year age groups (13 fewer deaths per 100 000 per annum for the 65- to 74-year age group; –1 per 100 000 per annum for the <75-year age group) when compared with the national mean and the West Midlands. There was also a significant benefit in the 65- to 74-year age group when the comparison group was peer local authorities. Results in the all age group and <65 years showed small but statistically significant increases against the national comparator (see Table 2).

Analyses of the other conditions show a mixed picture with small reductions in CKD in Stoke-on-Trent and small adverse trends for deaths from diabetes, chronic obstructive kidney disease and asthma.

### Effects of the introduction of the 2009 Quality Improvement Framework

The pre-2009 mean mortality rates were higher in Stoke-on-Trent across all conditions and age bands compared with the regional and national means, with smaller absolute differences than in 2004. With some exceptions, the mortality rates for the conditions analysed were generally lower in Stoke-on-Trent in 2009 than the mean mortality rates of the peer local authorities.

Mortality rates for most conditions and age groups showed a clear reduction associated in time with the introduction of the 2009 QIF in Stoke-on-Trent (see Table 3). Compared with the national mean, there was an additional reduction of about 9 deaths per 100 000 people for CHD (see Fig. 1, bottom left) and a reduction of ~14 deaths per 100 000 people for stroke (see Fig. 1, bottom right). This effect remains when compared with the regional mean, but there was no significant difference when compared with the mean of the peer regions.

Analyses of other conditions showed a small reduction in mortality from diabetes, asthma and CKD consistent across all comparison group means. On the other hand, epilepsy and COPD showed small increases.

## Discussion

The cardiovascular health of the population of Stoke-on-Trent improved faster from 2004, with statistically significant greater

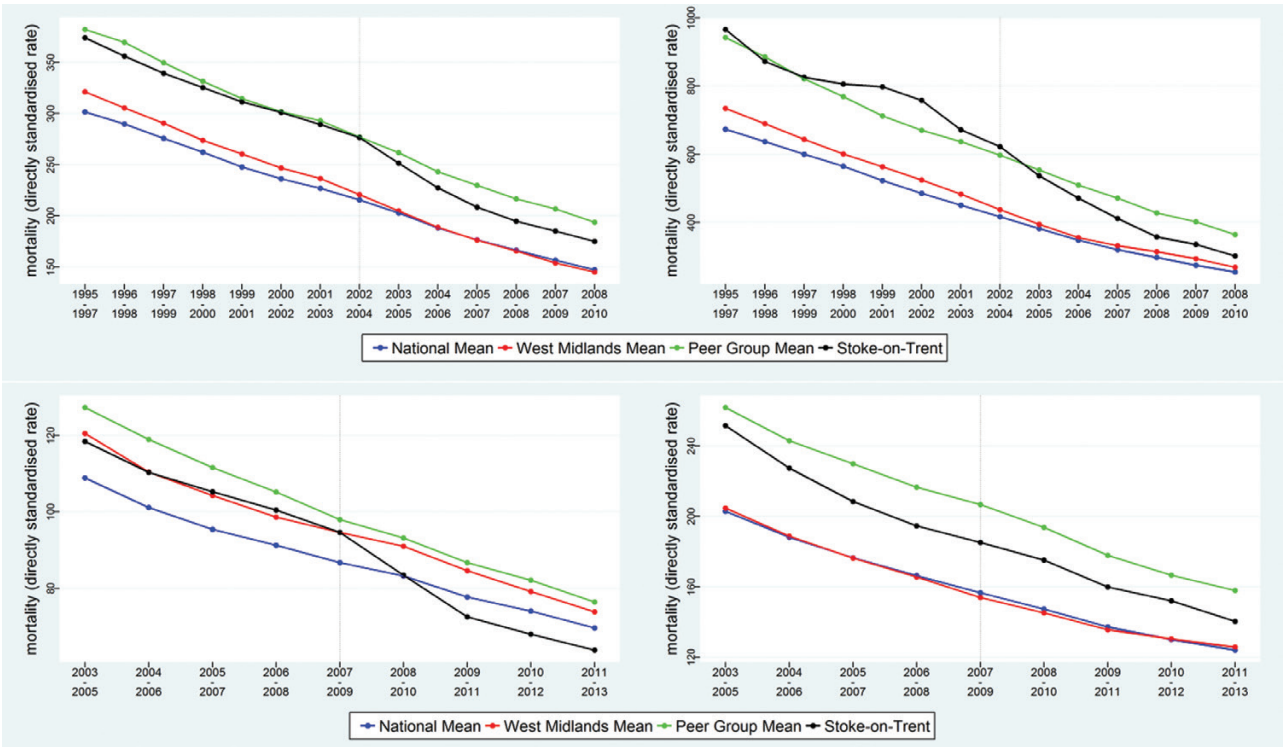


**Figure 1.** Time series of mortality rates from 1995 to 2013 for different geographic localities in England and for the 2004 Quality and Outcomes Framework and 2009 Quality and Outcomes Framework interventions. (a) The dashed black line refers to the 2004 national Quality and Outcomes Framework introduction—mortality time series are for cardiovascular heart disease in all age groups (left) and 65- to 74-year-old age groups (right). (b) The dashed black line refers to the 2009 local Quality Improvement Framework introduction—mortality time series are for stroke (left) and cardiovascular heart disease (right).

**Table 2.** Changes in mortality rates in Stoke-on-Trent associated with the 2004 Quality and Outcomes Framework (national, pay-for-performance scheme) in comparison to mortality rates in other geographic areas based on data from 1995 to 2013

| | National | West Midlands | Peer localities |
|---|---|---|---|
| | Intervention effect [95% confidence interval] | Intervention effect [95% confidence interval] | Intervention effect [95% confidence interval] |
| | *Mean difference to Stoke (pre-2004)* | *Mean difference to Stoke (pre-2004)* | *Mean difference to Stoke (pre-2004)* |
| CHD | | | |
| All age groups | **−35.85** [−37.87; −33.82] | **−24.69** [−30.73; −18.65] | **−13.58** [−28.24; 1.07] |
| | *63.01* | *53.35* | *−6.95* |
| <65 years | **−14.11** [−14.65; −13.57] | **−12.69** [−14.36; −11.03] | **−5.10** [−7.27; −2.93] |
| | *23.60* | *21.47* | *2.00* |
| <75 years | **−31.64** [−32.74; −30.55] | **−27.86** [−31.06; −24.66] | **−14.69** [−18.63; −10.75] |
| | 48.16 | *42.65* | *5.52* |
| 65–74 years | **−166.05** [−172.27; −159.84] | **−144.13** [−161.36; −126.90] | **−88.25** [−117.63; −58.88] |
| | *236.44* | *205.02* | *32.54* |
| Stroke | | | |
| All age groups | **3.60** [2.76; 4.44] | **6.21** [2.73; 9.70] | **6.12** [−0.01; 12.25] |
| | *5.97* | *−5.72* | *−14.29* |
| <65 years | **0.62** [0.41; 0.83] | **0.03** [−0.57; 0.63] | **1.96** [0.87; 3.05] |
| | *1.96* | *1.74* | *−2.74* |
| <75 years | **−0.96** [−1.27; −0.65] | **−1.33** [−2.54; −0.12] | **1.13** [−1.53; 3.79] |
| | *6.35* | *4.49* | *−3.99* |
| 65–74 years | **−13.10** [−15.30; −10.90] | **−11.78** [−19.90; −3.66] | **−5.27** [−27.14; 16.62] |
| | *39.98* | *25.54* | *−13.51* |
| Diabetes | | | |
| All age groups* | **1.90** [1.63; 2.17] | **1.39** [0.16; 2.62] | **1.27** [−0.63; 3.18] |
| | *0.87* | *−2.00* | *−0.54* |
| Epilepsy | | | |
| All age groups* | **0.06** [−0.02; 0.14] | **−0.12** [−0.45; 0.20] | **0.14** [−0.34; 0.63] |
| | *0.44* | *0.33* | *−0.18* |
| COPD | | | |
| All age groups* | **3.34** [2.75; 3.93] | **4.54** [2.37; 6.71] | **2.27** [−1.25; 5.79] |
| | *21.31* | *22.57* | *−8.00* |
| Asthma | | | |
| All age groups* | **2.18** [2.07; 2.28] | **1.74** [1.33; 2.14] | **2.66** [2.03; 3.30] |
| | *0.73* | *0.39* | *0.16* |
| CKD | | | |
| All age groups* | −1.29 [**−1.41**; **−1.16**] | −1.51 [**−1.95**; **−1.07**] | −1.21 [**−1.74**; **−0.67**] |
| | ***2.01*** | ***1.14*** | ***1.23*** |

The three columns represent differences-in-differences coefficients from different geographical comparison groups (national, West Midlands and peer regions) to compare mortality changes in Stoke-on-Trent. Negative interaction coefficients represent relatively improved mortality in Stoke-on-Trent following the 2004 QOF. Pre-QOF differences are in italics. Green highlight indicates a significant reduction in mortality; red highlight indicates a significant increase in mortality; and no highlight indicates a non-significant result.

improvements seen in Stoke-on-Trent when compared with most other populations. These were associated in time with the 2004 QOF and the 2009 QIF. The national improvement was a reduction of ~10 deaths per year per 100 000 of the general population (see Table 1); the additional effects associated with the QOF in Stoke-on-Trent per annum were 36 CHD deaths per 100 000 general population and 166 CHD deaths per 100 000 population of 65–74 years old (see Table 2).

Stroke mortality in the 65- to 74-year age group showed that in addition to the national effect of the 2004 QOF introduction of about −8.75 deaths per 100 000 population, there was an additional reduction of around −5 deaths per 100 000 per year in Stoke-on-Trent.

Following the 2009 local introduction of the QIF, there were further significant reductions of mortality rates for most conditions measured; again, these were largest for CHD and stroke. These effects remain when compared to the West Midlands but are not detectable in comparison with peer localities. A possible explanation for this is that when the QIF commenced, Stoke-on-Trent had improved its implementation of evidence-based interventions in response to the QOF to improve cardiovascular health better than those peer localities; therefore, the ability to further achieve a statistically significant reduction mortality was reduced because much of the available benefits had already been achieved. An alternative explanation is this might have occurred due to the statistical imprecision of the coefficients. The likely explanation of the failure to detect a statistically significant reduction in stroke in <65-year age group from 2009 (Table 3) is the low event rate at baseline and therefore the small number of potentially preventable events in that age group, especially in relatively small, sub-group population samples.

Benefits were greatest for the high-prevalence conditions amenable in the short term to evidence-based interventions—blood pressure lowering and lipid-lowering medicines, and the existing smoking cessation services and support. Self-reported, short-term

**Table 3.** Changes in mortality rates in Stoke-on-Trent following with the 2009 Quality Improvement Framework (local scheme) in comparison to mortality rates in other geographic areas based on data from 1995 to 2013

| Comparison | National | West Midlands | Peer regions |
|---|---|---|---|
| absolute mortality rate in Stoke pre-intervention | Intervention effect [95% confidence interval] | Intervention effect [95% confidence interval] | Intervention effect [95% confidence interval] |
| | *Mean difference to Stoke (pre-2009)* | *Mean difference to Stoke (pre-2009)* | *Mean difference to Stoke (pre-2009)* |
| CHD | | | |
| All age groups | **−8.85** [−10.11; −7.60] | **−10.77** [**−15.47; −6.07**] | **4.60** [−1.82; 11.01] |
| *321.52* | *29.21* | *31.65* | *−22.99* |
| <65 years | **−4.98** [−5.36; −4.60] | **−4.90** [−6.40; −3.40] | **−2.32** [−4.72; 0.09] |
| *63.30* | *12.43* | *11.59* | *−1.39* |
| <75 years | **−7.90** [−8.59; −7.20] | **−7.15** [−9.55; −4.74] | **0.06** [−4.46; 4.58] |
| *147.14* | *21.36* | *19.32* | *−7.89* |
| 65–74 years | **−30.25** [−34.42; −26.07] | **−24.36** [−37.44; −11.29] | **18.27** [−5.79; 42.34] |
| *789.96* | *89.84* | *78.59* | *−57.76* |
| Stroke | | | |
| All age groups | **−13.61** [−14.46; −12.77] | **−10.21** [−12.87; −7.55] | **−6.81** [−14.82; 1.21] |
| *132.56* | *7.97* | *−1.22* | *−8.15* |
| <65 years | **−0.06** [−0.25; 0.14] | **0.59** [0.01; 1.16] | **1.10** [−0.27; 2.47] |
| *13.64* | *1.87* | *1.27* | *−1.44* |
| <75 years | **−3.03** [−3.34; −2.72] | **−1.33** [−2.45; −0.22] | **−0.66** [−2.37; 1.04] |
| *35.41* | *4.44* | *2.55* | *−3.12* |
| 65–74 years | **−23.24** [−25.24; −21.25] | **−16.33** [−24.00; −8.67] | **−11.83** [−23.71; 0.05] |
| *202.36* | *21.62* | *12.65* | *−18.40* |
| Diabetes | | | |
| All age groups | **−3.59** [−3.82; −3.37] | **−2.08** [−3.02; −1.13] | **−2.63** [−3.26; −2.00] |
| *14.49* | *2.18* | *−1.22* | *0.76* |
| Epilepsy | | | |
| All age groups | **0.69** [0.62; 0.77] | **0.44** [0.12; 0.76] | **0.99** [0.48; 1.50] |
| *2.57* | *0.20* | *−0.10* | *−0.35* |
| COPD | | | |
| All age groups | **5.59** [5.07; 6.11] | **5.74** [3.74; 7.74] | **5.35** [1.60; 9.10] |
| *76.32* | *27.09* | *29.15* | *−3.01* |
| Asthma | | | |
| All age groups | **−0.88** [−0.97; −0.79] | **−0.62** [−1.00; −0.24] | **−0.93** [−1.32; −0.54] |
| *3.66* | 1.76 | 1.05 | 1.44 |
| CKD | | | |
| All age groups | **−1.10** [−1.22; −0.98] | **−0.78** [−1.31; −0.25] | **−0.85** [−1.80; 0.09] |
| *5.30* | *1.10* | *0.08* | 0.59 |

The three columns represent differences-in-differences coefficients from different geographical comparison groups (national, West Midlands and peer regions) to compare mortality changes in Stoke-on-Trent. Negative interaction coefficients represent relatively improved mortality in Stoke-on-Trent following the 2009 QIF. Pre-QIF differences are in italics.

smoking cessation rates were high in Stoke-on-Trent during this time (6), and sample data showed that hypertension and cholesterol levels improved locally during the relevant time period. By 2014/2015, detection and control of hypertension were better than comparable localities, while overall smoking, diet and activity indices in Stoke-on-Trent continued to be adverse (7,8). A differential increase in the effectiveness of the acute, secondary care treatment of myocardial infarction and stroke in Stoke-on-Trent compared with other localities is a possible but unlikely explanation for our findings. It seems far more plausible that the mortality of high prevalence chronic diseases such as CHD and stroke is more amenable to primary prevention interventions than secondary care interventions.

The differences in mortality rates were detectable and associated in time with the QIF that started in 2009, after 5 years of QOF, is a very interesting finding. For there to be detectable, small mortality benefits across several conditions—including diabetes, asthma and CKD where simple short-term therapeutic interventions are less likely to result in detectable improvements in mortality data—is notable.

The results demonstrate that some important outcomes that health care quality improvement schemes seek to address can be satisfactorily assessed using publicly available mortality data. However, there are well-known limitations to mortality data (notably diagnostic imprecision), and local authority populations do not map directly to patients registered with practices in clinical commissioning groups. Hence, the associations identified in the study should not necessarily be interpreted as causal effects. Also, given the large year-to-year variability in the data of the lesser prevalent conditions, the corresponding results should be treated with care, since eventual effects of any change could be concealed by random variation.

A detailed review of the literature evaluating pay-for-performance schemes was undertaken to inform this evaluation. The evidence that large, complex, pay-for-performance schemes improve the health of populations is mixed (9–14), and we did not find examples of local schemes similar to the Stoke-on-Trent QIF with its multifaceted approach combining P4P, professional and managerial support and monitoring, and educational co-initiatives. In summary, the concerns with P4P schemes are a lack of evidence of

benefits associated with the schemes, loss of focus on conditions outwith schemes, schemes not being relevant to local health priorities, mechanistic approaches to individual care as clinicians 'follow the rules' irrespective of whether the intervention is appropriate for that patient (including their values and preferences), and the sheer burden of administration and management on the workforce. Perhaps the most important finding in the many evaluations of the UK QOF is that it was associated with a reduction in health inequalities (15); this analysis supports that finding.

## Conclusion

It seems plausible that both the QOF from 2004 and QIF from 2009 may have contributed to reducing premature mortality from some important conditions in this specific locality. Given the limitations of large, national, pay-for-performance schemes, the question is what now replaces large-scale, complex, invasive, mandatory measurement as the dominant approach in some health systems to reduce unwarranted variation in provided care (16). Despite several inherent analytical limitations, a local, multifaceted scheme incorporating P4P alongside other locally agreed strategies may improve the health of populations. In the short term, benefits may only occur for common conditions for which there are simple, safe, effective, acceptable interventions in localities with high event rates. Benefits may be more difficult to achieve when disease-specific pathophysiology is more complex, and when event rates in the targeted diseases drop over time, presumably in part due to early gains resulting from the more consistent adoption of interventions in vulnerable populations. Nevertheless, local approaches, if they are well led and managed, may overcome many of the drawbacks of national schemes.

## Supplementary material

Supplementary data are available at *Family Practice* online.

## Declaration

Funding: N.M. received part-time funding for 12 months from Stoke-on-Trent Clinical Commissioning Group to support all aspects of the evaluation of the Quality Improvement Framework including these analyses. F.G. received an honorarium from Stoke-on-Trent Clinical Commissioning Group to support statistical analyses. The total funding was £26 000; no other funding was provided for this study.

Ethical approval: none.

Conflict of interest: all authors have read and understood the ICJME statement (http://www.icmje.org/coi_disclosure.pdf). F.G. declares that he received an honorarium from Stoke-on-Trent Clinical Commissioning Group for designing and undertaking statistical analyses. N.M. declares that he received personal, part-time research fees for 12 months from Stoke-on-Trent Clinical Commissioning Group to undertake an evaluation of pay-for-performance schemes of which these analyses form a part; R.C. declares her previous role as a GP partner at Furlong Medical Centre, Stoke-on-Trent, when her practice received payments from the national and local pay-for-performance schemes. Since 1 January 2016, R.C. has been Clinical Chair of Stoke-on-Trent Clinical Commissioning Group that has commissioned the local pay-for-performance scheme since replacing the Primary Care Trust in 2013. S.L. declares that he received no funding or payments in relation to this manuscript and that—during the past 5 years and for work unrelated to this manuscript—he received research funding from the US National Institutes of Health (project reference: R01-DE021678) and the European Commission (project reference: 635183), as well as an honorarium from the ADA Health Policy Institute for consultancy work in relation to Harvard University's Malaysia Health Systems Reform project. T.C. has nothing to declare.

## Data availability

Raw mortality data used in this study were extracted by Sam Dunn, Research and Information Analyst, Safer City Partnership, Directorate of Public Health, City of Stoke-on-Trent. The data are openly accessible via https://indicators.hscic.gov.uk/webview/.

## References

1. *Stoke-on-Trent Clinical Commissioning Group Annual Report, 2014/15*. https://www.stokeccg.nhs.uk/stoke-governance/annual-reports (accessed on 9 February 2018).
2. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. *Milbank Q* 2004; 82: 581–629.
3. Cox T. *Using a Quality Improvement Framework to Make Local Population Health Gains*. Shared Learning Database, National Institute for Health and Care Excellence. 2012. https://www.nice.org.uk/sharedlearning/using-a-quality-improvement-framework-to-make-local-population-health-gains (accessed on 9 February 2018).
4. Hoffmann TC, Glasziou PP, Boutron I *et al*. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014; 348: g1687.
5. Dimick JB, Ryan AM. Methods for evaluating changes in health care policy: the difference-in-differences approach. *JAMA* 2014; 312: 2401–2.
6. NHS England. Commissioning for Value. 2013. https://www.england.nhs.uk/2013/10/comm-val-data-pcks/ (accessed on 9 February 2018).
7. Cochrane T, Davey R, Iqbal Z *et al*. NHS health checks through general practice: randomised trial of population cardiovascular risk reduction. *BMC Public Health* 2012; 12: 944.
8. Public Health England. National Cardiovascular Intelligence Network. Cardiovascular Disease Profiles. https://www.gov.uk/guidance/cardiovascular-disease-data-and-analysis-a-guide-for-health-professionals (accessed on 9 February 2018).
9. Doran T, Kontopantelis E, Valderas JM *et al*. Effect of financial incentives on incentivised and non-incentivised clinical activities: longitudinal analysis of data from the UK quality and outcomes framework. *BMJ* 2011; 342. doi:10.1136/bmj.d3590
10. Campbell SM, Reeves D, Kontopantelis E, Sibbald B, Roland M. Effects of pay for performance on the quality of primary care in England. *N Engl J Med* 2009; 361: 368–78.
11. Roland M, Campbell S. Successes and failures of pay for performance in the United Kingdom. *N Engl J Med* 2014; 370: 1944–9.
12. Ryan AM, Krinsky S, Kontopantelis K, Doran T. Long-term evidence for the effect of pay-for-performance in primary care on mortality in the UK: a population study. *Lancet* 2016. doi:10.1016/S0140-6736(16)00276-2
13. Van Herck P, De Smedt D, Annemans L, Remmen R, Rosenthal M. Sermeus systematic review: effects, design choices, and context of pay-for-performance in health care. *BMW Health Serv Res* 2010; 315: 1329–30.
14. Petersen LA, Woodard LD, Urech T, Daw C, Sookanan S. Does pay-for-performance improve the quality of health care? *Ann Intern Med* 2006; 145: 265–72.
15. Doran T, Fullwood C, Kontopantelis E, Reeves D. Effect of financial incentives on inequalities in the delivery of primary clinical care in England: analysis of clinical activity indicators for the quality and outcomes framework. *Lancet* 2008; 372: 728–36.
16. Berwick DM. Era 3 for medicine and health care. *JAMA* 2016; 315: 1329–30.