

RESEARCH



Investigating the relationship between quality of primary care and premature mortality in England: a spatial whole-population study

OPEN ACCESS

Evangelos Kontopantelis *senior research fellow*^{1,2}, David A Springate *research fellow*^{2,3}, Mark Ashworth *senior lecturer*⁴, Roger T Webb *reader*⁵, Iain E Buchan *professor*¹, Tim Doran *professor*⁶

¹Centre for Health Informatics, Institute of Population Health, University of Manchester, Manchester, UK; ²NIHR School for Primary Care Research, Centre for Primary Care, Institute of Population Health, University of Manchester; ³Centre for Biostatistics, Institute of Population Health, University of Manchester; ⁴Primary Care and Public Health Sciences, King's College London, London, UK; ⁵Centre for Mental Health and Risk, University of Manchester; ⁶Department of Health Sciences, University of York, York, UK

Abstract

Objectives To quantify the relationship between a national primary care pay-for-performance programme, the UK's Quality and Outcomes Framework (QOF), and all-cause and cause-specific premature mortality linked closely with conditions included in the framework.

Design Longitudinal spatial study, at the level of the "lower layer super output area" (LSOA).

Setting 32482 LSOAs (neighbourhoods of 1500 people on average), covering the whole population of England (approximately 53.5 million), from 2007 to 2012.

Participants 8647 English general practices participating in the QOF for at least one year of the study period, including over 99% of patients registered with primary care.

Intervention National pay-for-performance programme incentivising performance on over 100 quality-of-care indicators.

Main outcome measures All-cause and cause-specific mortality rates for six chronic conditions: diabetes, heart failure, hypertension, ischaemic heart disease, stroke, and chronic kidney disease. We used multiple linear regressions to investigate the relationship between spatially estimated recorded quality of care and mortality.

Results All-cause and cause-specific mortality rates declined over the study period. Higher mortality was associated with greater area deprivation, urban location, and higher proportion of a non-white population. In general, there was no significant relationship between practice performance on quality indicators included in the QOF and all-cause or cause-specific mortality rates in the practice locality.

Conclusions Higher reported achievement of activities incentivised under a major, nationwide pay-for-performance programme did not seem to result in reduced incidence of premature death in the population.

Introduction

Primary care has enormous potential to improve population health outcomes—including mortality from common chronic conditions—through early intervention in the disease process^{1,2} and coordinated provision of care. Effective primary care is associated with reduced morbidity, increased longevity, and more equitable health outcomes,^{3,4} but quality of primary care varies widely between providers.^{5,6} Traditional physician payment systems have facilitated this variation, with fee-for-service systems potentially incentivising over-investigation and over-treatment, and capitation systems potentially incentivising under-utilisation. Neither approach directly rewards high quality care or investment in quality improvement.⁷⁻⁹

In order to improve patient outcomes, policymakers worldwide have attempted to link remuneration for providers to quality of care through pay-for-performance programmes. Multiple programmes have been implemented across a range of settings, but clear evidence for improved patient outcomes is yet to emerge.¹⁰⁻¹² In the United Kingdom a national primary care incentive scheme was introduced in 2004. The Quality and Outcomes Framework (QOF), one of largest pay-for-performance programmes in the world, links up to 25% of family practitioners' income to performance on over 100

Correspondence to: E Kontopantelis e.kontopantelis@manchester.ac.uk

Extra material supplied by the author (see <http://www.bmj.com/content/350/bmj.h904?tab=related#datasupp>)

Online appendix 1: supplementary tables A1 to A3

Online appendix 2: supplementary tables B1 to B5, supplementary figs B1 to B3

Online appendix 3: supplementary figs C1 to C6

publicly reported quality indicators. Several indicators relate to organisation of care and patient experience, but most relate to management of common chronic diseases,¹³ including the leading causes of death in the UK: coronary heart disease, cancer, chronic obstructive pulmonary disease, hypertension, stroke, and diabetes. The breadth of the scheme should not be underestimated: in 2011-12 the total register of the 19 conditions incentivised under the QOF represented data collected from approximately 62% of the total number of registered patients.

The QOF was associated with an initial improvement in incentivised processes of care and some intermediate outcomes, but there was little further improvement after the third year of the scheme (2006-7).¹⁴⁻¹⁶ Evidence for improved patient outcomes, however, is contradictory. For example, emergency hospital admission rates seem to have decreased for diabetes¹⁷ but not for some other incentivised conditions.¹⁸⁻¹⁹ Evidence on the benefits of several processes incentivised in the QOF suggests that improvements in performance of these processes should reduce population mortality.²⁰ However, no studies to date have examined the relationships between recorded practice performance under the QOF and premature death rates for conditions included in the programme.

In this study, we assessed these relationships for the first time. More specifically, we applied spatial analysis techniques to quantify the association between practices' performance on the QOF indicators, including improvement on these indicators, and all-cause premature mortality in the areas the practices serve, controlling for area and population characteristics. We also examined the relationship between performance on specific intermediate outcomes indicators that are more closely associated with patient outcomes (for example, blood pressure control) and premature mortality for closely related conditions (diabetes, heart failure, hypertension, ischaemic heart disease, stroke, and chronic kidney disease, the total register for which was approximately 27% of the total number of registered patients in 2011-12).

Methods

Data sources

We accessed various data sources to obtain information on population estimates, mortality, socioeconomic status, ethnicity, health status, urbanicity versus rurality, quality of healthcare, and spatial coordinates. Data were collected at the lowest available geographic level, the "lower layer super output area" (LSOA), which has a mean (and median) population size of around 1500 (first centile 1084, 99th centile 2931 residents), since higher levels would involve aggregation across heterogeneous areas and populations and obscure inferential analyses.²¹ An average primary care practice covers a population approximately equivalent to five LSOAs. Mortality counts at the 2001 LSOA by age, sex, and cause were obtained from the Office of National Statistics. Datasets for practice-level QOF achievement, prevalence rates, and list sizes were downloaded from the Health & Social Care Information Centre.²² Population counts by gender and age group, as well as information on disability, ethnicity, and health status were obtained from the 2011 census, at the 2011 LSOA.²³⁻²⁴ Revised population estimates for prior years (based on 2001 and 2011 census information) were also downloaded.²⁵ Digital vector boundaries for the 2001 LSOAs, generalised to 20 metres and clipped to the coastline to reduce size and improve visualisation, were obtained from the Office of National Statistics open geography portal.²⁶ Area deprivation, as measured by the 2010 Index of

Multiple Deprivation,²⁷ and urban classification²⁸ were available at the 2001 LSOA level.²⁹⁻³¹

By definition, LSOAs are geographical areas mainly determined by population size and were first developed using data from the 2001 census. After the 2011 census some changes were introduced to reflect large population increases or reductions in localities, although 97.5% of LSOAs in England and Wales remained unchanged.²¹ Since most of the collected information was reported at the 2001 LSOA level, we proceeded to merge and analyse data using these geographies. Census data for 2011 were reported only for 2011 LSOAs, and we had to convert to the 2001 geographies. For the small percentage of areas that had merged or split, we used population weights to estimate data, and for some of these (approximately 1% of all LSOAs) population estimates were outside the 1000-3000 range, as would be expected.

Standardised mortality and final dataset

To allow for more meaningful comparisons across areas in terms of mortality, standardised mortality rates were calculated, using the 2011 census population data and all-cause and condition-specific mortality counts.³² Rather than standardising by the 1976 European Standard Population, we used 2011 census data to estimate the true age and sex population breakdown for England, which allowed us to produce yearly age and sex standardised mortality rates, from 2005 to 2012, per 100 000 people within each gender. Death among some age-sex subgroups can be a rare outcome at the LSOA level, so we also calculated two-year rates over the same time period.

The final dataset contained the following variables used in the regression analyses, complete for the 32 482 English LSOAs of 2001: all-cause one-year deaths and standardised mortality rates (yearly, 2005-12), all-cause two-year deaths and standardised mortality rates (yearly, 2005-11), condition-specific (diabetes, heart failure, hypertension, ischaemic heart disease, stroke, and chronic kidney disease; see online appendix 1 table A1 for details) one-year deaths and standardised mortality rates (yearly, 2005-12), condition-specific two-year deaths and standardised mortality rates (yearly, 2005-11), total population (yearly, 2005-12), percentage of whites (2011), non-urban (2004), index of multiple deprivation (2010). The dataset also included variables that were not used in the analyses, mainly due to collinearity, but were explored descriptively: percentage of population with day-to-day activities limited a lot or a little (2011), and percentage reporting bad or very bad health (2011). Using the Stata *shp2dta* command and the vector boundaries, we calculated the centroid for each LSOA in the British National Grid format.³³ These were converted from British National Grid easting and northing to longitude and latitude in degrees.³⁴ The manual conversion process was double checked for correctness using R, where a relevant command was available.³⁵

Practice-level data were aggregated at the 2001 LSOA level, to account for multiple practices in a locality, and the number of these primary care practice "hubs" (a LSOA with one or more practices) varied between 6569 in the third year of the QOF (2006-07) and 6413 in the eighth (2011-12). Our decision to exclude the first two years of the scheme was driven by practicality: first, not all required information was released for 2004-05 (for example, the number of exceptions (patients for whom indicators were deemed inappropriate by practices)); second, the scheme underwent a major revision after 2005-06 with many disease domains added and others considerably modified. Available information from 2006-07 to 2011-12

included number of practices, total list size, and prevalence rates for 19 QOF clinical domains (atrial fibrillation, asthma, hypertension, cancer, coronary heart disease, chronic kidney disease, chronic obstructive pulmonary disease, dementia, depression, diabetes, epilepsy, heart failure, learning disabilities, left ventricular dysfunction, mental health, obesity, palliative care, stroke, and hypothyroidism).

To better capture true levels of QOF quality of care we calculated population achievement,^{36 37} defined as $PA = (\sum N_i) / (\sum (D_i + E_i))$, where N_i , D_i and E_i the numerator, denominator and exceptions for QOF indicator i respectively. Three population achievement measures were calculated, within each QOF year: (a) overall, across all indicators (PA_{oval}); (b) across all intermediate outcome indicators (PA_{outc}); and (c) across a subset of nine intermediate outcome indicators in the clinical domains linked with condition-specific mortality (PA_{outc} ; blood pressure control in hypertension, coronary heart disease, chronic kidney disease, diabetes, and stroke; cholesterol control in coronary heart disease, diabetes, and stroke; HbA_{1c} control in diabetes). Details on all the indicators used, over time, are provided in online appendix 1 table A2. To quantify the morbidity “burden” in each practice hub, we defined morbidity load as the sum of QOF clinical domain denominators over the total list size for: (a) all indicators in all 19 domains (ML_{tot}); and (b) the nine outcome indicators in the five clinical domains linked with condition-specific mortality (ML_9).

Spatial weighted estimates datasets

Approximately 80% of LSOAs do not contain a general practice, so we used spatial analysis methods to estimate healthcare data at the LSOA level. Current UK policy only allows patients to register with primary care services in the locality they reside, and general practice choice is therefore relatively limited. However, that limitation allows us to use practice-level information and attribute it to geographical areas close to it. The LSOA centroid coordinates (longitude and latitude) were inputted to the *spmat* command in Stata and a 32 482×32 482 matrix of inverse distance (in miles) was obtained.³⁸ The matrix, which holds a detailed distance mapping of each LSOA with all other LSOAs, allowed us to quantify geographical connectivity and proximity, with closer areas having larger values (or weights). This was then used to generate the prevalence, quality of care, and morbidity load measures reported previously, for all LSOAs from 2006-07 to 2011-12. Two general approaches, under different assumptions, were used for this purpose.

Complete local attendance

Under this approach, for practice-hub LSOAs (approximately 20% containing one or more practices), we assumed that the population in each attended the local practice or practices and therefore practice-level aggregates were directly applied to the population level. For each of the remaining LSOAs not containing a general practice (approximately 80%), weighted estimation followed these steps: (a) the product P_{LD} of total list size and (centroid) inverse distance from the area was calculated and ranked for all practice-hub LSOAs; (b) the five “closest” practice-hubs (in terms of size and proximity) were selected; (c) each measure was calculated as the mean of the respective measures in the five selected practice-hubs weighted by P_{LD} . In the few cases (<50 in each year) where the area population N_A in a practice-hub was smaller than the total list size N_L , we used the same weighted estimation process to account for $N_L - N_A$ while N_A was accounted for by the local practice(s).

2014 attribution dataset

The complete local attendance assumption can be difficult to justify for all the patients in all areas (especially in high density urban areas where practice options are numerous), and we used an alternative approach to generate the weighted estimates. In 2014, the Health & Social Care Information Centre released information on the attribution of general practice populations to LSOAs and vice versa, completely linking practice registers and LSOAs for the first time.³⁹ Although the dataset covered only 2014, and some changes in registration practice occurred in 2012,⁴⁰ we used it as a blueprint to generate annual attribution datasets starting from 2011-12 and moving back to 2006-07. Of the 8122 practices in our analyses, 7932 were identified in the attribution dataset (97.7%), 190 had closed down or merged by 2014, while 77 new practices had emerged. The algorithm used to quantify quality of care and morbidity load proceeded along these steps, for each LSOA:

- 1) If two or more practices were linked to the LSOA population in 2014, Poisson and negative binomial regression models were fitted to the data, with list size and distance to the practice as predictors. The best modelling strategy between the two was identified by comparing the standard errors of the predictions.
- 2) For practices present both in our analyses and the attribution dataset, attributed population over time was adjusted for the practice’s list size in the respective year (that is, assumed a constant attribution rate over time).
- 3) For practices present in our analyses but not in the attribution dataset, we generated estimates using the model selected in step 1 (that is, predictions based on list size and distance) across all years.

Some LSOAs were served completely by a single practice, and we assumed that this remained the case in previous years. It should be noted that the 77 newer practices are not included in the attribution estimates for 2006-07 to 2011-12 and are only used to model the 2014 attribution in the area. The patients they served in 2014 were effectively re-distributed to the active practices within each year according to the selected regression model. Similarly, for practices that ceased to exist or merged by 2014, patients were re-distributed to them in their years of activity, according to their characteristics. Finally, the estimated attribution counts across practices and within each year were used to generate the weighted mean estimates for quality of care and morbidity load (and the potential differences between total estimated attribution size and total LSOA population, due to the appearance or disappearance of practices, becomes irrelevant).

Statistical analyses

Three sets of multiple linear regression models were used to investigate the relationship between QOF quality of care and mortality. Following spatial weighted estimation, data were complete for all 32 482 English 2001 LSOAs. Each analysis set was applied to both spatial weighted estimation approaches. The first set of models examined the relationship between QOF quality of care and 2011-12 standardised mortality rates, all-cause and condition-specific. As explained previously, the two-year estimates were considered more reliable because of the low incidence of deaths for younger subgroups. We chose 2011 as our population baseline from which to calculate mortality rates, since that was the census year and population counts were more reliable. Different models were used for each of the three population achievement measures (PA_{oval} , PA_{outc} ,

PA_{outx}) and for each QOF year from 2008-09 to 2011-12. All-cause standardised mortality rates were regressed with PA_{oval} and PA_{outc} , condition-specific standardised mortality rates with PA_{outx} . In all models we controlled for 2010 deprivation, urbanity versus rurality, ethnicity, and morbidity load (ML_{tot} for PA_{oval} , PA_{outc} ; ML_9 for PA_{outx}).

A second set of models investigated the relationship between changes in QOF quality of care over a three or five year period and 2011-12 standardised mortality rates. Changes in the three population achievement measures were used as predictors in separate regression models. For each measure we constructed and analysed change between: 2011-12 and 2008-09, 2010-11 and 2007-08, and 2009-10 and 2006-07 (three years); and between 2011-12 and 2006-07 (five years). Models were controlled for the same covariates as in the first set, and all-cause standardised mortality rates were regressed with PA_{oval} and PA_{outc} changes, while condition-specific standardised mortality rates were regressed with PA_{outx} changes.

A third set of models were used as sensitivity analyses, to assess the relationship between QOF quality of care and mortality over time (longitudinally). The outcome was one-year standardised mortality rates from 2007 to 2012, and we investigated the relationship for each of the three population achievement measures, in the same or future years (up to lag of three years). This zero lag model linked population achievement to mortality in the same year (such as QOF performance in 2011-12 to standardised mortality rate in 2012) and included data for all six years, while at the other end the three-year lag model necessarily used a subset of the data and standardised mortality rates in years 2010 to 2012 (linked to population achievement from 2006-07 to 2008-09). The same covariates were used as before, with the addition of year. These analyses addressed the same question as the first set, longitudinally rather than cross sectionally and with more data, but possibly a less reliable outcome (standardised mortality rates based on between-census population predictions) and were used as sensitivity analyses. Stata v13.1⁴¹ was used for data management and all analyses, with an alpha value of 1% to better protect against type I error inflation because of the numerous tests we performed. However, we focused on effect sizes rather than P values since statistical significance is more likely and can be less meaningful in large datasets such as the one we analysed.⁴² Analyses sets 1 and 2 were run with the *regress* command since their cross sectional nature did not require longitudinal random-effects modelling. For analysis set 3, however, we used the *xtreg* command, to account for the correlated nature of observations within each LSOA, over time (that is, by fitting a random effect for LSOAs). Although we collected more potential covariates (such as reported health) they tended to be strongly correlated with deprivation and were excluded from analyses because of collinearity. The distributions of 1-year and 2-year standardised mortality rates are linked to a Poisson process,⁴³ and we observed skew-normal distributions, although not extreme. Linear regressions are known to be robust in such scenarios of small or moderate skew, especially for large sample sizes,⁴⁴ but we performed additional sensitivity analyses with the log transformed outcomes.

Moreover, we repeated the three sets of analyses across each of the two spatial datasets using maximum likelihood spatial regressions with the *spre* command,⁴⁵ to account for potential spatial auto-correlation (that is, mortality in one LSOA affecting mortality in neighbouring LSOAs). An additional sensitivity analysis focussed on standardised mortality rates for those aged 50 or over, to assess whether any potential relationships, which

we would expect mainly for this group, were attenuated in the main analyses due to “noise” from younger, healthier people.

More sensitivity analyses assessed the sensitivity of the findings to the spatial weighted estimation assumptions. Further regressions with untransformed standardised mortality rates were run, using data from practice-hub LSOAs only. Finally three more sets of sensitivity analyses were executed for the complete local attendance approach, with a three practice-hub estimation limit, a 10 practice-hub estimation limit, and no limit at all, to assess how sensitive results were to the five practice-hub weighted estimation limit.

Results

We present results from the complete local attendance approach in the main paper and results from the attribution approach in online appendix 2. We mainly discuss the former but also highlight different results between the two methods and the other sensitivity analyses.

Descriptive statistics on mortality, demographics, reported health, deprivation, prevalence rates, QOF morbidity load, and quality of care (population achievement) are provided in table 1[¶] and online appendix 1 table A3, over geographical areas and time respectively. The spatial weighted estimation method is summarily described, in parallel with an example, in figure 1[¶]. Spatial maps for morbidity load and population achievement on outcome indicators are provided in figures 2[¶], 3[¶], and 4[¶] for the largest population centres (Greater London, West Midlands (Birmingham), and Greater Manchester).⁴⁶ Morbidity load represents the proportion of the population with the relevant chronic diseases, and therefore reflects demographic characteristics of resident populations (for example, age distribution) in addition to premature morbidity.

The North East and North West regions had the highest median all-cause death and standardised mortality rates (SMRs). Crude condition-specific death rates were more uniform across England, with the exception of London and its younger population, where rates were 33% below the national average in 2010-11. Condition-specific standardised mortality rates where highest for the North West and Yorkshire & Humber regions. Overall, median population QOF achievement varied from 83.6% (South East region) to 85.2% (North East). Population achievement for the subgroup of nine outcome indicators was lower and varied from 71.7% (London) to 74.8% (North East). Median morbidity load was highest in the North East and lowest in London, both for all indicators and for the nine outcome indicator subgroup.

Over time, all mortality metrics seemed to decrease, especially condition-specific standardised mortality rates, but late registration of Office of National Statistics deaths contributed to this (the probability of a death being “missing” is not time-invariant). Nevertheless, decreases in condition-specific deaths and standardised mortality rates and all-cause standardised mortality rates can be observed in the 2007-09 period, for which we would expect few unregistered deaths. The size of populations served by a practice-hub increased over time. Median recorded prevalence for most conditions and overall morbidity load (ML_{tot}) increased over time, but the subgroup morbidity load (ML_9) remained stable. Prevalence and performance rates estimated under the attribution approach were similar (online appendix 2, tables B1 and B2 and figs B1, B2, and B3).

Regression analyses

Levels of QOF quality of care and 2011-12 standardised mortality rates

All regression analyses in the first set of models failed to identify a significant relationship between quality of care (population achievement) and all-cause or condition-specific standardised mortality rates (table 2⇓). Area deprivation in 2010, as measured by the Index of Multiple Deprivation, was by far the strongest predictor of mortality across all models, with higher deprivation linked to higher standardised mortality rates in 2011-12. To put the effects into context, using QOF year 8 models, a change in the Index of Multiple Deprivation from the median (17.25) to the 90th centile (44.88) would correspond to an all-cause standardised mortality rate increase from the median to the 82nd centile. For the condition-specific standardised mortality rate models the effect was weaker, and a change in Index of Multiple Deprivation from the median to the 90th centile would correspond to a condition-specific standardised mortality rate increase from the median to the 73rd centile. In non-urban areas, all-cause standardised mortality rates were estimated to be lower, but not condition-specific standardised mortality rates. A non-urban LSOA (18.6%) would correspond to an all-cause standardised mortality rate change of -44.5, compared with an urban LSOA, which would take the median all-cause standardised mortality rate from 513.4 to 468.9 (40th centile). Ethnicity was also a significant but less strong predictor in all models, with larger standardised mortality rates observed in areas with larger proportions of non-white residents.

In the models where morbidity load across all 19 QOF domains (ML_{tot}) was used as a predictor, it was found to be negatively associated with standardised mortality rates, while a positive association with standardised mortality rates was observed for the subgroup morbidity load (ML_g) in the condition-specific mortality models. However, these relationships were relatively weak, with a change in ML_{tot} from the 50th centile (1.94) to the 90th (2.30) corresponding to an all-cause standardised mortality rate move from the median to the 44th centile, and a change in ML_g from the 50th centile (0.41) to the 90th (0.50) corresponding to a condition-specific standardised mortality rate move from the median to the 52nd centile. The models explained approximately 29% of the all-cause and around 8% of the variance in condition-specific standardised mortality rates.

Changes in levels of QOF quality of care and 2011-12 standardised mortality rates

Findings were similar in the second set of analyses, where changes in population achievement (over three or five years) were not found to be significantly or modestly associated with all-cause or condition-specific standardised mortality rates (table 3⇓). The only exception was the relationship between three year change in overall achievement (from year 5 to 8) and all-cause mortality, where a small but statistically significant effect was observed. Estimates were close to the first set of analyses, and area deprivation was again the strongest predictor. Urbanicity versus rurality and ethnicity were weaker predictors. Similar relationships to the first set of analyses were observed between morbidity load and standardised mortality rate: negative for ML_{tot} and positive for ML_g .

Attribution dataset analyses

Results from the two spatial approaches generally agreed across all analyses. Across each attribution model, the rate of explained variance (R^2) was very close to the explained variance in the respective complete local attendance model. Deprivation was

again the strongest predictor by far, across all models and the relationship between QOF population achievement and outcomes was generally negligible (online appendix 2, tables B3, B4, and B5). There were a few analyses where the relationship seemed to be statistically significant (high performance linked to reduction in mortality, but the reverse in some cases). However, these relationships were not verified in the statistically more robust longitudinal random-effects models (third set), where all relationships between population achievement and outcomes were negligible.

Sensitivity analyses

In all sensitivity analyses (third set of analyses: longitudinal model with yearly standardised mortality rates from 2007 to 2012; log-transformed standardised mortality rates; practice-hub only and varying assumptions for the practice-hub estimation limit; spatial regressions and standardised mortality rates for those aged ≥ 50 years) findings were similar (third set of analyses reported in table 4⇓ and table B5 in online appendix 2, for the two spatial approaches; other results available from the authors). In almost all models and analyses the relationship between QOF population achievement and all-cause or condition-specific standardised mortality rates was negligible or, in a few cases, small but statistically significant (and often in the opposite direction to the expected). This small number of statistically significant findings is not unusual given the number of models executed and the size of the dataset, and we feel they can be discarded as spurious. Area deprivation was by far the strongest predictor in all sensitivity analyses. Spatial parameters estimated with *spreg* were often statistically significant but negligible when compared with the error terms.

Discussion

The Quality and Outcomes Framework (QOF) was introduced in 2004 across all family practices in the UK to reward high quality primary care, and over £10bn has since been invested in the programme. Although the incentive scheme does not fully capture quality in primary care, it does cover many important aspects, especially within its clinical domain. Between the third and eighth years of the QOF programme mortality rates in England decreased by 14%, but we could not identify a relationship between practice performance on the clinical aspect of the QOF and mortality outcomes in the practice locality.

Strengths and limitations of the study

The main strength of this analysis is that it was conducted at the population level and used two different novel spatial estimation techniques to analyse data for the whole of the primary care population in England.

However, there are several limitations. First, standardised mortality rates are imprecise at the level of small geographical areas such as lower layer super output areas (LSOAs). However, we felt that analysis at this level was required to meaningfully examine the relationship between recorded quality of care and patient outcomes. Aggregating data at higher levels introduces heterogeneity and potential confounding that would be very difficult to account for. In addition, use of standardised mortality rates should not pose a threat to the validity of our findings: even with an imprecise measure, we should have identified any existing modest link between performance and mortality. Second, because of late registration of some deaths by the Office of National Statistics, we expect the rates reported in the later years to be slightly underestimated. However, we observed a reduction in standardised mortality rates between 2007 and 2010

(although less pronounced) and we would not expect a relationship between late death registration and QOF quality of care that would bias the regression analyses results. Third, the index of multiple deprivation was collinear with various measures of poor health and deprivation collected in the 2011 census—unsurprisingly since it is an aggregate of income, employment, health, education, housing, crime, and environmental deprivation in the locality. We excluded census measures from the analyses because the coefficients were uninterpretable, although they explained some additional variability in the outcomes. Fourth, not everyone resident in a locality at time of death is registered with a participating general practice. However, over 95% of the UK population is estimated to be registered with a practice, and over 99% of registered patients attend practices participating in the QOF.⁴⁷ Fifth, we were unable to attribute practice performance to LSOAs for the first two years of the QOF (2004/5 and 2005/6), the period when practices made the greatest improvements in performance.⁴⁸ Sixth, the effect of the scheme might be delayed and the three year lag between care and outcome could be too narrow a time window. However, analyses with larger lag periods are prone to methodological problems, such as in accounting for migration.

In addition, we made numerous spatial weighted estimation assumptions. Under our first approach, we modelled patients as attending practices in their locality, which we considered a realistic assumption given existing restrictions on registration. Although patient dissatisfaction with access to primary care during work hours was identified as early as 2010,⁴⁹ policy has not yet changed to allow registration with a practice close to work rather than close to home (a trial scheme was launched in April 2012,⁴⁰ after the end of our study period). However, we had to assume that patients will always register with the practice(s) in their residing LSOA, provided one (or more) exists, an assumption that might be difficult to justify in urban areas where there can be a few options within walking distance. According to the 2014 attribution dataset, that was the case for around 60% of patients and not all, although we would expect that rate to be higher before the introduction of the 2012 trial scheme. If we assume that the 60% figure applies to all the years we analysed, a considerable amount of noise would have been introduced in the spatial estimates, implying that we would be unable to detect weak relationships between them and the outcome. Other characteristics of this approach included an estimation limit of the five “closest” practice-hubs, which, although arbitrary, we argue is reasonable; it might be too limiting to have fewer than five in areas surrounded by practice-hubs, and it is unrealistic for practices in one city to affect levels of care in another. In assigning weights to practice-hubs, we decided that ranking on the product of list size and proximity was appropriate as this limits the effect of the few very small practices. We repeated the analyses using a proximity ranking strategy and obtained almost identical results. Under our second approach, we used the 2014 attribution dataset to calculate the contribution of primary care practices and estimated the quality of care provided to the population of each locality. We assumed these contribution rates remained constant over time, and we used regression models to estimate the contribution of practices that had closed down or merged by 2014—assumptions and estimates that are not infallible. However, we arrived at similar results and conclusions from two very different starting points.

Under both spatial approaches, there is uncertainty in the estimates which we could not include in the models because of methodological and software limitations. However, we conducted numerous sensitivity analyses under different

assumptions to address this issue. Overall, these limitations and assumptions might have attenuated the relationship between recorded quality of care and mortality, but we would expect a relationship of reasonable strength to have been detectable, as was the case for estimated morbidity burden.

The spatial weighted estimation approach is necessary when access to the Primary Care Mortality Database is not available,⁵⁰ and it allows the analyses to be accurately controlled for the two strongest predictors of mortality (urbanicity v rurality and area deprivation in the patient locality) and other census-measured covariates, while accounting for spatial auto-correlation. An analysis at the practice level would have to use practice location proxies for deprivation and urbanicity v rurality, and census measured covariates would be unavailable, but would not have to make any spatial estimation assumptions for quality of care and the morbidity burden. We aim to verify our findings with the practice-level analyses if and then the data become available to researchers.

Although, from a statistical point of view, spatial regressions are more suitable for our data, we decided to use outputs from standard regression models as our main results for several reasons: (1) not all spatial regression models converged; (2) the auto-correlation estimates, although statistically significant in most cases, were extremely low compared with the error term (which is reasonable: deaths in an LSOA are generally not “spatially” linked to deaths in neighbouring LSOAs); (3) each model takes days or weeks to run in a dedicated modern server; (4) findings across all analyses, including spatial regressions, were almost identical; and (5), with the standard regression models, we could report easily interpretable model-fit measures (such as R^2) and avoid further complicating the paper with spatial parameters and their interpretation. We also considered Poisson regression analyses with death counts as the outcome, but that approach was more problematic than linear regressions with standardised mortality rates, since collinearity issues did not allow us to accurately control the analyses for the distribution of the population within each LSOA.

Finally, this is an observational study and the potential for unmeasured confounding is always a possibility. For example, secondary care could be consistently very good where primary care is very poor and vice versa; thus the aggregate population effect from the two levels would be relatively constant, and we would fail to observe an effect for primary care. Although such a scenario is unlikely, the effect of high quality primary care on mortality will be attenuated by the (unknown) quality of secondary care in the LSOA.

Findings

We found that overall quality of care provided by practices—as measured by achievement across all clinical QOF indicators—was not associated with mortality rates in their localities for conditions covered by the QOF. There remained no association when potential effects were lagged for up to three years. This finding seems to contradict previous evidence that processes incentivised in the QOF increase longevity,²⁰ but there are several possible explanations for this null effect. For most QOF indicators, levels of performance improved at the fastest rate over the first two years of the programme and were generally high from the third year of the scheme onwards, with relatively little variation between practices. It is therefore possible that we failed to detect a global mortality dividend predominantly gained in the first two years of the programme. It is also possible that performance has been exaggerated by practices responding to financial incentives and that actual levels

of achievement are lower than reported. However, when we examined the associations between performance on intermediate outcomes indicators for specific diseases and related mortality we also found no effect. This applied both to indicators measured by practices and to indicators measured by third parties (for example HbA_{1c} levels). Variation in practice performance is also substantial for most intermediate outcomes indicators (for example, the interquartile range for HbA_{1c} ≤7.5% was 57.3% to 68.2% in 2011-12), and we would therefore expect to detect associations between performance on these indicators and mortality.

It might be the case that the indicators of the scheme need to be reconsidered and better aligned with existing evidence. For example, clinical trial findings indicate that intensive glucose control is associated with increased mortality,⁵¹ especially risk of cardiovascular death in younger patients,⁵² while observational studies have generally demonstrated U-shaped relationships between levels of HbA_{1c} in diabetic patients and death.⁵³⁻⁵⁵ Similar U-shaped relationships have been observed for other biometric measurements, including blood pressure and total cholesterol levels.^{55, 56} These non-linear patterns might suggest that target values (such as ≤7.5 mm Hg for HbA_{1c} in 2011-12) are suboptimal measures of high quality of care and that target ranges might be more suitable.

Our regression models suggest that area characteristics, such as material deprivation and urbanicity versus rurality, or unmeasured factors associated with these characteristics have a greater impact on mortality than variations in quality of care provided by general practices. However, our models explained relatively little variation in mortality rates. Given that mortality rates fell substantially for several QOF conditions during the course of the programme, in particular cerebrovascular disease and coronary heart disease, this suggests that improvements were due to factors outside primary care, or at least not incentivised under the QOF. For coronary heart disease, previous studies suggest that the main drivers behind mortality reductions have been population improvements in risk factors, mainly declining rates of smoking.⁵⁷ While the QOF included indicators relating to several risk factors, many of these incentivized processes (for example, recording smoking status and offering cessation support) rather than improvements in outcomes (such as reducing smoking rates).

Conclusions

We found that perceived improvements in performance incentivised under a nationwide pay-for-performance programme were not associated with a subsequent reduction in premature mortality rates. This suggests that the impact of the incentive scheme has fallen far short of previous estimates,²⁰ although it is possible that there have been significant population benefits in terms of reduced morbidity incidence or improved quality of life, and that longer term mortality reductions will ultimately accrue. The apparent lack of large effect on mortality over the medium term may suggest that the QOF may not have been an optimal investment of health service resources, but mortality rates need to be further investigated within primary care practices, accounting for the quality of local secondary care services.⁵⁸ If incentive schemes continue to be used in primary care with the intention of improving population outcomes, indicators will need to be reconsidered and better aligned with evidence on which activities contribute to reduction of premature mortality.

We thank the Health and Social Care Information Centre and the Office of National Statistics for the wealth of information they have collected and systematically organised, which made this study possible.

Contributors: EK and TD designed the study. EK extracted the data from all sources and performed the analyses, while some sensitivity analyses were performed by DAS. EK and TD wrote the manuscript. DAS, MA, RTW, and IEB critically edited the manuscript. EK is guarantor of this work and had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Competing interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: EK was partly supported by a NIHR School for Primary Care Research fellowship in primary healthcare; TD was supported by a NIHR Career Development Fellowship. The views expressed are those of the authors and not necessarily those of the NHS, the National Institute for Health Research, or the Department of Health. No other relationships or activities could appear to have influenced the submitted work.

Transparency: EK affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Data sharing: Most of the data used in this study are freely available, and the authors are happy to share an organised and cleaned final dataset, except for the mortality data, which were obtained under a sharing agreement from the Health and Social Care Information Centre.

- Bellamy D, Smith J. Role of primary care in early diagnosis and effective management of COPD. *Int J Clin Pract* 2007;61:1380-9.
- Hung DY, Rundall TG, Tallia AF, Cohen DJ, Halpin HA, Crabtree BF. Rethinking prevention in primary care: applying the chronic care model to address health risk behaviors. *Milbank Q* 2007;85:69-91.
- Starfield B, Shi LY, Macinko J. Contribution of primary care to health systems and health. *Milbank Q* 2005;83:457-502.
- Engstrom S, Foldevi M, Borgquist L. Is general practice effective? A systematic literature review. *Scand J Prim Health* 2001;19:131-44.
- McGlynn EA, Asch SM, Adams J, Keesey J, Hicks J, DeCristofaro A, et al. The quality of health care delivered to adults in the United States. *N Engl J Med* 2003;348:2635-45.
- Campbell S, Steiner A, Robison J, Webb D, Raven A, Roland M. Is the quality of care in general medical practice improving? Results of a longitudinal observational study. *Br J Gen Pract* 2003;53:298-304.
- Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am Psychol* 2000;55:68-78.
- Le Grand J. *Motivation, agency, and public policy: of knights and knaves, pawns and queens*. Oxford University Press, 2006.
- May C, Finch T, Mair F, Ballini L, Dowrick C, Eccles M, et al. Understanding the implementation of complex interventions in health care: the normalization process model. *BMC Health Serv Res* 2007;7:1-7.
- 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.
- Flodgren G, Eccles MP, Shepperd S, Scott A, Parmelli E, Beyer FR. An overview of reviews evaluating the effectiveness of financial incentives in changing healthcare professional behaviours and patient outcomes. *Cochrane Database Syst Rev* 2011:CD009255.
- Scott A, Sivey P, Ait Ouakrim D, Willenberg L, Naccarella L, Furler J, et al. The effect of financial incentives on the quality of health care provided by primary care physicians. *Cochrane Database Syst Rev* 2011:CD008451.
- Roland M. Linking physicians' pay to the quality of care—a major experiment in the United Kingdom. *N Engl J Med* 2004;351:1448-54.
- 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.
- Doran T, Kontopantelis E, Valderas JM, Campbell S, Roland M, Salisbury C, 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:d3590.
- Kontopantelis E, Reeves D, Valderas JM, Campbell S, Doran T. Recorded quality of primary care for patients with diabetes in England before and after the introduction of a financial incentive scheme: a longitudinal observational study. *BMJ Qual Saf* 2013;22:53-64.
- Dusheiko M, Doran T, Gravelle H, Fullwood C, Roland M. Does higher quality of diabetes management in family practice reduce unplanned hospital admissions? *Health Serv Res* 2011;46:27-46.
- Purdy S, Griffin T, Salisbury C, Sharp D. Emergency respiratory admissions: influence of practice, population and hospital factors. *J Health Serv Res Policy* 2011;16:133-40.
- Bottle A, Gnani S, Saxena S, Aylin P, Mainous AG 3rd, Majeed A. Association between quality of primary care and hospitalization for coronary heart disease in England: national cross-sectional study. *J Gen Intern Med* 2008;23:135-41.
- Fleetcroft R, Parekh-Burke S, Howe A, Cookson R, Swift L, Steel N. The UK pay-for-performance programme in primary care: estimation of population mortality reduction. *British J Gen Pract* 2010;60:e345-52.

What is already known on this topic

The Quality and Outcomes Framework (QOF) is a UK pay-for-performance programme that has cost approximately £1bn per annum since 2004

Under the scheme, performance on numerous clinical indicators is incentivised, and achievement rates have increased—though partially due to pre-2004 trends

Evidence for improved patient outcomes seems contradictory, with decreased emergency hospital admission rates for diabetes but not for other incentivised conditions

No studies to date have examined the relationships between recorded practice performance under the QOF and death rates for conditions included in the programme

What this study adds

Material deprivation was by far the strongest predictor of mortality, and, when controlling for this and other covariates, we failed to observe a relationship between QOF achievement and mortality

This suggests that the impact of the incentive scheme has fallen short of previous estimates and indicators will need to be reconsidered and better aligned with evidence

- 21 Office for National Statistics. Changes to output areas and super output areas in England and Wales, 2001 to 2011. ONS, 2012. www.ons.gov.uk/ons/guide-method/geography/products/census/report-changes-to-output-areas-and-super-output-areas-in-england-and-wales--2001-to-2011.pdf
- 22 Health and Social Care Information Centre. The Quality and Outcomes Framework. 2013. www.ic.nhs.uk/qof.
- 23 Office for National Statistics. Super output area mid-year population estimates for England and Wales, mid-2011 (census based). 2013. www.ons.gov.uk/ons/publications/re-reference-tables.html?edition=tcm%3A77-285629.
- 24 Office for National Statistics. Nomis official labour market statistics. 2013. www.nomisweb.co.uk/census/2011/quick_statistics.
- 25 Office for National Statistics. Super output area mid-year population estimates for England and Wales (experimental), mid-2002 to mid-2010 revised. 2013. www.ons.gov.uk/ons/publications/re-reference-tables.html?edition=tcm%3A77-285154.
- 26 Office for National Statistics. Open Geography Portal. 2013. <https://geoportals.statistics.gov.uk/geoportals/>
- 27 Communities and Local Government. The English indices of deprivation 2010: technical report. Government DfCaL, 2011. www.communities.gov.uk/documents/statistics/pdf/1870718.pdf.
- 28 Bibby P, Shepherd J. Developing a new classification of urban and rural areas for policy purposes—the methodology. DEFRA, 2004. http://collection.eurparchive.org/ea/20070108123845/http://www.statistics.gov.uk/geography/downloads/Methodology_Report.pdf
- 29 Office for National Statistics. Neighbourhood statistics. 2013. www.neighbourhood.statistics.gov.uk/.
- 30 Communities and Local Government. Indices of deprivation 2010. 2013. www.communities.gov.uk/communities/research/indicesdeprivation/deprivation10/.
- 31 Edina. UKBORDERS: Boundary datasets and geographic look up tables of the United Kingdom. 2013. <http://edina.ac.uk/ukborders/>.
- 32 Office for National Statistics. Mortality statistics: metadata. ONS, 2013:55. www.ons.gov.uk/ons/guide-method/user-guidance/health-and-life-events/mortality-metadata.pdf.
- 33 SHP2DATA: Stata module to converts shape boundary files to Stata datasets [program]. S456718 version: Boston College Department of Economics, 2006.
- 34 Ordnance Survey. A guide to coordinate systems in Great Britain. ONS, 2013:43. www.ordnancesurvey.co.uk/docs/support/guide-coordinate-systems-great-britain.pdf.
- 35 R: A language and environment for statistical computing [program]. 3.0.3 version. R Foundation for Statistical Computing, 2014.
- 36 Kontopantelis E, Buchan I, Reeves D, Checkland K, Doran T. Relationship between quality of care and choice of clinical computing system: retrospective analysis of family practice performance under the UK's quality and outcomes framework. *BMJ Open* 2013;3.
- 37 Kontopantelis E, Springate D, Reeves D, Ashcroft DM, Valderas JM, Doran T. Withdrawing performance indicators: retrospective analysis of general practice performance under UK Quality and Outcomes Framework. *BMJ* 2014;348:g330.
- 38 Drukker DM, Peng H, Prucha IR, Raciborski R. Creating and managing spatial-weighting matrices with the `spmat` command. *Stata J* 2013;13:242-86.
- 39 Health and Social Care Information Centre. Numbers of Patients Registered at a GP Practice—January 2014. 2014. www.hscic.gov.uk/catalogue/PUB13365.
- 40 Boseley S. Patients can choose GP near work or school under trial scheme. *Guardian* 2011 Dec 30.
- 41 Stata Statistical Software for Windows [program]. 13.1 version, 2013.
- 42 Lin MF, Lucas HC, Shmueli G. Too big to fail: large samples and the p-value problem. *Inform Syst Res* 2013;24:906-17.
- 43 Liddell FD. Simple exact analysis of the standardised mortality ratio. *J Epidemiol Commun Health* 1984;38:85-8.
- 44 Li X, Wong W, Lamoureux EL, Wong TY. Are linear regression techniques appropriate for analysis when the dependent (outcome) variable is not normally distributed? *Invest Ophthalmol Vis Sci* 2012;53:3082-3.
- 45 Drukker DM, Prucha IR, Raciborski R. Maximum likelihood and generalized spatial two-stage least-squares estimators for a spatial-autoregressive model with spatial-autoregressive disturbances. *Stata J* 2013;13:221-41.
- 46 SPMPAP: Stata module to visualize spatial data [program]. S456812 version: Boston College Department of Economics, 2007.
- 47 Health and Social Care Information Centre. Quality and outcomes framework achievement, prevalence and exceptions data, 2012/13: Frequently asked questions. HSCIC, 2013:20. www.hscic.gov.uk/catalogue/PUB12262/qual-outc-fram-freq-asked-ques-2012-13-gui.pdf.
- 48 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.
- 49 Kontopantelis E, Roland M, Reeves D. Patient experience of access to primary care: identification of predictors in a national patient survey. *BMC Fam Pract* 2010;11.
- 50 Health and Social Care Information Centre. Primary care mortality database. 2014. www.hscic.gov.uk/pcmdatabase.
- 51 Action to Control Cardiovascular Risk in Diabetes Study Group, Gerstein HC, Miller ME, Byington RP, Goff DC Jr, Bigger JT, et al. Effects of intensive glucose lowering in type 2 diabetes. *N Engl J Med* 2008;358:2545-59.
- 52 Miller ME, Williamson JD, Gerstein HC, Byington RP, Cushman WC, Ginsberg HN, et al. Effects of randomization to intensive glucose control on adverse events, cardiovascular disease and mortality in older versus younger adults in the ACCORD trial. *Diabetes Care* 2014;37:634-43.
- 53 Currie CJ, Peters JR, Tynan A, Evans M, Heine RJ, Bracco OL, et al. Survival as a function of HbA(1c) in people with type 2 diabetes: a retrospective cohort study. *Lancet* 2010;375:481-9.
- 54 Andersson C, van Gaal L, Caterson ID, Weeke P, James WP, Coutinho W, et al. Relationship between HbA1c levels and risk of cardiovascular adverse outcomes and all-cause mortality in overweight and obese cardiovascular high-risk women and men with type 2 diabetes. *Diabetologia* 2012;55:2348-55.
- 55 Kontopantelis E, Springate D, Reeves D, Ashcroft D, Rutter M, Buchan I, et al. Glucose, blood pressure and cholesterol levels and their relationships to clinical outcomes in type 2 diabetes: a retrospective cohort study. *Diabetologia* 2015;58:505-18.
- 56 Vamos EP, Harris M, Millet C, Pape UJ, Khunti K, Curcin V, et al. Association of systolic and diastolic blood pressure and all cause mortality in people with newly diagnosed type 2 diabetes: retrospective cohort study. *BMJ* 2012;345:e5567.
- 57 Unal B, Critchley JA, Capewell S. Explaining the decline in coronary heart disease mortality in England and Wales between 1981 and 2000. *Circulation* 2004;109:1101-7.
- 58 Goodacre S, Campbell M, Carter A. What do hospital mortality rates tell us about quality of care? *Emerg Med J* 2013. doi:10.1136/emered-2013-203022.

Accepted: 13 January 2015

Cite this as: *BMJ* 2015;350:h904

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

RESEARCH

Tables

Table 1 | Area and population characteristics by strategic health authority*: Quality and Outcomes Framework (QOF) year 7 (2010-11); census data for 2011; deaths for 2011 and 2012

	Strategic health authority*									
	North East	North West	Yorkshire & Humber	East Midlands	West Midlands	East England	London	South East	South Central	South West Coast
Aggregates										
Deaths (%):										
All-cause (2011)	0.55	0.53	0.5	0.48	0.49	0.45	0.37	0.46	0.44	0.48
Condition-specific (2011)	0.19	0.2	0.2	0.19	0.19	0.19	0.12	0.19	0.16	0.2
All-cause (2011-12)	1.09	1.05	1.01	0.96	0.98	0.91	0.74	0.93	0.87	0.96
Condition-specific (2011-12)	0.39	0.4	0.4	0.39	0.38	0.38	0.24	0.37	0.33	0.4
Population size	2 596 445	7 055 961	5 288 218	4 537 448	5 608 670	5 862 419	8 204 407	4 475 798	4 176 986	5 300 881
No of practices	404	1267	808	633	971	791	1505	634	505	726
Census information: medians (IQRs) across 2001 LSOAs										
SMR:										
All-cause (2011)	572 (437-748)	577 (433-752)	541 (407-701)	500 (383-645)	523 (393-675)	461 (350-597)	553 (403-740)	459 (343-591)	480 (367-641)	451 (342-563)
Condition-specific (2011)	172 (97-282)	194 (107-298)	188 (110-287)	178 (102-268)	176 (96-280)	162 (88-250)	156 (72-269)	152 (86-232)	148 (73-245)	152 (87-247)
All-cause (2011-12)	574 (455-717)	580 (448-737)	541 (435-689)	508 (405-635)	528 (415-663)	466 (377-587)	563 (440-716)	456 (363-573)	482 (383-626)	448 (362-522)
Condition-specific (2011-12)	184 (122-259)	198 (133-281)	194 (132-272)	184 (124-257)	184 (123-259)	167 (111-237)	166 (101-253)	154 (104-219)	155 (103-229)	158 (108-222)
Limiting disability (%)	21.9 (18.5-25.7)	20.2 (16.9-23.9)	18.6 (15.5-22.3)	18.6 (15.4-22.1)	19.0 (16.2-21.9)	16.4 (13.5-19.8)	14.1 (11.9-16.5)	16.2 (13.0-19.8)	14.4 (11.9-17.6)	18.4 (15.3-21.5)
Bad or very bad health (%)	7.5 (5.2-9.7)	6.4 (4.6-8.7)	5.7 (4.1-7.7)	5.3 (3.9-6.9)	5.9 (4.4-7.7)	4.4 (3.3-5.9)	4.7 (3.7-6.0)	4.3 (3.2-5.8)	3.7 (2.8-5.0)	4.9 (3.7-6.4)
White ethnicity (%)	98.0 (95.9-98.9)	96.5 (92.1-98.1)	96.6 (90.6-98.2)	96.6 (89.7-98.0)	92.6 (79.1-97.3)	94.7 (89.0-97.3)	63.0 (45.8-77.5)	94.7 (90.3-97.0)	93.7 (86.5-96.8)	97.5 (95.1-98.8)
IMD 2010	24.3 (12.8-39.0)	21.8 (11.4-40.2)	19.9 (11.1-36.7)	15.9 (9.1-27.8)	20.0 (11.5-37.0)	12.8 (7.8-21.1)	23.8 (14.1-34.9)	12.9 (7.3-20.8)	10.3 (5.8-18.9)	14.9 (9.5-20.6)
QOF information: medians (IQRs) across practice-hubs										
List size	7899 (4457-11 857)	6561 (3729-11 016)	7502 (4193-11 106)	7657 (4407-11 413)	6567 (3809-10 337)	7835 (4523-11 892)	6305 (3857-10 145)	7941 (4518-11 829)	9304 (5791-12 800)	7886 (4877-11 288)
QOF information (spatially estimated): medians (IQRs) across 2001 LSOAs										
Atrial fibrillation	1.5 (1.4-1.7)	1.5 (1.2-1.7)	1.5 (1.2-1.7)	1.5 (1.2-1.7)	1.5 (1.3-1.7)	1.5 (1.3-1.7)	0.8 (0.6-1.1)	1.6 (1.4-1.9)	1.3 (1.1-1.5)	1.7 (1.5-2.0)
Asthma	6.1 (5.8-6.5)	6.3 (5.9-6.7)	6.1 (5.7-6.6)	6.1 (5.7-6.5)	6.2 (5.8-6.5)	6.2 (5.7-6.7)	4.8 (4.3-5.2)	5.6 (5.3-6.0)	6.0 (5.5-6.4)	6.4 (6.0-6.7)
Hypertension	14.9 (13.5-16.1)	14.0 (12.9-15.0)	13.8 (12.0-15.1)	13.9 (12.6-15.0)	14.7 (13.5-16.2)	13.6 (12.5-14.9)	11.1 (9.3-12.7)	13.4 (12.3-15.1)	12.2 (11.1-13.4)	14.1 (12.7-15.5)
Cancer	1.6 (1.4-1.8)	1.6 (1.4-1.8)	1.5 (1.3-1.8)	1.6 (1.3-1.7)	1.7 (1.4-1.9)	1.6 (1.4-1.9)	1.1 (0.9-1.4)	1.7 (1.5-2.0)	1.5 (1.3-1.8)	1.9 (1.6-2.1)
Coronary heart disease	4.5 (4.0-4.9)	4.0 (3.6-4.5)	4.0 (3.5-4.4)	3.5 (3.1-4.1)	3.6 (3.2-4.0)	3.2 (2.9-3.7)	2.1 (1.7-2.7)	3.0 (2.7-3.6)	2.8 (2.5-3.2)	3.5 (3.0-4.0)
Chronic kidney disease (age ≥18)	3.7 (3.1-4.4)	3.5 (2.9-4.1)	3.4 (2.9-4.2)	3.7 (3.1-4.8)	3.5 (3.0-4.2)	3.1 (2.7-3.8)	2.1 (1.5-2.7)	3.4 (3.0-3.9)	3.0 (2.4-3.5)	3.8 (3.1-4.4)
COPD	2.4 (2.1-2.8)	2.1 (1.8-2.4)	1.9 (1.6-2.2)	1.7 (1.3-2.0)	1.6 (1.4-1.9)	1.5 (1.3-1.8)	1.0 (0.8-1.2)	1.4 (1.2-1.7)	1.2 (1.0-1.5)	1.6 (1.4-1.8)
Dementia	0.5 (0.4-0.6)	0.5 (0.4-0.6)	0.5 (0.4-0.6)	0.5 (0.4-0.6)	0.5 (0.4-0.5)	0.5 (0.4-0.6)	0.3 (0.2-0.4)	0.5 (0.4-0.6)	0.4 (0.3-0.6)	0.5 (0.4-0.6)
Depression (age ≥18)	11.9 (9.9-13.8)	10.4 (8.5-11.8)	9.0 (6.8-10.8)	9.2 (7.4-11.1)	9.0 (7.3-10.7)	8.8 (6.8-10.7)	6.3 (5.1-7.6)	8.7 (7.2-10.2)	8.8 (7.4-10.3)	9.8 (8.5-11.1)
Diabetes (age ≥17)	4.6 (4.2-5.0)	4.7 (4.3-5.1)	4.5 (3.9-5.0)	4.6 (4.1-5.0)	4.8 (4.4-5.4)	4.2 (3.9-4.8)	4.2 (3.6-5.0)	4.1 (3.7-4.7)	3.9 (3.4-4.4)	4.3 (3.9-4.6)
Epilepsy (age ≥18)	0.7 (0.7-0.8)	0.7 (0.6-0.8)	0.7 (0.6-0.8)	0.6 (0.6-0.7)	0.7 (0.6-0.8)	0.6 (0.5-0.7)	0.5 (0.4-0.5)	0.6 (0.5-0.7)	0.6 (0.5-0.6)	0.7 (0.6-0.7)
Heart failure	0.8 (0.7-1.0)	0.8 (0.7-0.9)	0.7 (0.6-0.8)	0.8 (0.7-0.9)	0.8 (0.7-0.9)	0.7 (0.6-0.8)	0.5 (0.4-0.6)	0.6 (0.5-0.7)	0.6 (0.5-0.7)	0.8 (0.7-0.9)

BMJ: first published as 10.1136/bmj.h904 on 2 March 2015. Downloaded from <http://www.bmj.com/> on 23 April 2015 by guest. Protected by copyright.

Table 1 (continued)

	Strategic health authority*									
	North East	North West	Yorkshire & Humber	East Midlands	West Midlands	East England	London	South East	South Central	South West Coast
Learning disabilities (age ≥18)	0.4 (0.4–0.5)	0.4 (0.3–0.4)	0.4 (0.3–0.5)	0.4 (0.3–0.5)	0.3 (0.3–0.4)	0.3 (0.3–0.4)	0.2 (0.2–0.3)	0.3 (0.2–0.4)	0.3 (0.2–0.3)	0.4 (0.3–0.5)
Left ventricular dysfunction	0.5 (0.4–0.6)	0.4 (0.4–0.5)	0.4 (0.3–0.5)	0.4 (0.3–0.5)	0.4 (0.3–0.5)	0.4 (0.3–0.4)	0.2 (0.2–0.3)	0.3 (0.2–0.4)	0.3 (0.2–0.4)	0.4 (0.4–0.5)
Mental health	0.8 (0.7–0.9)	0.9 (0.7–1.0)	0.7 (0.6–0.8)	0.7 (0.6–0.8)	0.7 (0.6–0.9)	0.7 (0.6–0.8)	0.9 (0.8–1.2)	0.7 (0.6–0.8)	0.7 (0.6–0.8)	0.7 (0.6–0.9)
Obesity (age ≥16)	10.6 (9.2–11.7)	9.3 (8.2–10.5)	9.0 (7.4–10.7)	8.7 (7.5–10.1)	9.5 (8.1–11.0)	8.2 (6.9–9.6)	7.4 (6.2–8.9)	7.1 (6.0–8.6)	7.3 (6.4–8.4)	8.1 (7.0–9.2)
Palliative care	0.2 (0.1–0.2)	0.2 (0.1–0.2)	0.2 (0.1–0.2)	0.2 (0.1–0.3)	0.2 (0.1–0.2)	0.2 (0.1–0.2)	0.1 (0.1–0.2)	0.1 (0.1–0.2)	0.1 (0.1–0.2)	0.1 (0.1–0.2)
Stroke	2.1 (1.9–2.3)	1.9 (1.7–2.1)	1.9 (1.7–2.1)	1.7 (1.5–2.0)	1.8 (1.6–2.0)	1.6 (1.5–1.9)	1.1 (0.8–1.3)	1.7 (1.5–2.0)	1.5 (1.3–1.7)	2.0 (1.7–2.3)
Hypothyroidism	3.5 (3.1–4.1)	3.0 (2.7–3.4)	3.0 (2.6–3.4)	2.9 (2.5–3.2)	3.2 (2.8–3.6)	3.3 (2.9–3.9)	2.2 (1.7–2.8)	3.4 (3.0–3.9)	2.8 (2.4–3.1)	3.2 (2.8–3.6)
Population achievement:										
All 70 included indicators†	85.2 (84.4–85.8)	84.2 (83.2–85.3)	84.6 (83.6–85.4)	84.1 (83.1–85.1)	84.5 (83.5–85.4)	84.2 (83.0–85.1)	83.7 (82.3–85.0)	83.6 (82.4–84.8)	84.2 (83.3–85.2)	84.3 (83.1–85.5)
All 12 outcome indicators:	74.9 (73.7–76.2)	74.4 (72.6–76.2)	73.8 (71.7–75.4)	73.2 (71.6–74.9)	73.6 (72.3–75.0)	73.0 (71.3–74.6)	72.1 (70.1–73.7)	73.3 (71.6–74.9)	72.9 (71.3–74.5)	73.0 (71.3–74.7)
Subset of 9‡	74.8 (73.4–76.0)	73.9 (72.2–75.7)	73.4 (71.4–75.1)	73.0 (71.5–74.7)	73.0 (71.7–74.6)	72.7 (70.8–74.4)	71.7 (69.8–73.5)	72.7 (70.8–74.3)	72.7 (71.1–74.5)	72.5 (70.8–74.2)
Morbidity load:										
Across all 19 clinical domains	2.3 (2.1–2.5)	2.2 (2.1–2.3)	2.2 (2.0–2.3)	2.1 (2.0–2.3)	2.2 (2.0–2.4)	2.0 (1.9–2.2)	1.7 (1.5–1.9)	2.0 (1.8–2.2)	1.8 (1.7–2.0)	2.1 (1.9–2.3)
Across subset of 5§	0.46 (0.42–0.49)	0.44 (0.41–0.46)	0.43 (0.38–0.47)	0.42 (0.38–0.46)	0.44 (0.40–0.47)	0.39 (0.36–0.43)	0.33 (0.27–0.37)	0.39 (0.36–0.44)	0.35 (0.32–0.39)	0.42 (0.38–0.46)
Other information across 2001 LSOAs										
Median (IQR) distance to practice (miles)	0.5 (0.3–0.9)	0.5 (0.3–0.7)	0.5 (0.3–0.9)	0.6 (0.3–1.1)	0.4 (0.3–0.7)	0.6 (0.3–1.1)	0.3 (0.1–0.4)	0.5 (0.3–1.0)	0.6 (0.3–1.0)	0.6 (0.3–1.2)
Rural (%)	18.5	11.4	19.1	28.3	15.2	30.1	0.1	21.4	22	32.5

IQR=interquartile range. SMR=standardised mortality rate. LSOA=Lower Super Output Area. IMD=Index of Multiple Deprivation. COPD=chronic obstructive pulmonary disease.

*Strategic health authorities were the highest geographical organisational level for NHS England during the study period. This structure was abolished in 2013.

†For details on all included indicators please see online appendix 1, table A2.

‡Blood pressure control for coronary heart disease, chronic kidney disease, diabetes, hypertension, and stroke; cholesterol control for coronary heart disease, diabetes, and stroke; for diabetes.

§Clinical domains associated with the subset of 9 outcome indicators: coronary heart disease, chronic kidney disease, hypertension, diabetes, and stroke.

Table 2| Regression analyses set 1, effect of Quality and Outcomes Framework (QOF) population achievement on 2011-12 standardised mortality rate (SMR). Values are reporting coefficients (99% CIs), and P values (t values)

	QOF year 8 (2011-12)*	QOF year 7 (2010-11)*	QOF year 6 (2009-10)*	QOF year 5 (2008-09)*
Outcome: all-cause SMR. QOF predictors: overall population achievement, overall morbidity load				
Index of Multiple Deprivation 2010	7.44 (7.24 to 7.65), <0.001 (93.67)	7.45 (7.24 to 7.65), <0.001 (93.60)	7.41 (7.21 to 7.62), <0.001 (93.25)	7.46 (7.26 to 7.67), <0.001 (93.40)
Rural (v urban)	-44.52 (-52.17 to -36.86), <0.001 (-14.98)	-44.20 (-51.86 to -36.53), <0.001 (-14.86)	-44.62 (-52.28 to -36.96), <0.001 (-15.01)	-43.76 (-51.42 to -36.10), <0.001 (-14.72)
% White population	-0.45 (-0.63 to -0.28), <0.001 (-6.60)	-0.42 (-0.60 to -0.25), <0.001 (-6.13)	-0.47 (-0.65 to -0.29), <0.001 (-6.82)	-0.44 (-0.62 to -0.27), <0.001 (-6.48)
% Population achievement (PA _{overall})†	0.73 (-0.60 to 2.07), 0.158 (1.41)	0.47 (-0.93 to 1.87), 0.39 (0.86)	-0.06 (-1.36 to 1.24), 0.903 (-0.12)	-0.26 (-1.67 to 1.15), 0.636 (-0.47)
Morbidity load (ML _{total})†	-77.38 (-86.91 to -67.84), <0.001 (-20.90)	-72.06 (-80.97 to -63.15), <0.001 (-20.84)	-69.55 (-78.42 to -60.69), <0.001 (-20.20)	-82.27 (-92.11 to -72.43), <0.001 (-21.54)
Model intercept	530.88 (421.25 to 640.50), <0.001 (12.47)	548.14 (431.23 to 665.04), <0.001 (12.08)	588.22 (480.14 to 696.31), <0.001 (14.02)	609.38 (490.74 to 728.03), <0.001 (13.23)
Adjusted R ² (%)	29.1	29.1	29.0	29.1
Outcome: all-cause SMR. QOF predictors: overall outcome population achievement, overall morbidity load				
Index of Multiple Deprivation 2010	7.45 (7.25 to 7.66), <0.001 (93.65)	7.44 (7.24 to 7.65), <0.001 (93.37)	7.41 (7.20 to 7.61), <0.001 (93.25)	7.48 (7.27 to 7.68), <0.001 (93.94)
Rural (v urban)	-44.40 (-52.06 to -36.74), <0.001 (-14.94)	-44.19 (-51.86 to -36.53), <0.001 (-14.85)	-44.64 (-52.30 to -36.98), <0.001 (-15.01)	-43.63 (-51.29 to -35.97), <0.001 (-14.68)
% White population	-0.46 (-0.64 to -0.29), <0.001 (-6.72)	-0.42 (-0.60 to -0.24), <0.001 (-6.03)	-0.47 (-0.64 to -0.29), <0.001 (-6.68)	-0.46 (-0.63 to -0.28), <0.001 (-6.66)
% Population achievement (PA _{outcome})†	0.79 (-0.21 to 1.79), 0.041 (2.05)	-0.01 (-1.01 to 0.99), 0.975 (-0.03)	-0.11 (-1.09 to 0.86), 0.764 (-0.30)	0.81 (-0.18 to 1.80), 0.036 (2.10)
Morbidity load (ML _{total})†	-78.65 (-88.42 to -68.89), <0.001 (-20.75)	-71.67 (-80.79 to -62.56), <0.001 (-20.26)	-69.38 (-78.39 to -60.37), <0.001 (-19.83)	-84.37 (-94.33 to -74.40), <0.001 (-21.81)
Model intercept	535.48 (464.16 to 606.80), <0.001 (19.34)	587.46 (517.62 to 657.30), <0.001 (21.67)	590.77 (522.92 to 658.63), <0.001 (22.43)	531.58 (459.93 to 603.23), <0.001 (19.11)
Adjusted R ² (%)	29.1	29.1	29.0	29.2
Outcome: condition-specific SMR. QOF predictors: 9 indicator outcome population achievement, 5 domains morbidity load				
Index of Multiple Deprivation 2010	2.41 (2.27 to 2.55), <0.001 (43.48)	2.40 (2.26 to 2.54), <0.001 (43.30)	2.41 (2.27 to 2.55), <0.001 (43.49)	2.41 (2.27 to 2.55), <0.001 (43.43)
Rural (v urban)	-3.67 (-9.11 to 1.77), 0.082 (-1.74)	-3.81 (-9.25 to 1.64), 0.072 (-1.80)	-3.70 (-9.14 to 1.75), 0.08 (-1.75)	-3.61 (-9.06 to 1.83), 0.087 (-1.71)
% White population	-0.32 (-0.45 to -0.20), <0.001 (-6.68)	-0.34 (-0.46 to -0.21), <0.001 (-6.82)	-0.32 (-0.44 to -0.19), <0.001 (-6.43)	-0.32 (-0.45 to -0.20), <0.001 (-6.59)
% Population achievement (PA _{outcome})†	0.26 (-0.47 to 0.98), 0.359 (0.92)	0.11 (-0.57 to 0.78), 0.688 (0.40)	-0.08 (-0.74 to 0.59), 0.768 (-0.30)	0.21 (-0.46 to 0.87), 0.419 (0.81)
Morbidity load (ML ₉)†	32.37 (4.87 to 59.87), 0.002 (3.03)	40.39 (12.02 to 68.76), <0.001 (3.67)	34.76 (6.64 to 62.89), 0.001 (3.18)	33.09 (3.20 to 62.98), 0.004 (2.85)
Model intercept	140.23 (87.17 to 193.29), <0.001 (6.81)	150.12 (102.79 to 197.45), <0.001 (8.17)	163.96 (117.51 to 210.40), <0.001 (9.09)	144.46 (97.14 to 191.77), <0.001 (7.86)
Adjusted R ² (%)	7.8	7.8	7.8	7.8

*Year of reference for the two QOF variables in the model.

†Time-varying QOF variables across different models. All other variables do not vary over time.

Table 3| Regression analyses set 2, effect of change in Quality and Outcomes Framework (QOF) population achievement on 2011-12 standardised mortality rate (SMR). Values are reporting coefficients (99% CIs), and P values (t values)

	QOF year 8 (2011-12)*		QOF year 7 (2010-11)*	QOF year 6 (2009-10)*
	5 year change in achievement (year 8 to 3)†	3 year change in achievement (year 8 to 5)‡	3 year change in achievement (year 7 to 4)†	3 year change in achievement (year 6 to 3)†
Outcome: all-cause SMR. QOF predictors: change in overall population achievement, overall morbidity load				
Index of Multiple Deprivation 2010	7.43 (7.22 to 7.63), <0.001 (93.43)	7.43 (7.22 to 7.63), <0.001 (93.46)	7.44 (7.23 to 7.64), <0.001 (93.55)	7.41 (7.21 to 7.61), <0.001 (93.53)
Rural (v urban)	-44.45 (-52.11 to -36.79), <0.001 (-14.95)	-44.40 (-52.05 to -36.74), <0.001 (-14.94)	-44.13 (-51.79 to -36.46), <0.001 (-14.84)	-44.62 (-52.28 to -36.96), <0.001 (-15.01)
% White population	-0.44 (-0.62 to -0.27), <0.001 (-6.44)	-0.46 (-0.64 to -0.28), <0.001 (-6.69)	-0.42 (-0.60 to -0.24), <0.001 (-6.09)	-0.47 (-0.64 to -0.29), <0.001 (-6.81)
% Change in population achievement (PA _{overall})‡	0.86 (-0.33 to 2.05), 0.063 (1.86)	1.79 (0.19 to 3.40), 0.004 (2.87)	0.78 (-0.73 to 2.29), 0.183 (1.33)	0.11 (-1.23 to 1.45), 0.833 (0.21)
Morbidity load (ML _{tot})‡	-76.81 (-86.28 to -67.34), <0.001 (-20.9)	-76.57 (-86.04 to -67.10), <0.001 (-20.83)	-71.69 (-80.54 to -62.85), <0.001 (-20.88)	-69.61 (-78.40 to -60.81), <0.001 (-20.39)
Model intercept	590.56 (571.03 to 610.09), <0.001 (77.90)	594.36 (574.46 to 614.26), <0.001 (76.93)	587.03 (567.73 to 606.32), <0.001 (78.37)	583.14 (563.95 to 602.34), <0.001 (78.26)
Adjusted R ² (%)	29.1	29.1	29.1	29.0
Outcome: all-cause SMR. QOF predictors: change in overall outcome population achievement, overall morbidity load				
Index of Multiple Deprivation 2010	7.44 (7.23 to 7.64), <0.001 (93.74)	7.44 (7.23 to 7.64), <0.001 (93.74)	7.44 (7.23 to 7.64), <0.001 (93.67)	7.41 (7.21 to 7.61), <0.001 (93.60)
Rural (v urban)	-44.49 (-52.15 to -36.84), <0.001 (-14.97)	-44.50 (-52.16 to -36.84), <0.001 (-14.97)	-44.25 (-51.91 to -36.59), <0.001 (-14.88)	-44.62 (-52.28 to -36.96), <0.001 (-15.01)
% White population	-0.44 (-0.62 to -0.27), <0.001 (-6.44)	-0.45 (-0.63 to -0.27), <0.001 (-6.53)	-0.42 (-0.60 to -0.24), <0.001 (-6.07)	-0.47 (-0.65 to -0.29), <0.001 (-6.84)
% Change in population achievement (PA _{outcome})‡	0.35 (-0.61 to 1.30), 0.35 (0.94)	0.41 (-0.82 to 1.63), 0.39 (0.86)	-0.94 (-2.10 to 0.23), 0.038 (-2.07)	-0.74 (-1.82 to 0.34), 0.077 (-1.77)
Morbidity load (ML _{tot})‡	-77.01 (-86.50 to -67.52), <0.001 (-20.89)	-76.94 (-86.43 to -67.46), <0.001 (-20.90)	-71.28 (-80.14 to -62.42), <0.001 (-20.73)	-69.38 (-78.18 to -60.58), <0.001 (-20.31)
Model intercept	589.71 (570.18 to 609.24), <0.001 (77.79)	590.43 (570.87 to 609.99), <0.001 (77.76)	584.50 (565.04 to 603.97), <0.001 (77.35)	582.22 (562.98 to 601.46), <0.001 (77.96)
Adjusted R ² (%)	29.1	29.1	29.1	29.0
Outcome: condition-specific SMR. QOF predictors: change in 9 indicator outcome population achievement, 5 domains morbidity load				
Index of Multiple Deprivation 2010	2.41 (2.27 to 2.55), <0.001 (43.47)	2.41 (2.27 to 2.55), <0.001 (43.46)	2.40 (2.26 to 2.54), <0.001 (43.31)	2.41 (2.27 to 2.55), <0.001 (43.54)
Rural (v urban)	-3.71 (-9.15 to 1.74), 0.079 (-1.75)	-3.72 (-9.16 to 1.72), 0.078 (-1.76)	-3.85 (-9.30 to 1.59), 0.068 (-1.82)	-3.68 (-9.13 to 1.76), 0.081 (-1.74)
% White population	-0.33 (-0.45 to -0.20), <0.001 (-6.66)	-0.32 (-0.45 to -0.20), <0.001 (-6.61)	-0.33 (-0.46 to -0.21), <0.001 (-6.84)	-0.32 (-0.45 to -0.20), <0.001 (-6.65)
% Change in population achievement (PA _{outcome})‡	-0.17 (-0.83 to 0.49), 0.508 (-0.66)	0.00 (-0.85 to 0.86), 0.988 (0.01)	-0.22 (-1.02 to 0.57), 0.47 (-0.72)	-0.57 (-1.31 to 0.17), 0.046 (-2.00)
Morbidity load (ML ₅)‡	34.88 (8.03 to 61.73), 0.001 (3.35)	34.53 (7.72 to 61.35), 0.001 (3.32)	41.55 (13.77 to 69.32), <0.001 (3.85)	34.47 (6.79 to 62.14), 0.001 (3.21)
Model intercept	159.27 (145.95 to 172.59), <0.001 (30.80)	158.54 (145.35 to 171.73), <0.001 (30.97)	156.99 (143.92 to 170.06), <0.001 (30.94)	159.01 (145.98 to 172.04), <0.001 (31.43)
Adjusted R ² (%)	7.8	7.8	7.8	7.8

*Year of reference for QOF morbidity load in the models.

†Years of reference for QOF population achievement change in the models.

‡Time-varying QOF variables across different models. All other variables do not vary over time.

Table 4| Regression analyses set 3, effect of Quality and Outcomes Framework (QOF) population achievement on yearly standardised mortality rate (SMR), 2007–11, over time. Values are reporting coefficients (99% CIs), and P values (t values)

	Model lag between population achievement year and outcome year			
	Lag=0*	Lag=1†	Lag=2‡	Lag=3§
Outcome: all-cause SMR. QOF predictors: overall population achievement, overall morbidity load				
Year:				
2007	Reference	—	—	—
2008	-9.05 (-14.23 to -3.87), <0.001 (-4.5)	Reference	—	—
2009	-38.08 (-43.34 to -32.82), <0.001 (-18.65)	-28.98 (-33.95 to -24.02), <0.001 (-15.04)	Reference	—
2010	-46.42 (-52.15 to -40.68), <0.001 (-20.85)	-36.33 (-41.70 to -30.95), <0.001 (-17.41)	-7.59 (-12.66 to -2.52), <0.001 (-3.86)	Reference
2011	-72.37 (-78.32 to -66.43), <0.001 (-31.35)	-63.40 (-68.92 to -57.88), <0.001 (-29.58)	-34.09 (-39.38 to -28.80), <0.001 (-16.59)	-26.03 (-30.64 to -21.42), <0.001 (-14.54)
2012	-74.55 (-80.14 to -68.97), <0.001 (-34.4)	-64.14 (-69.27 to -59.01), <0.001 (-32.21)	-35.34 (-40.21 to -30.46), <0.001 (-18.67)	-26.33 (-31.05 to -21.61), <0.001 (-14.36)
Index of Multiple Deprivation 2010	7.75 (7.55 to 7.94), <0.001 (103.84)	7.72 (7.53 to 7.91), <0.001 (103.89)	7.66 (7.46 to 7.85), <0.001 (102.09)	7.67 (7.47 to 7.87), <0.001 (98.36)
Rural (v urban)	-50.73 (-57.94 to -43.52), <0.001 (-18.12)	-50.34 (-57.51 to -43.17), <0.001 (-18.08)	-49.82 (-57.05 to -42.59), <0.001 (-17.74)	-48.51 (-56.02 to -41.00), <0.001 (-16.64)
% White population	-1.12 (-1.29 to -0.96), <0.001 (-17.48)	-1.01 (-1.17 to -0.84), <0.001 (-15.79)	-0.86 (-1.02 to -0.69), <0.001 (-13.25)	-0.69 (-0.87 to -0.52), <0.001 (-10.3)
% Population achievement (PA _{oval})¶	-0.90 (-1.87 to 0.07), 0.016 (-2.4)	-0.46 (-1.46 to 0.54), 0.24 (-1.18)	-0.41 (-1.46 to 0.64), 0.312 (-1.01)	-0.87 (-2.01 to 0.26), 0.047 (-1.99)
Morbidity burden (ML _{tot})¶¶	-79.47 (-87.83 to -71.10), <0.001 (-24.47)	-78.26 (-86.58 to -69.93), <0.001 (-24.21)	-76.23 (-84.62 to -67.84), <0.001 (-23.39)	-74.74 (-83.40 to -66.08), <0.001 (-22.23)
Model intercept	809.09 (727.94 to 890.24), <0.001 (25.68)	750.84 (666.93 to 834.76), <0.001 (23.05)	702.02 (614.47 to 789.56), <0.001 (20.66)	715.34 (620.86 to 809.81), <0.001 (19.5)
Overall R ² (%)	19.2	19.7	19.8	20.2
rho (%)††	26.8	27.2	27.6	29.0
Outcome: all-cause SMR. QOF predictors: overall outcome population achievement, overall morbidity load				
Year:				
2007	Reference	—	—	—
2008	-9.87 (-15.03 to -4.71), <0.001 (-4.92)	Reference	—	—
2009	-39.20 (-44.43 to -33.98), <0.001 (-19.33)	-29.51 (-34.45 to -24.56), <0.001 (-15.37)	Reference	—
2010	-45.68 (-51.49 to -39.86), <0.001 (-20.23)	-36.89 (-42.22 to -31.56), <0.001 (-17.81)	-7.95 (-12.99 to -2.90), <0.001 (-4.06)	Reference
2011	-72.11 (-78.13 to -66.10), <0.001 (-30.87)	-62.89 (-68.49 to -57.29), <0.001 (-28.91)	-34.60 (-39.85 to -29.35), <0.001 (-16.98)	-26.78 (-31.37 to -22.20), <0.001 (-15.05)
2012	-73.80 (-79.32 to -68.28), <0.001 (-34.45)	-64.05 (-69.21 to -58.89), <0.001 (-31.96)	-34.98 (-39.92 to -30.04), <0.001 (-18.24)	-27.52 (-32.19 to -22.84), <0.001 (-15.15)
Index of Multiple Deprivation 2010	7.76 (7.57 to 7.95), <0.001 (104.14)	7.73 (7.54 to 7.92), <0.001 (104.15)	7.67 (7.47 to 7.86), <0.001 (102.44)	7.69 (7.49 to 7.89), <0.001 (98.87)
Rural (v urban)	-50.75 (-57.96 to -43.54), <0.001 (-18.13)	-50.33 (-57.50 to -43.16), <0.001 (-18.08)	-49.82 (-57.05 to -42.59), <0.001 (-17.74)	-48.57 (-56.08 to -41.06), <0.001 (-16.66)
% White population	-1.13 (-1.30 to -0.97), <0.001 (-17.58)	-1.02 (-1.18 to -0.85), <0.001 (-15.82)	-0.87 (-1.03 to -0.70), <0.001 (-13.3)	-0.70 (-0.88 to -0.53), <0.001 (-10.38)
% Population achievement (PA _{oval})¶	0.24 (-0.50 to 0.97), 0.408 (0.83)	0.23 (-0.53 to 0.98), 0.437 (0.78)	0.23 (-0.56 to 1.01), 0.457 (0.74)	0.16 (-0.69 to 1.00), 0.635 (0.48)
Morbidity burden (ML _{tot})¶¶	-80.79 (-89.26 to -72.32), <0.001 (-24.57)	-79.09 (-87.50 to -70.68), <0.001 (-24.22)	-76.99 (-85.45 to -68.52), <0.001 (-23.43)	-75.72 (-84.45 to -66.99), <0.001 (-22.34)
Model intercept	719.13 (666.22 to 772.05), <0.001 (35.01)	697.82 (643.48 to 752.17), <0.001 (33.08)	652.79 (596.18 to 709.40), <0.001 (29.7)	633.21 (572.58 to 693.83), <0.001 (26.9)
Overall R ² (%)	19.2%	19.7%	19.8%	20.2%
rho (%)††	26.8%	27.2%	27.6%	29.0%

BMJ: first published as 10.1136/bmj.h904 on 2 March 2015. Downloaded from <http://www.bmj.com/> on 23 April 2024 by guest. Protected by copyright.

Table 4 (continued)

	Model lag between population achievement year and outcome year			
	Lag=0*	Lag=1†	Lag=2‡	Lag=3§
Outcome: condition-specific SMR. QOF predictors: 9 indicator outcome population achievement, 5 domains morbidity load				
Year:				
2007	Reference	—	—	—
2008	-11.58 (-15.66 to -7.49), <0.001 (-7.29)	—	—	—
2009	-33.22 (-37.35 to -29.09), <0.001 (-20.71)	-21.71 (-25.60 to -17.83), <0.001 (-14.4)	—	—
2010	-47.52 (-51.66 to -43.39), <0.001 (-29.58)	-36.04 (-39.95 to -32.12), <0.001 (-23.69)	-14.60 (-18.30 to -10.90), <0.001 (-10.16)	—
2011	-76.64 (-80.81 to -72.47), <0.001 (-47.34)	-65.22 (-69.12 to -61.31), <0.001 (-42.99)	-43.89 (-47.63 to -40.16), <0.001 (-30.27)	-29.27 (-32.83 to -25.71), <0.001 (-21.19)
2012	-73.49 (-77.90 to -69.08), <0.001 (-42.92)	-62.18 (-66.11 to -58.24), <0.001 (-40.73)	-40.58 (-44.29 to -36.87), <0.001 (-28.18)	-26.33 (-29.92 to -22.74), <0.001 (-18.88)
Index of Multiple Deprivation 2010	2.82 (2.71 to 2.94), <0.001 (62.4)	2.74 (2.62 to 2.87), <0.001 (58.82)	2.69 (2.56 to 2.81), <0.001 (55.89)	2.62 (2.49 to 2.75), <0.001 (51.41)
Rural (v urban)	-6.54 (-10.98 to -2.09), <0.001 (-3.79)	-6.50 (-11.09 to -1.92), <0.001 (-3.65)	-5.69 (-10.42 to -0.97), 0.002 (-3.1)	-5.74 (-10.74 to -0.73), 0.003 (-2.95)
% White population	-0.50 (-0.60 to -0.39), <0.001 (-12.44)	-0.48 (-0.58 to -0.37), <0.001 (-11.63)	-0.45 (-0.56 to -0.34), <0.001 (-10.65)	-0.39 (-0.51 to -0.27), <0.001 (-8.66)
% Population achievement (PA _{total})¶	-0.03 (-0.51 to 0.46), 0.891 (-0.14)	-0.01 (-0.51 to 0.49), 0.956 (-0.06)	0.24 (-0.28 to 0.76), 0.231 (1.2)	-0.02 (-0.57 to 0.53), 0.943 (-0.07)
Morbidity burden (ML _{tot})¶¶	0.00 (-23.17 to 23.18), 1 (0)	6.10 (-17.51 to 29.71), 0.506 (0.67)	12.46 (-11.61 to 36.53), 0.182 (1.33)	22.21 (-3.04 to 47.46), 0.023 (2.27)
Model intercept	260.01 (225.70 to 294.32), <0.001 (19.52)	245.30 (210.15 to 280.45), <0.001 (17.98)	201.76 (165.25 to 238.26), <0.001 (14.24)	198.13 (159.24 to 237.03), <0.001 (13.12)
Overall R ² (%)	6.5%	6.2%	5.9%	5.6%
rho (%)††	13.3%	15.1%	16.1%	17.0%

*Population achievement in QOF year 3 (2006-07) corresponds to SMRs in 2007, in QOF year 4 corresponds to SMRs in 2008, etc.
 †Population achievement in QOF year 3 (2006-07) corresponds to SMRs in 2008, in QOF year 4 corresponds to SMRs in 2009, etc.
 ‡Population achievement in QOF year 3 (2006-07) corresponds to SMRs in 2009, in QOF year 4 corresponds to SMRs in 2010, etc.
 §Population achievement in QOF year 3 (2006-07) corresponds to SMRs in 2010, in QOF year 4 corresponds to SMRs in 2011, etc.
 ¶Varying values over time, all other predictors are constant over time.
 ¶¶No time lag for morbidity load, corresponds to "same" year SMRs (such as morbidity load in QOF year 8 (2011-12) corresponds to SMRs in 2012).
 ††Fraction of error variance due to random effects component.

BMJ: first published as 10.1136/bmj.h904 on 2 March 2015. Downloaded from http://www.bmj.com/ on 23 April 2024 by guest. Protected by copyright.

Figures

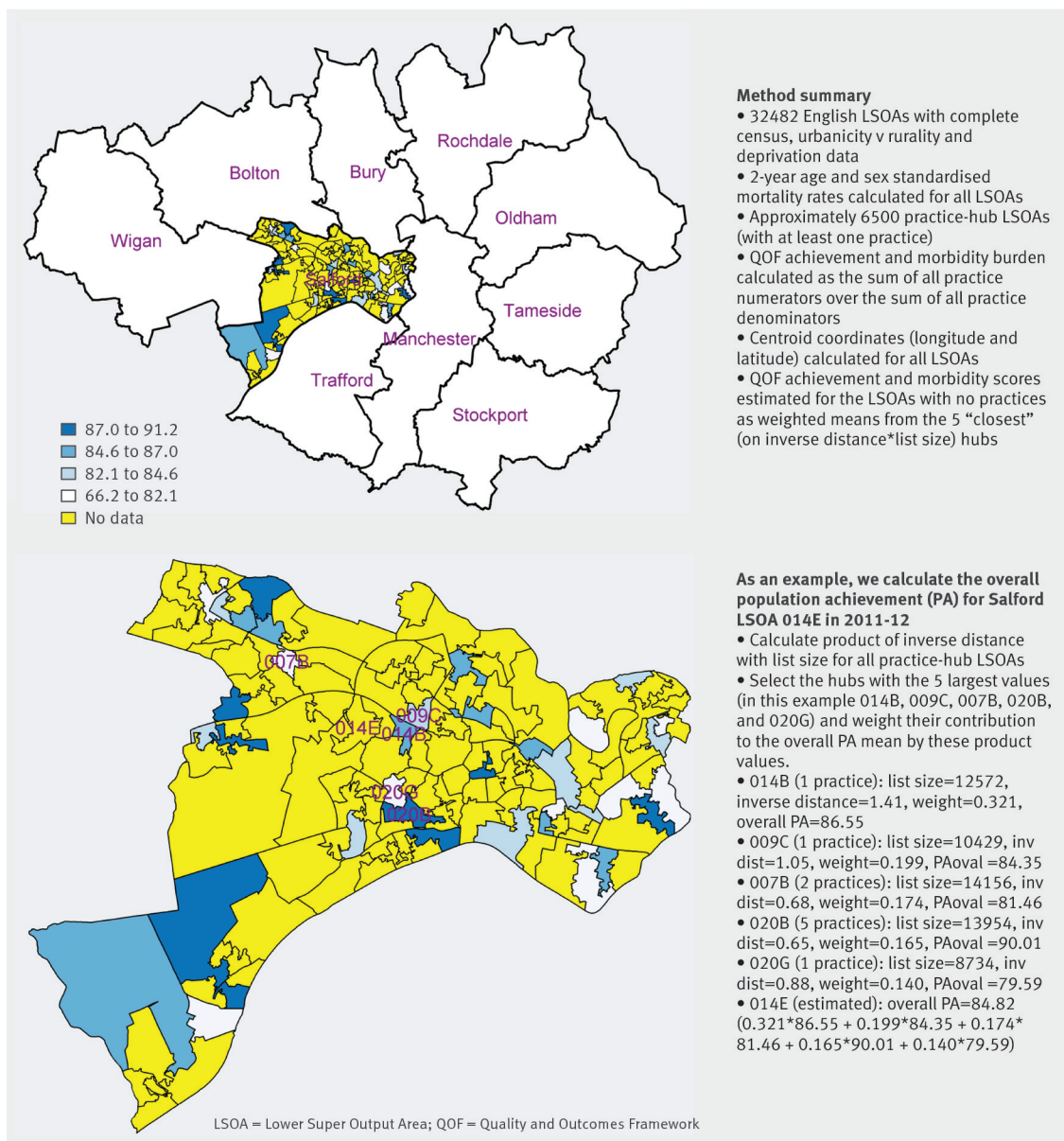


Fig 1 Spatial weighted estimation method summary and example.

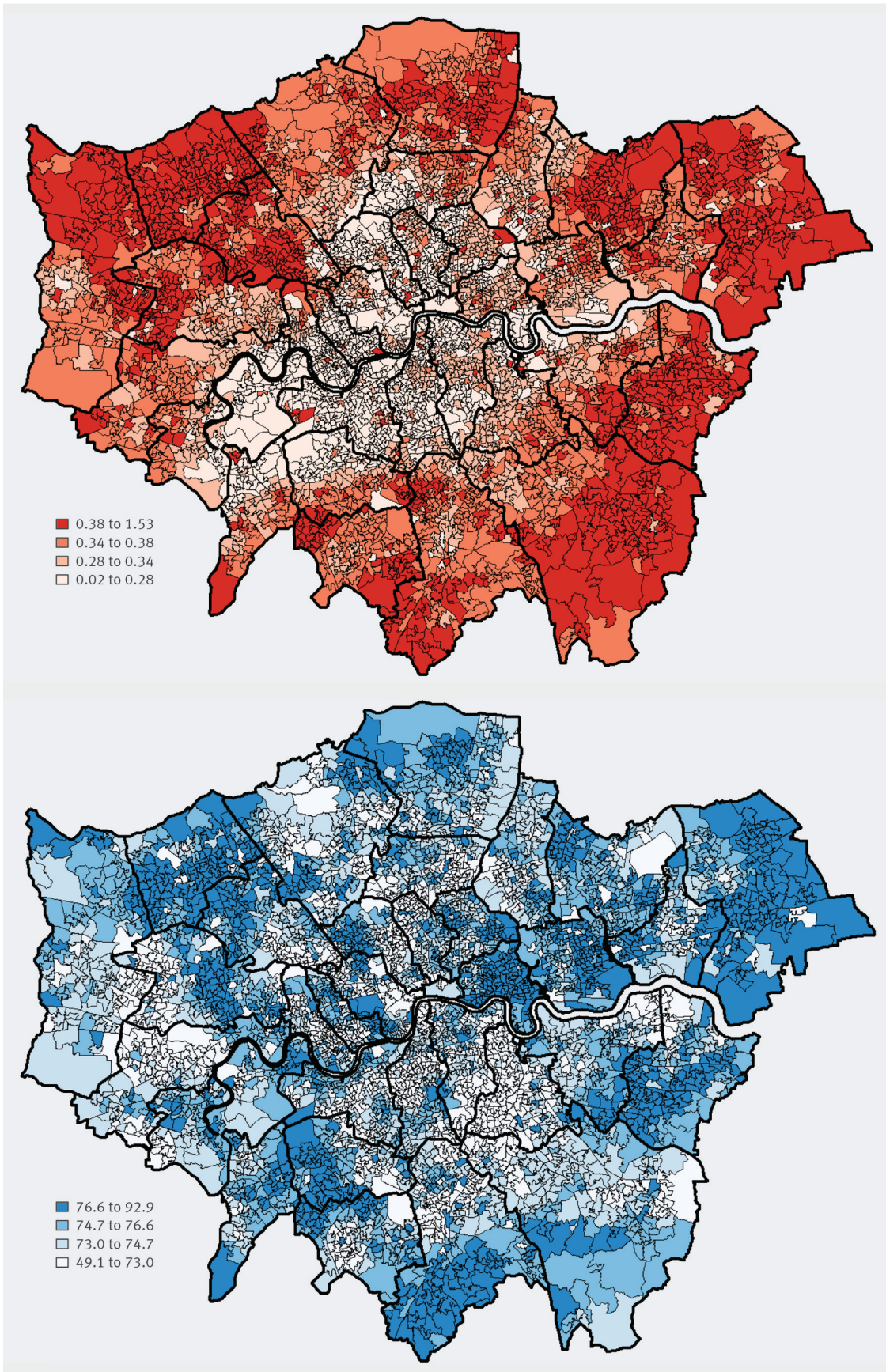


Fig 2 Complete local attendance spatial maps for morbidity load ML_9 (red) and population achievement PA_{outx} (blue) in Greater London for specific conditions and indicators. ML_9 is the sum of QOF registers for nine specific intermediate outcome indicators over the practice list size (blood pressure control for coronary heart disease, chronic kidney disease, diabetes, hypertension, and stroke; cholesterol control for coronary heart disease, diabetes, and stroke; and HbA_{1c} control for diabetes).

PA_{outx} is the aggregate QOF population achievement (that is, numerator over the sum of the denominator and exceptions) across the nine intermediate outcome indicators.

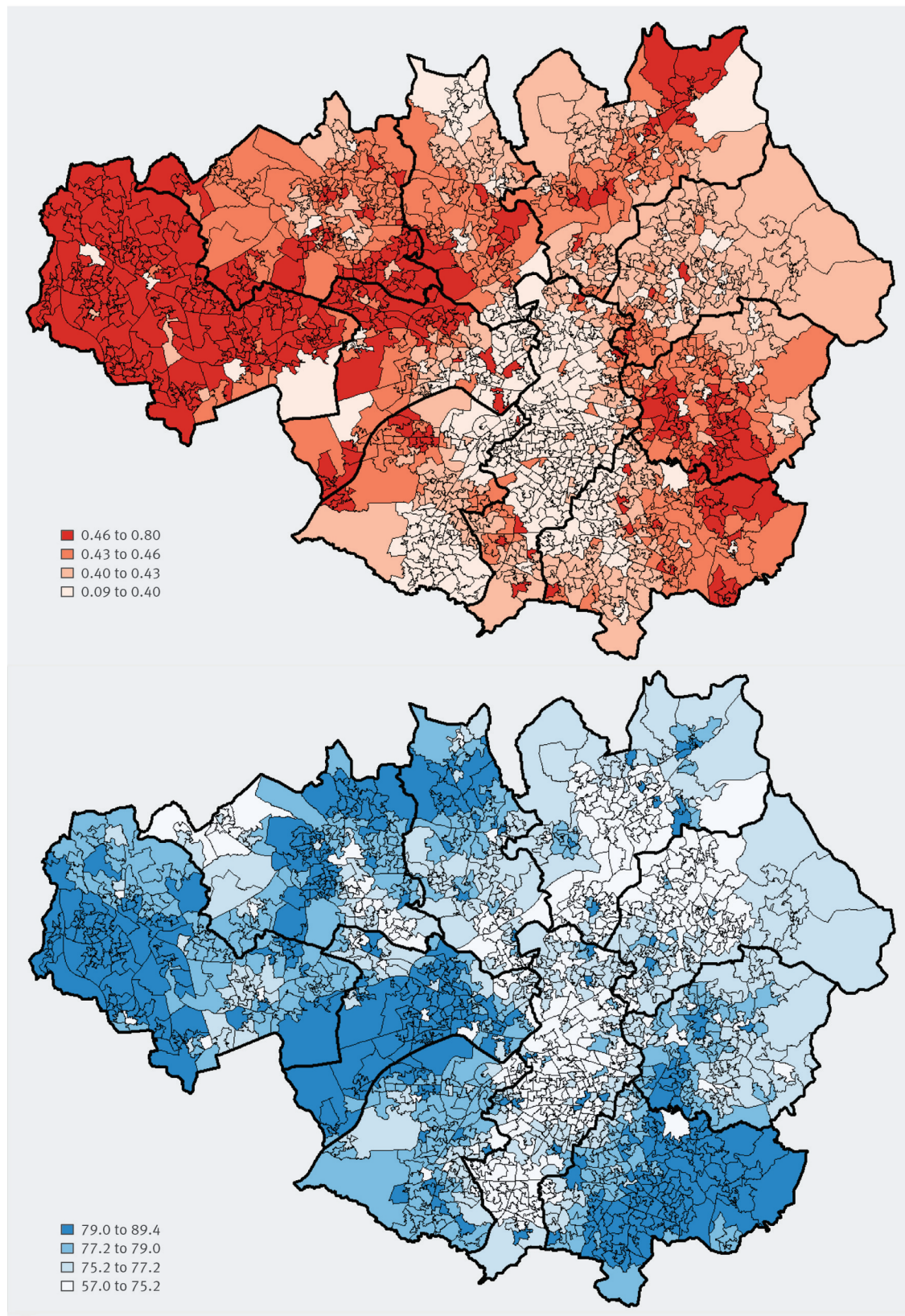


Fig 3 Complete local attendance spatial maps for morbidity load ML_9 (red) and population achievement PA_{outx} (blue) in Greater Manchester for specific conditions and indicators. ML_9 is the sum of QOF registers for nine specific intermediate outcome indicators over the practice list size (blood pressure control for coronary heart disease, chronic kidney disease, diabetes, hypertension, and stroke; cholesterol control for coronary heart disease, diabetes, and stroke; and HbA_{1c} control for diabetes). PA_{outx} is the aggregate QOF population achievement (that is, numerator over the sum of the denominator and exceptions) across the nine intermediate outcome indicators.

BMJ: first published as 10.1136/bmj.h904 on 2 March 2015. Downloaded from <http://www.bmj.com/> on 23 April 2024 by guest. Protected by copyright.

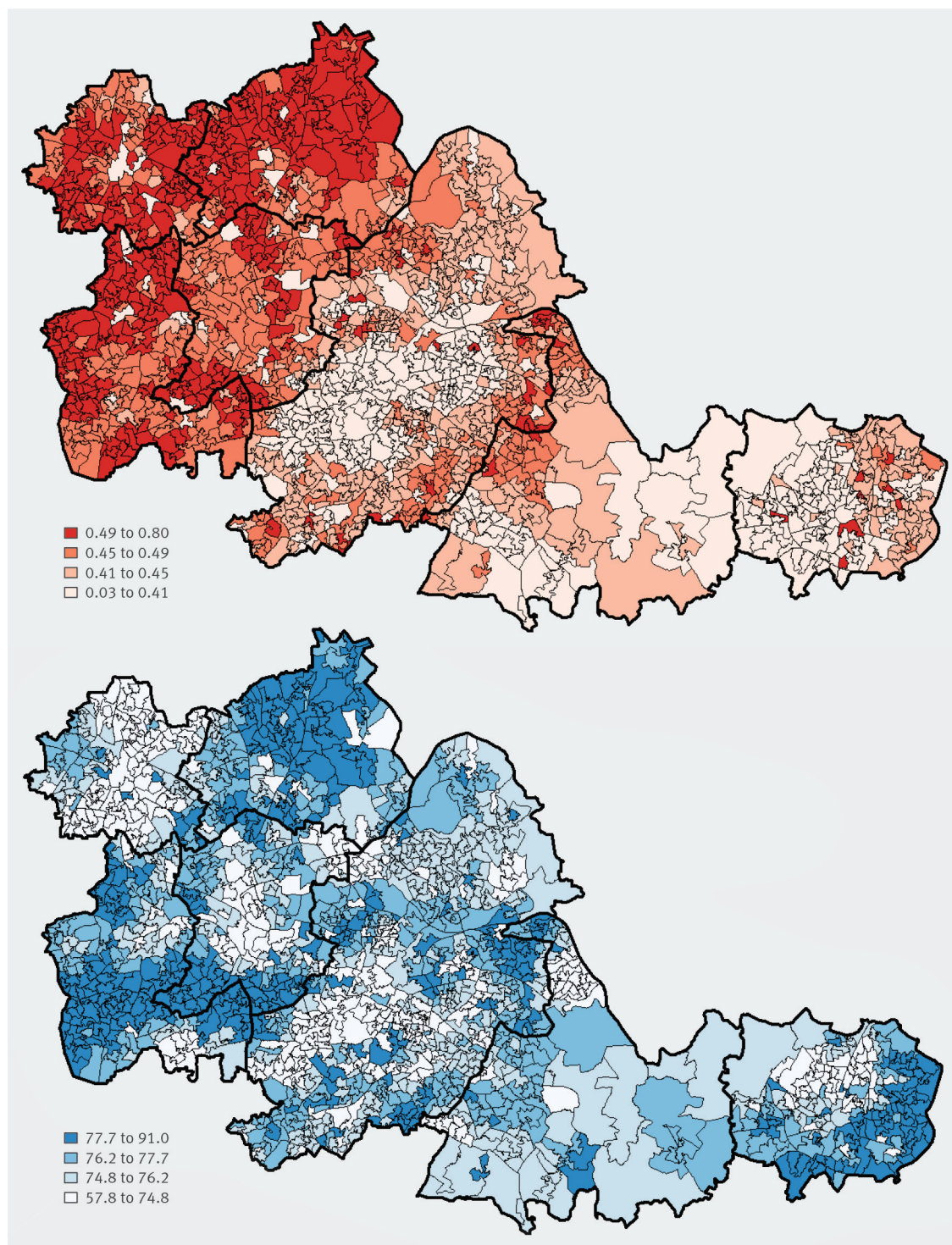


Fig 4 Complete local attendance spatial maps for morbidity load ML_9 (red) and population achievement PA_{outx} (blue) in the West Midlands for specific conditions and indicators. ML_9 is the sum of QOF registers for nine specific intermediate outcome indicators over the practice list size (blood pressure control for coronary heart disease, chronic kidney disease, diabetes, hypertension, and stroke; cholesterol control for coronary heart disease, diabetes, and stroke; and HbA_{1c} control for diabetes). PA_{outx} is the aggregate QOF population achievement (that is, numerator over the sum of the denominator and exceptions) across the nine intermediate outcome indicators.

BMJ: first published as 10.1136/bmj.h904 on 2 March 2015. Downloaded from <http://www.bmj.com/> on 23 April 2024 by guest. Protected by copyright.