Socioeconomic deprivation scores as predictors of variations in NHS practice payments: a longitudinal study of English general practices 2013–2017

INTRODUCTION
Socioeconomic deprivation strongly predicts adverse health outcomes, and is associated with earlier and greater multimorbidity, including up to an 18-year difference in disability-free life expectancy between most and least deprived populations. Disparities in health outcomes persist, despite an absolute and relative decrease in all-cause mortality in lower socioeconomic groups between 1990 and 2010. In 1971, Julian Tudor Hart suggested an 'inverse care law', where those most in need of medical care were least likely to receive it. This was most likely to arise from income inequalities (between users of health care), but practices serving populations with higher levels of socioeconomic deprivation have greater workloads. Although, in real terms, NHS general practice funding has been increasing since 2012–2013 and is projected to continue, funding for secondary care services increased at a more rapid rate than primary care in the first half of this decade. A King’s Fund report estimated that there was a 15% increase in the overall number of general practice consultations between 2010–2011 and 2014–2015. Over the same period, the proportion of NHS funding spent on general practice declined by 0.4%. This undercuts the role primary care can play in reducing the impact of socioeconomic factors on health, and in reducing health inequalities between population subgroups.

When the General Medical Services (GMS) contract was introduced in 2004, it included a formula, the Carr-Hill formula (known as the global sum allocation formula), to ensure that funding reflects 'the contractor’s workload' and to reimburse the ‘unavoidable costs of delivering ... care to the local population’. Weighting includes adjustments for age and sex structure, morbidity and mortality measures, list turnover, individuals living in nursing and residential homes, and staff expenses and rurality. A Minimum Practice Income Guarantee (MPIG), a correction payment to prevent sudden drops in practices’ core funding, was introduced in 2004.

This core funding was supplemented by a pay-for-performance component, the Quality and Outcomes Framework (QOF); one of the aims of QOF was to reduce variations in quality between providers. Calculations of payment for achievement against clinical QOF targets used adjusted rather than actual prevalence until 2009 to protect practices’ incomes. Using adjusted prevalence and the phasing out of MPIG from 2014 to equalise weighted funding per patient across all practices has caused many practices in deprived areas to face financial hardship.

In addition to the GMS contract, two alternative primary care contracts were introduced: the Personal Medical Services (PMS) and the Alternative Provider Medical Services (APMS) contracts, often with [Abstract content]
greater funding for additional services, but not containing the global sum.\textsuperscript{20} GPs may be either profit-sharing partners or salaried to the practice. While PMS are provided only by NHS GP partnership-based practices, APMS can be provided by outside contractors. Providers under these contracts are still eligible for performance payments.

A formula mainly or strongly based on age has some merits, including simplicity, but, without adequate adjustment for other predictors of cross-sectional variations in practice payments. In this longitudinal study, deprivation scores weakly predicted variations in the increase of adjusted total payments per registered patient between 2013 and 2017 at practice level, after accounting for the effects of confounding variables. Other organisational and population factors either did not predict or only weakly predicted, with very small effect sizes, the increase in payments.

**How this fits in**

Socioeconomic deprivation strongly predicts adverse health and is associated with earlier and greater multimorbidity, and consequently greater GP workload. NHS general practice payments in England are designed to compensate for workload, but do not include specific deprivation measures in their weighting. Variables related to population health needs had been found to be poor predictors of cross-sectional variations in practice payments. In this longitudinal study, deprivation scores weakly predicted variations in the increase of adjusted total payments per registered patient between 2013 and 2017 at practice level, after accounting for the effects of confounding variables. Other organisational and population factors either did not predict or only weakly predicted, with very small effect sizes, the increase in payments.

Medical Association’s General Practitioners Committee and NHS England were announced in August 2016.\textsuperscript{21} A House of Commons briefing paper in September 2018 stated that ‘a new funding formula will be developed to better reflect practice workload, including deprivation.’\textsuperscript{22} However, only two small and unrelated changes have been made to the formula in the 2019–2020 financial year.\textsuperscript{23}

In a previous study, variables related to population health needs were found to be poor predictors of cross-sectional variations in practice payments.\textsuperscript{24} Subsequently, the additional years of payments data now available allow a longitudinal analysis. The aim of this study was to examine whether deprivation scores predict variations in the slope over time of NHS practice payments, after adjusting for organisational and population factors.

**METHOD**

**Setting**

The setting was all NHS general practices in England.

**Type of study and study period**

An observational longitudinal study was carried out, covering four consecutive financial years, 2013–2014 to 2016–2017 inclusive.

**Data sources**

The practice-level data used in this study were published as spreadsheets by Public Health England,\textsuperscript{30} NHS England, NHS Digital,\textsuperscript{31–34} the GP Patient Survey,\textsuperscript{35} and the Department of Health. These were combined using the common unique practice identification code into a single database for analysis of practices in England.

**Dependent variable**

The dependent variable used in the study model was the slope over the four study years of adjusted total payments per patient to practices. As in the authors’ previous study,\textsuperscript{25} payments were calculated for each year by taking all payments due to a practice for providing NHS services in a financial year, then subtracting pensions, levies, prescription charge income, and premises payments, and finally dividing the remainder by the number of registered patients in the practice for that year. These deductions remove funding streams that distort workload, performance, and structural income. These data on payments to each practice were published on the NHS England website.\textsuperscript{33–34}

Despite recommendations to update the Carr-Hill formula, it has not been significantly altered since its introduction. Detailed negotiations between the British Medical Association’s General Practitioners Committee and NHS England were announced in August 2016.\textsuperscript{21} A House of Commons briefing paper in September 2018 stated that ‘a new funding formula will be developed to better reflect practice workload, including deprivation.’\textsuperscript{22} However, only two small and unrelated changes have been made to the formula in the 2019–2020 financial year.\textsuperscript{23}
Independent variables
In accordance with the study aim, the main predictor was each practice’s deprivation score, derived from the Index of Multiple Deprivation (IMD), a set of relative measures of deprivation for small areas (lower-layer super output areas) in seven domains: income deprivation; employment deprivation; education, skills, and training deprivation; health deprivation and disability; crime; barriers to housing and services; and living environment deprivation. Individual practice scores were calculated by the Office for National Statistics, using patients’ postcodes. The most recent calculation was used, which was in 2015. It was found that practice IMD 2015 scores are highly correlated with their IMD scores in 2012.

Other factors might act as confounders in this relationship, and, thus, needed to be included in the analysis. Based on plausibility guided by the authors’ conceptual framework for population health research, and on data availability at practice level, variables were included as confounders, which were subdivided into organisational and population variables.

Organisational variables: These were list size, whole-time equivalent (WTE) doctors/1000 patients (2014 and 2016), WTE practice nurses/1000 patients (2014 and 2016), and NHS contract type.

Population variables: These were geographical region, percentage >75 years, percentage black ethnicity, percentage South Asian ethnicity, percentage smokers, percentage claiming disability, percentage with chronic condition, urban versus rural (2017), percentage in nursing home (2015), percentage on coronary heart disease register, percentage on hypertension register, percentage on stroke register, and percentage on diabetic register.

Unless specified, values for all 4 years were used in the model (that is, they were allowed to vary over time).

Missing data
There was not a full set of values for all of the variables in each year studied. However, because the proportion of missing values never exceeded 6% for any variable (Table 1), it was decided not to undertake multiple imputation.

Statistical analysis
Descriptive statistics of the variables, univariable analyses of the relationships between pairs of variables, and finally multivariable analyses were all carried out.

From previous studies, it was anticipated that the distribution of practice payments would be positively skewed and, if so, then a natural logarithmic transformation would be undertaken.

The clustering of measurements within practices was adjusted for by fitting a random intercepts and random slopes model. The a priori rationale for selecting this model was that it was not possible to assume similarities between the intercepts or slopes for all practices’ levels of payments over time. The linear effect of each independent variable on payment over time was modelled by fitting interaction terms, formed as the product of each independent variable with year. The significance of the interaction term between deprivation and year would indicate whether or not deprivation independently predicted the slope of payments over time. Statistical significance was set at the 5% level. Post-estimation statistics were used to generate random effects values, predicted values for payments, and to calculate the intraclass correlation (ICCs) for a null model (containing no predictors).

Stata (version 14) was used for all analyses.

RESULTS
Number of practices used in the analysis
Payments data were published for 8178 practices in the financial years 2013–2014 to 2016–2017. A small number of practices received implausibly large or small (including negative) payments per patient and were treated as erroneous and omitted. On this basis, practices with fewer than 500 patients or with total adjusted payments per patient of £10 or >£500 in either year were excluded. A small number of practices whose deprivation score (the main predictor related to the study aim) was not available were also excluded. A total of 6900 practices (84.4%) were included for analysis in all 4 years. Further practices were excluded because of missing data, with numbers per analysis provided in the tables.

Descriptive statistics
Table 1 gives the distribution of values and the percentage of missing values in each year for all variables used in the analysis. Some had a normal distribution, while others were skewed.

The median of adjusted total NHS payments per patient steadily rose from £102.77 in 2013–2014 to £115.43 in 2016–2017, an increase of 12.3%.

Across all 4 years, mean practice IMD scores were consistently and markedly
higher in APMS practices than in either GMS or PMS practices, which were fairly similar. There was a progressive increase in the numbers and proportion of GMS practices, while the numbers of PMS practices declined.

Univariable analyses

To check correlations between the dependent variable and the main predictor, deprivation, Spearman correlation was used because distribution of payments was skewed. The Spearman’s rank-order correlation coefficient between payments and IMD was very small in all years (Table 2).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Normally distributed</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD); % missing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Practice list aged ≥75 years [PHE]</td>
<td>7.62 (3.14); 0.00</td>
<td>7.68 (3.18); 0.00</td>
<td>7.69 (3.20); 0.00</td>
<td>7.68 (3.28); 1.76</td>
</tr>
<tr>
<td>% Self-reported smokers [GPPS Q59]</td>
<td>17.74 (6.85); 0.22</td>
<td>16.96 (6.83); 0.27</td>
<td>16.97 (6.53); 0.08</td>
<td>16.14 (6.54); 1.38</td>
</tr>
<tr>
<td>% Practice list on hypertension register</td>
<td>13.96 (3.59); 1.31</td>
<td>14.04 (3.59); 1.31</td>
<td>14.06 (3.59); 1.31</td>
<td>14.04 (3.75); 3.43</td>
</tr>
<tr>
<td>% Practice list on diabetes register</td>
<td>6.47 (1.84); 1.31</td>
<td>6.65 (1.90); 1.38</td>
<td>6.84 (1.96); 1.31</td>
<td>6.95 (2.08); 3.43</td>
</tr>
<tr>
<td>% Practice list on CHD register</td>
<td>3.32 (1.14); 1.31</td>
<td>3.28 (1.13); 1.39</td>
<td>3.23 (1.12); 1.31</td>
<td>3.19 (1.13); 3.43</td>
</tr>
<tr>
<td>% Practice list on stroke register</td>
<td>1.69 (0.65); 1.31</td>
<td>1.71 (0.65); 1.40</td>
<td>1.72 (0.66); 1.31</td>
<td>1.73 (0.68); 3.43</td>
</tr>
<tr>
<td>Skewed distribution</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (IQR); % missing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total payments per patient [NHS Digital]</td>
<td>102.77 (92.31–118.81); 0.00</td>
<td>106.09 (96.69–122.47); 0.00</td>
<td>107.71 (97.69–123.10); 0.00</td>
<td>115.43 (103.48–132.35); 0.00</td>
</tr>
<tr>
<td>List size [QOF]</td>
<td>6509 (3991–9825); 0.00</td>
<td>6618 (4076–9924); 0.00</td>
<td>6721 (4182–10 095); 0.00</td>
<td>6947 (4317–10 354); 0.00</td>
</tr>
<tr>
<td>% Black ethnicity [GPPS Q52]</td>
<td>0.49 (0–3.71); 4.04</td>
<td>0 (0–3.13); 4.22</td>
<td>0 (0–3.38); 4.23</td>
<td>0.64 (0–3.51); 2.11</td>
</tr>
<tr>
<td>% South Asian ethnicity [GPPS Q52]</td>
<td>1.09 (0–4.90); 3.47</td>
<td>1.10 (0–4.94); 3.66</td>
<td>1.32 (0–5.73); 2.26</td>
<td>1.35 (0–5.64); 3.69</td>
</tr>
<tr>
<td>% Permanently sick or disabled [GPPS Q53]</td>
<td>4.00 (2.22–6.44); 0.73</td>
<td>3.86 (2.13–6.33); 0.78</td>
<td>3.78 (2.02–6.09); 0.32</td>
<td>3.69 (2.01–6.02); 1.70</td>
</tr>
<tr>
<td>Categorical values</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Numbers (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contract type: GMS</td>
<td>4141 (55.97)</td>
<td>GMS 4205 (56.83)</td>
<td>GMS 4692 (63.41)</td>
<td>GMS 5175 (71.73)</td>
</tr>
<tr>
<td>PMS</td>
<td>3198 (43.22)</td>
<td>PMMS 3005 (60.61)</td>
<td>PMMS 2515 (53.99)</td>
<td>PMMS 2029 (28.12)</td>
</tr>
<tr>
<td>APMS</td>
<td>60 (0.81)</td>
<td>APMS 189 (2.55)</td>
<td>APMS 192 (2.59)</td>
<td>APMS 10 (0.14)</td>
</tr>
<tr>
<td>IMD scores by contract type: GMS</td>
<td>22.59 (11.67)</td>
<td>22.54 (11.61)</td>
<td>22.86 (11.72)</td>
<td>22.73 (11.76)</td>
</tr>
<tr>
<td>PMS</td>
<td>24.50 (11.76)</td>
<td>24.42 (11.72)</td>
<td>24.12 (11.59)</td>
<td>24.84 (11.49)</td>
</tr>
<tr>
<td>APMS</td>
<td>35.26 (11.48)</td>
<td>31.56 (13.09)</td>
<td>31.49 (13.19)</td>
<td>30.35 (12.51)</td>
</tr>
</tbody>
</table>

Logarithmic transformations

As the dependent variables, payments, were positively skewed in all 4 years, natural logarithm transformations were undertaken. This transformation generally improved normality, and the kurtosis of the dependent variable was reduced but remained high. Although the distribution of the main independent variable, IMD score, was not skewed, this was also transformed to facilitate interpretation of the effect size (avoiding the need to interpret geometric means). The remaining independent variables, the confounders, were not transformed as some were not continuous, and the primary interest was whether they had a predictive effect to justify inclusion in the model rather than interpreting their effect sizes.

Multivariable analyses

Before running the multilevel regression, single-level cross-sectional linear regressions were undertaken for each of the 4 years to check for multicollinearity and to ascertain what proportion of the variation in payments could be explained by each deprivation decile between 2013–2014 and 2016–2017 (Figure 1).
by the set of explanatory variables available (Table 3). The adjusted $R^2$ slowly increased across the 4 years, from 0.41 in 2013–2014 to 0.43 in 2016–2017. There was no multicollinearity. However, it cannot be assumed that independent variables will predict similarly in both cross-sectional and longitudinal models.

Organisational variables. These were list size, WTE doctors/1000 patients, WTE practice nurses/1000 patients, and NHS contract type.

Population variables. These were percentage black ethnicity, percentage with chronic condition, percentage claiming disability, and urban versus rural.

Sensitivity analysis
Two sensitivity analyses were run:

1. There was a possible negative impact of the high kurtosis of the dependent variables (ranged from 12.9 to 14.2 across the 4 years). After removing practices with payments that were outside the 1% and 99% percentiles (28 practices or 0.04%), the kurtosis values were reduced across the 4 years (ranged from 9.0 to 10.1). When the model was re-run, IMD was no longer significant ($P = 0.061$) (Table 4). Nine of the 17 confounders were significant, although their coefficients were very again small. The newly significant predictors were percentage on coronary heart disease register, percentage on stroke register, and percentage on diabetic register, but WTE practice nurses/1000 patients and percentage claiming disability were no longer significant.

2. It was assessed whether the predictive effects of deprivation on variations in payment slopes would be different between the two main contract types, GMS and PMS. The predictive effect of IMD was slightly greater in PMS practices (coefficient = 0.0089, compared with 0.0045 in GMS, and 0.0060 in all practices), but the effect sizes remained small (Table 4).

Intraclass correlations
The ICC from the null model measured the proportion of the variation in practice payments at the practice level. This was 0.81, dropping to 0.75 when excluding outliers and was fairly consistent between GMS and PMS practices (0.82 and 0.77, respectively) (Table 4).

DISCUSSION
Summary
In relation to the aim of the study, the principal finding was that practice IMD score was a very weak predictor of longitudinal variations in total payments’ slopes (0.06% more per 10% increase in IMD). Eight of the 17 confounder variables were...

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of practices</td>
<td>5333</td>
<td>6305</td>
<td>6899</td>
<td>8395</td>
</tr>
<tr>
<td>Adjusted R² value</td>
<td>0.41</td>
<td>0.41</td>
<td>0.42</td>
<td>0.43</td>
</tr>
<tr>
<td>Log IMD score, beta-coefficient</td>
<td>(0.033 to 0.070)</td>
<td>(0.052 to 0.086)</td>
<td>(0.078 to 0.113)</td>
<td>(0.070 to 0.103)</td>
</tr>
<tr>
<td>Log IMD score, P-value</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

*aDependent variable: logarithm of adjusted total payments per registered patient. IMD = Index of Multiple Deprivation (1 = most deprived and 10 = least deprived).

Table 4. Results of mixed-effects multilevel linear regression

<table>
<thead>
<tr>
<th>Variable</th>
<th>Main analysis</th>
<th>Sensitivity analysis excluding practices outside the 1st and 99th centiles (extreme outliers)</th>
<th>Sensitivity analysis GMS/PMS practices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations Groups (practices)</td>
<td>25,614</td>
<td>25,333</td>
<td>15,745/9,359</td>
</tr>
<tr>
<td>Log IMD score, beta-coefficient</td>
<td>0.060</td>
<td>0.060</td>
<td>0.045/0.089</td>
</tr>
<tr>
<td>Log IMD score, P-value (95% CI)</td>
<td>0.002 (0.010)</td>
<td>0.041 (0.062)</td>
<td>0.034/0.016</td>
</tr>
<tr>
<td>Log IMD score effect size: change in slope over 4 years (%) if variable increases by 10%</td>
<td>0.060%</td>
<td>0.029%</td>
<td>0.045/0.089</td>
</tr>
<tr>
<td>ICC</td>
<td>0.81</td>
<td>0.75</td>
<td>0.82/0.77</td>
</tr>
</tbody>
</table>

*aDependent variable = adjusted total NHS payments per patient. *Some practices changed contract type during the study and were included in both the GMS and PMS analyses only during the years during which they had that contract. This accounts for why the sum of each group exceeded the total number of groups, while the sum of each group’s observations did not exceed the total number of observations. GMS = General Medical Services. ICC = intraclass correlation. IMD = Index of Multiple Deprivation (1 = most deprived and 10 = least deprived). PMS = Personal Medical Services.

Predictors, but had very small effects on the slopes, suggesting that their significance was mainly due to the large sample size. Unfortunately, none of the 18 independent variables (IMD and the 17 confounding variables) that were used predicted much of the variability between the payments’ slopes. There was no multicollinearity between the independent variables.

Strengths and limitations

This study’s strengths are that it covered the whole of England, using recent time-matched variables in a model focused on a specific research question. The study’s findings are consistent with both the authors’ previous study, which found that practice IMD scores had a very weak positive predictive effect on cross-sectional variations in NHS practice payments in 2013–2014 and 2014–2015, and a cross-sectional study of Scottish general practices that found no association between total practice funding and deprivation.

It could be argued that the dependent variable should have included only capitation-based payments. However, performance-related payments (including QOF and all enhanced services) accounted for only 21.9% of total payments after deductions in 2016–2017 (while capitation accounted for 59.5%). Using total payments allows for the possible predictive effect of deprivation on variations across all components of payment. Although variations between practices in performance have narrowed significantly since the introduction of QOF, it cannot be assumed that these ‘gaps’ have completely disappeared. The authors’ previous study found a weak negative correlation between IMD scores and enhanced service payments levels.

The GP Patient Survey has low response rates, raising questions about its usefulness in quantitative studies. However, the methodology for the GP Patient Survey now uses ‘proportionately stratified, unclustered samples drawn from each practice,’ and data are weighted to account for unequal probability of selection, differences between responders and non-responders, and the demographic characteristics of the eligible population.

Comparison with existing literature

The findings of the current study are consistent with those of the authors’ previous study, in which variables related to population health needs were poor predictors of cross-sectional variations in practice payments. They are also consistent with a cross-sectional spatial analysis, which showed that primary care funding had large variability between regions in England and had only a modest association with morbidity, and that the correlation between chronic morbidity and funding was very weak across the country.

No publications were found that described longitudinal analyses of variations in UK general practice funding.

Implications for research and practice

The findings raise the question of what the payment system is intended to achieve and, indeed, what the NHS’s priorities are, given the persistent and wide disparity in health outcomes between the more and less privileged. The NHS Five Year Forward View promises to ‘Continue to increase investment in GP services…’ but it is not clear how this will be allocated.
What predicts most of the variations in payments, both cross-sectionally and longitudinally (that is, slopes over time), between practices is not known. If practice deprivation scores are not included in the Carr–Hill formula, then further investigation needs to identify suitable measures that will enable the weighting formula to better reflect the health needs of practice populations, as well as the associated differences in workload.

Not only does the existing NHS payment formula for general practices operate less than optimally to reflect practice population health needs, but it also has little redistributive potential and is unlikely to lead to any substantial narrowing of gaps in funding between practices with differing workloads due to the characteristics of their populations. The continuing absence of any measure of deprivation in the Carr-Hill formula means that some practices, particularly those working with socioeconomically disadvantaged populations, are not currently or imminently likely to receive the necessary additional funding they require to handle a greater workload. The revised funding formula introduced for the 2019–2020 financial year, with the promise of ‘... better reflect deprivation’, still has no specific measures of deprivation.

Weighting is an established mechanism for adjusting practice payments to take into account differences between practice populations. Before undertaking weighting, however, two key issues must be considered:

1. What components of payments should be weighted: capitation, performance, or both? Capitation broadly predicts the scale of need, while performance corresponds to the level of activity. Performance achievement targets are endpoints. However, practices working in deprived areas usually start further away from these endpoints than those in more affluent areas. Therefore, it would seem appropriate that performance payments should include some measure of the ‘distance travelled’.

2. What factors should be included in weighting calculations? These factors need to quantify population health needs. As socioeconomic deprivation is an important predictor of health needs, it is logical to include either IMD scores or a factor reliably and strongly correlated with deprivation.

Only once these issues have been addressed will primary care be adequately supported to play its important role in reducing health inequalities, one of the goals set out in the NHS Five Year Forward View.

Funding
This work had no dedicated funding. No funding organisations had a role in the design and conduct of the study, in the collection, analysis, and interpretation of the data, or in the preparation, review, or approval of the manuscript.

Ethical approval
Not applicable.

Provenance
Freely submitted; externally peer reviewed.

Competing interests
The authors have declared no competing interests.

Discuss this article
Contribute and read comments about this article: bjgp.org/letters

