BMJ Open

Heterogeneous association of health with patient and general practice characteristics by region, age and chronic condition: pooled cross-sectional study of patient-level data from England

Yan Feng ![ORCID],[1] Hugh Gravelle,[2] Vladimir Sergeevich Gordeev[1]

[1]Centre for Evaluation and Methods, Wolfson Institute of Population Health, Queen Mary University of London, London, UK
[2]Centre for Health Economics, University of York, York, UK

Correspondence to
Dr Yan Feng;
yan.feng@qmul.ac.uk

ABSTRACT

Objectives To investigate the association of individual patient health with their characteristics, including income deprivation, ethnicity and gender, with the size, quality and staffing of their general practices, and how these associations and income-related health inequality vary across chronic conditions, regions and age bands.

Design Using observational pooled cross-sectional individual patient survey data linked with data on general practice clinical quality and staffing and deprivation at Lower-Layer Super Output Area level. Ordinary least-squares multiple regression models of patient health estimated on the full sample and on eight condition-specific, nine region-specific and six age-specific subsamples. Three concentration indices embodying different value judgements summarise income-related health inequality in the full sample and subsamples.

Setting Primary care in England.

Participants Over 1 million adult patients in 6426 general practices in 2015/2016 and 2016/2017.

Primary outcome measures Patient-reported health (the 5-level EQ-5D version or EQ-5D-5L).

Results Patients who are younger, male, more satisfied with their practice, have fewer chronic conditions and live in less-income or education-deprived areas report better health. White ethnicity is associated with worse health up to age 64 and better health from age 65, with better health in five of the eight chronic condition samples, and in the regional samples except for London and Yorkshire and Humber regions. Practice clinical quality is positively associated with health in the full sample but only in 4 of the 23 subsamples. Income-related health inequality is worst for patients with a mental health problem, residents in the northwest and northeast regions and is greatest for those aged 55–64. The three concentration indices are highly positively correlated across chronic condition and age-band samples. One index has a much weaker correlation relationship with the other two indices in the region-specific samples.

Conclusion Income-related health inequality and the associations of health with patient and practice

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ A large random individual patient sample enables comparison of the associations of health with patient and general practice characteristics across subsets of the English population.
⇒ The health measure (EQ-5D-5L) is a generic measure of health widely used in observational studies, population health surveys and clinical trials.
⇒ Income-related health inequality is measured with three summary measures embodying different value judgements.
⇒ The data are a pooled cross-section, rather than a panel, so that it is not possible to control for unobserved persistent patient and practice characteristics correlated with health.
⇒ Deprivation is measured at small area level and so measures individual deprivation with error, possibly resulting in estimated associations of health, with deprivation being biased towards zero.

characteristics are heterogeneous by patients' chronic condition, age and region.

INTRODUCTION

There is a large literature attesting to the association of greater deprivation with worse health in England[1] and many other countries.[2 3] Fewer studies examine how the association varies across subsets of the population. Marmot *et al*[1] report that the association of childhood obesity rates with deprivation quintiles was greatest in London and smallest in the southwest region and that the mortality gradient by social class is greater in the northwest and northeast regions and smallest in the east, southeast and southwest regions. Doran *et al*[4] found that the social class difference in self-reported general health for those aged 25–64 was greater in London than the other

English regions. Schneider *et al*[5] find that deprivation-related health inequality emerges around age 35 and peaks for those aged 60–64.

In this article, we examine the association of patient health with deprivation and other patient and general practice characteristics using a large (1 million+) sample of individuals who responded to the General Practice Patient Survey (GPPS). The patient-level GPPS data include the EQ-5D-5L health measure and a rich set of patient characteristics, including age, gender, ethnicity, chronic conditions, smoking status and patient views on their experience in their general practice. We also have information on income and educational deprivation in the nearly 33 000 small areas in England in which patients live.

General practitioners (GPs) provide preventative care and lifestyle advice, manage chronic conditions and act as gatekeepers to non-emergency healthcare. Visits to general practices account for around 90% of patient contacts with the National Health Service (NHS). It is therefore plausible that the staffing and clinical quality of general practices will be associated with patient health[6] and there is some evidence that increased supply of GPs will improve patient health.[7] There are marked differences in the number of GPs across areas and more deprived areas have fewer GPs.[8] We link the individual patient GPPS data to administrative data on the characteristics of their 6426 practices, including measures of staffing and clinical quality, to examine how patient health varies with the characteristics of their practices after controlling for patient characteristics.

We examine the association of patient health with deprivation and other characteristics using multiple OLS regressions of patient health on patient and practice characteristics. We capitalise on the large 1 million sample to examine how these associations vary across eight condition-specific, nine region-specific and six age band-specific samples.

We also use three summary measures of income-related health inequality to examine if comparisons of the degree of income-related health inequality in population groups are sensitive to the different value judgements about absolute and relative inequality embodied in the measures. We use the measures to compare income-related health inequality across eight condition-specific, nine region-specific and six age-band-specific samples.

INSTITUTIONAL BACKGROUND
Primary medical care in the English NHS is provided by general practices, most of which are owned and run by GPs. General practices manage chronic conditions and are the first point of contact with the NHS for most patients for non-emergency medical care. Primary care is free at the point of use, apart from a small charge for around 10% of prescriptions. In 2016/2017, there were 7392 general practices, with an average list of 7850 patients and 3.8 full-time equivalent (FTE) GPs. Practices are paid by a mix of need-weighted capitation, lump sums, items of service fees and quality incentives. They are reimbursed for the costs of their premises but must cover all other expenses, such as the employment of non-partner salaried GPs, nurses and support staff. General practices were grouped in just over 200 Clinical Commissioning Groups (CCGs) which received needs-weighted capitation budgets from the Department of Health and Social Care to purchase healthcare from secondary care providers. CCGs also commissioned some services directly from general practices.

DATA
General Practice Patient Survey
The GPPS is an annual survey of patients in all general practices in England. The aim is to provide a representative sample of adult patients aged 18 and over who are registered with a GP in England. In each financial year (April–March), a new random sample of 5% patients aged 18 or over in each general practice is sent the GPPS questionnaire. The survey was distributed in two waves (July–September and January–March) up to 2015/2016 and in one wave (January–March) from 2016/2017. Data collection was mainly done by postal paper questionnaires with options to respond online or over the telephone. The mailout strategy was changed in 2015/2016 with redesigned cover letters in each full survey mailing and a postcard reminder being sent to all sampled patients 1 week after the first survey pack mailing. These design changes led to increases in the response rate from 32.5% in 2014/2015 to 38.9% in 2015/2016 and 37.6% in 2016/2017, and to changes in the mix of respondents.[9] We control for the characteristics of the individual respondents in the patient-level regression modelling.

The GPPS collects information on individual patient gender, age, ethnicity, smoking status, chronic conditions, views on practice opening hours, experience in making appointments and overall experience with the practice. The EQ-5D-5L health instrument was included in the GPPS from 2012/2013 to 2016/2017. Patients were asked to report problems on five dimensions of their health (mobility, self-care, usual activities, pain/discomfort, anxiety/depression) in five severity levels (none, slight, moderate, severe, extreme problems). We calculate the EQ-5D-5L index for a patient by applying a value set to the reported severity levels for the five health dimensions.[10] The resulting cardinal measure has a maximum of 1 (no problems on any dimension) and a minimum of −0.594 (extreme problems on all dimensions). We rescale the EQ-5D-5L index (subtracting −0.594 and dividing by 1.594=1−(−0.594)), so that the health measure has a more intuitive maximum of 1 and a minimum of 0. The rescaling affects the magnitude of the coefficients in the regression models but not their statistical significance.

Indices of Multiple Deprivation

The English Indices of Multiple Deprivation (IMD) 2015 are measures of the relative deprivation of residents of 32 844 Lower-Layer Super Output Areas (LSOAs) in England. LSOAs have an average of around 1500 inhabitants. The IMD income score is the proportion of LSOA residents who are receiving social security payments for which they qualify because they have a low income. We measure income as *Income*=1 − *IMD income score* so that higher values indicate LSOAs with a smaller proportion of residents having low income.

Given the evidence on the effects of education on health,[11 12] we use the LSOA rank of the IMD Education Deprivation Index to construct a percentile measure of education. The LSOA rank runs from 1 (most deprived) to 32 844 (least deprived). We divide the LSOA rank by 32 844, so that the rescaled variable is akin to a percentile, ranging from 1/32 844 to 1. A higher value indicates that the respondent lives in an area with a better-educated population and so is more likely to be better educated.

Quality and outcomes framework

The quality and outcomes framework (QOF) is an annual pay-for-performance scheme introduced in 2004/2005 to improve the management of chronic conditions by general practices. Although the scheme was voluntary, almost all English practices took part. Payment to a general practice was linked to points earned for their achievement of quality indicators. In 2016/2017, practices could score up to 559 points, being paid, on average, £165 per point. The bulk of the points were for indicators of clinical quality of care for chronic conditions, such as the proportion of patients with hypertension whose blood pressure was controlled.

We use achievement of QOF clinical indicators to measure the clinical quality of general practices. The population achievement rate for a QOF clinical indicator for a condition is the proportion of patients with the condition for whom the indicator was achieved. We measure overall practice population achievement as the average population achievement rates over all conditions, weighted by the maximum points available for each indicator. Clinical quality for a specific condition is the maximum points weighted average of population achievement rates for the indicators relevant to that condition.

We also use practice-level information on the numbers of FTE nurses and other staff, and the FTE numbers, gender mix, age bands and country of qualification of GPs.

Data linkage

We merged the GPPS data with NHS Digital QOF data and NHS Digital GP workforce data by practice and year. The linked data set was then merged with the deprivation data using the LSOA of the practice. Online supplemental appendix 1 reports the variables extracted from each of the four data sets.

Samples

To avoid the impacts of GPPS design changes in 2015, we use data for 2015/2016 and 2016/2017 which were the last years in which the EQ-5D-5L health instrument was included in the GPPS. There were 1 620 636 GPPS respondents in 2015/2016 and 2016/2017. We have data on the EQ-5D-5L health measure, IMD, practice population achievement rate, list size, rurality, region and wave for an incomplete case sample of 1 494 866 individuals, and full information on all patient and practice covariates for 1 089 398 complete cases. We also exploit the large size of the GPPS data set to construct three sets of complete case subsamples for eight chronic conditions, nine English Government Office Regions and six age bands.

Details of data sources and construction of variables are provided in online supplemental appendix 1. Definitions of the eight chronic conditions are provided in online supplemental appendix 2.

METHODS
Patient health models

We examine the association of patient and practice characteristics with patient health using pooled cross-section ordinary least-square (OLS) regression. OLS is the most common regression method for the analysis of EQ-5D values despite the bounded nature of the health measure.[13] Pullenayegum *et al*[14] suggest that OLS with robust SEs is less likely to lead to biased results than methods such as Tobit or censored least absolute deviation intended to deal with the apparent truncation of EQ-5D at its upper bound of 1.

The data are pooled across three GPPS samples (the July–September and January–March waves in 2015/2016 and the single wave in January–March in 2016/2017). Each patient i, on the list of practice g, resident in region a, in year t, wave w, is observed only once, in the year t and wave w in which they responded to the GPPS. For the full sample, the OLS model is

$$h_{igatw} = \beta_0 + \beta_y y_i + \sum_k \beta_k^P x_{ik} + \sum_j \beta_j^G x_{gtj}$$
$$+ \sum_a \beta_a^A D_i^a + \beta^t D_i^t + \beta^w D_i^w + \varepsilon_{igatw} \quad (1)$$

where h_{igatw} is the EQ-5D-5L health index for patient i in practice g, region a, in the year t and wave w in which patient i responded to the GPPS. y_i is income for patient i, and x_{ik} are other patient characteristics for patient i, x_{gtj} are characteristics of practice g in year t, D_i^a is an indicator for the region a in which patient i lives, D_i^t and D_i^w are indicators for the survey year t and wave w in which patient i responded to the GPPS.

We also examine how these associations vary with patients' chronic condition, region and age. Rather than adding a large number of interaction terms to the full sample model (1), we split the sample into nine condition-specific, eight regional-specific and six age band-specific subsamples and estimate a similar OLS model on each subsample. For the condition-specific models, we include

both condition-specific and overall clinical quality as measured by QOF population achievement.

We use STATA/MP 18.0 for all analyses and cluster robust SEs by general practice.

Robustness checks

In addition to the OLS models, we use the Stata *mixed* command to estimate multilevel mixed-effects models in which the error term ε is decomposed into three parts reflecting the nesting of patients within practices and practices within CCGs.

The complete case sample (n=1 089 398) has full data on all patient and practice explanatory variables, while the incomplete case sample has 1 494 866 observations. Practice workforce variables have the highest missing percentages (GP gender: 14.4%; GP age and country of qualification: 12.3%). The patient variables with the highest missingness are the number of conditions (7.5%), experience with making appointments (4.4%), ethnicity (1.8%), gender and age (1.4%). Online supplemental appendix 3 reports the percentage of missing values for all variables.

We create five imputed data sets (STATA command: *mi impute chained*) using logistic and ordered logistic regressions to impute missing binary and ordered categorical variables, and linear regression for missing continuous variables.[15] We use OLS to estimate the health model on each of the five imputed data sets (STATA command: *mi estimate: reg*) and compare the means of the estimated coefficients with the coefficients from the model estimated on the complete case sample.

Measures of income-related health inequality

We use the Stata *conindex*[16] package to compute summary measures of income-related health inequality for each set of subsamples to examine how income-related inequality varies across patients with different health conditions, across regions and across age bands. We compare three concentration index measures of income-related health inequality to investigate the extent to which comparison of income-related health inequality across different patient groups is sensitive to the value judgements embodied in the different measures.[17–19]

The measures are based on the covariance $Cov\left(h_i, R_i\right)$, where h_i is the health of individual i and R_i—the fractional rank of i in the income distribution—is the proportion of individuals who have an income no greater than i. All the measures indicate pro-rich health inequity if those with higher incomes have better health ($Cov\left(h_i, R_i\right) > 0$).

The *absolute concentration index* (ACI)

$$\text{ACI} = 2Cov\left(h_i, R_i\right) \qquad (2)$$

is not affected by equal absolute change in health for all individuals but would be affected by an equal proportion change in health for all.

The s*tandard concentration index* (SCI) is

$$\text{SCI} = 2Cov\left(\frac{h_i}{\bar{h}}, R_i\right) = \frac{2Cov\left(h_i, R_i\right)}{\bar{h}} = \frac{\text{ACI}}{\bar{h}} \qquad (3)$$

where $\bar{h}$ is mean health. It is akin to the Gini coefficient measure of income inequality and embodies a concern about relative inequity in that changes the *share* of total health enjoyed at different income ranks will change income-related health inequality. In particular, an equal absolute increase in health for all individuals will reduce the SCI since it will increase mean health but not alter $Cov\left(h_i, R_i\right)$. An equal proportionate increase in health will not affect the SCI.

The *partial concentration index* (PCI)[20]

$$\text{PCI} = \beta_y \text{SCI}\left(y_{igatw}, R_i\right) \qquad (4)$$

is relevant if we wish to focus solely on the income-related health inequality which is due to the combined effect of income on health β_y and the inequality of income as measured by the concentration index of income against income rank $SCI\left(y_{igatw}, R_i\right)$, otherwise known as the Gini coefficient.

We examine the extent to which the three measures of income-related health inequality correlated with each other by calculating their Pearson correlation coefficients in the condition-, region- and age-specific samples.

Patient and public involvement

Patients or the public were not involved in the design, conduct, reporting or dissemination plans of this research.

RESULTS

Summary statistics

Table 1 shows the summary statistics for the complete case sample (n=1 089 398). 31.9% have at least one chronic condition, and 30.8% are multimorbid. 88.6% of respondents report their ethnicity as white ethnicity. The distribution of EQ-5D-5L index values for the complete case sample is reported in online supplemental appendix figure A1.

Health and patient characteristics

Table 2 shows the results from OLS health model for the complete case pooled sample. Results from the multiple imputation model and mixed-effects model are very similar to the pooled OLS model (online supplemental appendix 4). Mixed-effects model suggests that the CCG-level and practice-level random effects accounted for very small proportions of the residual variance.

Health is better, ceteris paribus, for patients who have fewer reported health conditions, have better experience with their practices, are younger, do not smoke and live in small areas with less income and education deprivation. The positive marginal effect for income is from a quadratic function of income which performed better in terms of the Bayesian information criteria than linear and cubic functions, though all specifications had the same adjusted $\bar{R}^2$ to four decimal places. Women have worse health except for those who report a mental health condition. In the pooled sample, those of white ethnicity have worse health, though in five of the condition-specific samples they have better health.

Table 1 Summary statistics for pooled estimation sample

Patient characteristics		
EQ-5D-5L index, mean (SD)	0.794	(0.235)
Rescaled EQ-5D-5L index, mean (SD)	0.871	(0.148)
Health conditions, n (%)		
No health condition	406 900	(37.4%)
One health condition	347 376	(31.9%)
Multimorbid (>1 health condition)	335 122	(30.8%)
Overall experience, n (%)		
Very/fairly poor or neutral	130 354	(12.0%)
Very/fairly good	959 044	(88.0%)
Make appointment experience, n (%)		
Very/fairly poor	240 834	(22.1%)
Very/fairly good	848 564	(77.9%)
Satisfaction with open hours, n (%)		
Very/fairly poor, neutral don't know	221 536	(20.3%)
Very/fairly good	867 862	(79.7%)
Gender, n (%)		
Male	485 481	(44.6%)
Female	603 917	(55.4%)
Age, n (%)		
18–44 years old	281 193	(25.8%)
45–64 years old	417 856	(38.4%)
≥65 years old	390 349	(35.8%)
Smoking status, n (%)		
Occasional/regular smoker	147 631	(13.6%)
Never/former smoker	941 767	(86.5%)
Ethnicity, n (%)		
Non-white ethnicity	123 842	(11.4%)
White ethnicity	965 556	(88.6%)
Income, mean (SD)	0.860	(0.101)
Education deprivation, mean (SD)	0.516	(0.283)
Practice characteristics		
Quality and outcomes framework population achievement, mean (SD)	0.786	(0.074)
List size (1000s), mean (SD)	8.052	(4.694)
GPs/1000 patients, mean (SD)	0.559	(0.222)
Nurses/1000 patients, mean (SD)	0.259	(0.147)
Other staff/1000 patients, mean (SD)	1.261	(0.512)
		Continued

Table 1 Continued

GPs' age, mean (SD)		
Proportion up to 34 years	0.143	(0.189)
Proportion 35–49 years	0.459	(0.273)
Proportion≥ 50 years	0.399	(0.284)
GPs' gender, mean (SD)		
Proportion of female	0.486	(0.250)
Proportion of male	0.514	(0.250)
GPs' qualification, mean (SD)		
Proportion of UK qualified	0.754	(0.318)
Proportion of Europe (not-UK) qualified	0.043	(0.122)
Proportion of non-European qualified	0.203	(0.303)
Rural/Urban area, n (%)		
Rural area	190 820	(17.5%)
Urban area	898 578	(82.5%)
Region, n (%)		
East Midlands	86 793	(8.0%)
Eastern	115 231	(10.6%)
London	159 319	(14.6%)
Northeast	62 695	(5.8%)
Northwest	167 431	(15.4%)
Southeast	163 080	(15.0%)
Southwest	116 811	(10.7%)
West Midlands	118 423	(10.9%)
Yorkshire and Humber	99 615	(9.1%)
Survey information		
GPPS wave 1 (July–September), n (%)	274 000	(25.2%)
GPPS wave 2 (January–March), n (%)	815 398	(74.8%)
2015/2016, n (%)	551 288	(50.6%)
2016/2017, n (%)	538 110	(49.4%)

Complete case sample. n=1 089 398. Income=1 – IMD income score. IMD income score is the proportion of LSOA residents drawing social security payments on the grounds of low income. IMD education deprivation rank of LSOA of residence (higher rank is less deprivation). White ethnicity: English, Welsh, Scottish, Northern Irish, British, Irish, Gypsy, Traveller or Irish Traveller, any other white background.
GP, general practitioner; GPPS, General Practice Patient Survey.

There are marked regional health differences in the pooled sample after controlling for other patient and practice characteristics. Residence in the London, southeast and eastern regions is associated with the better health, while residence in the northeast and northwest regions is associated with worse health.

Online supplemental appendix 5 reports results from OLS regressions for the eight condition-specific samples.

Table 2 Health models: pooled sample

	Coefficient	SE
Patient characteristics		
One condition	−0.0630***	0.0003
Multimorbid	−0.1672***	0.0005
Experience: very/fairly good	0.0140***	0.0005
Make appointment: very/fairly good	0.0110***	0.0004
Open hours: very/fairly good	−0.0001	0.0003
Female	−0.0141***	0.0003
Age 45–64	−0.0152***	0.0003
Age ≥65 years old or above	−0.0217***	0.0004
Never/former smoker	0.0302***	0.0004
White ethnicity	−0.0054***	0.0005
Income	0.1388***	0.0034
Education	0.0099***	0.0009
Practice characteristics		
Quality and outcomes framework (population achievement)	0.0056**	0.0021
List size (1000s)	0.0001*	0.0000
GPs/1000 patients	0.0012	0.0007
Nurses/1000 patients	−0.0029*	0.0013
Other staff/1000 patients	−0.0004	0.0004
GPs' age 35–49	0.0007	0.0009
GPs' age ≥50 years	0.0022*	0.0009
Proportion male GPs	−0.0021**	0.0007
GPs Europe qualified	−0.0001	0.0013
GP non-Europe qualified	−0.001	0.0006
Urban	−0.0034***	0.0004
East Midlands	−0.0046***	0.0007
Eastern	0.0004	0.0006
Northeast	−0.0105***	0.0008
Northwest	−0.0084***	0.0006
Southeast	−0.0005	0.0006
Southwest	−0.0013*	0.0006
West Midlands	−0.0043***	0.0006
Yorkshire and Humber	−0.0042***	0.0007
Survey information		
Wave	−0.0005	0.0003
2016/2017	−0.0007*	0.0003
Constant	0.8468***	0.0087
R^2	0.2707	
Observations	1 089 398	

Ordinary least-square models. EQ-5D-5L∈ [0,1]. Income: average marginal effect of quadratic function of income. Robust SEs clustered on practices. Omitted categories: *no condition* for number of conditions; *very/fairly poor or neutral* for overall experience with the practice; *very/fairly poor experience with making appointment* for making appointment with the practice; *very/ fairly poor, neutral, or don't know* is the omitted category for satisfaction with practice open hours; *male* for patient gender; *18–44 years old* for patient age; *occasional/regular smoker* for patient smoking status; *non-white* for patient ethnicity; *GPs under 34 years old* for GP age; *female* for GP gender; *GPs UK qualified* for GPs qualification; *rural* for area type; *London* for region; *wave 1* for General Practice Patient Survey wave; *2015/2016* for financial year.
***p<0.001; **p<0.01; *p<0.05.
GP, general practitioner.

As highlighted in figure 1, higher income is associated with better health across all eight conditions. Women have worse health across all conditions except for those who report a mental health condition. White ethnicity is associated with better health in five of the eight chronic condition samples.

Online supplemental appendix 6 reports OLS models for the nine regional samples. Figure 2 shows that higher income and male gender are associated with better health in all nine regions. As in the pooled sample, white ethnicity is associated with worse self-reported health, except in the London and Yorkshire and Humber regions.

Online supplemental appendix 7 has OLS results from the six age-specific samples. The most obvious difference from the pooled model is that there is a clear age-related trend in the association of white ethnicity with health. As shown in figure 3, the coefficient on white ethnicity with health is negative up to 64 years old and turns positive and increasing in age across the older groups. The reduction in health for women relative to men increases with age.

Health and practice characteristics
In the pooled sample, better health is reported by patients in practices which have better overall QOF clinical quality (table 2). But in the condition-specific models better condition-specific QOF achievement is not associated with better health, except for patients with diabetes (online supplemental appendix 5). Health is better in practices with higher proportions of female GPs in the pooled sample and in five of the condition-specific samples. Patients in larger practices have better health, especially in the condition-specific samples. Staff/patient ratios are not associated with health in any of the models.

For six of the eight chronic condition samples in online supplemental appendix 5, patients report worse health if they are in practices with a greater proportion of GPs who qualified outside Europe. There is also a clear age-related trend in online supplemental appendix 7 in the association of health with the proportion of GPs who qualified outside Europe: it is positive for the youngest 18–44 patient age group, vanishes for those aged 45–54 and becomes increasingly negative in the older age groups.

Income-related health inequality
The health models indicate that health is strongly positively associated with income (see also the comparison of health across quintiles of income in online supplemental appendix 8). Table 3 reports mean health, mean income and three measures of income-related health inequality for the pooled, regional, condition-specific and age-specific complete case samples. Those with a chronic condition, unsurprisingly, report worse mean health than respondents in the pooled sample. Patients with dementia have the worst health across the condition-specific samples, followed by patients with a mental health problem. Patients with hypertension have the best mean health among those with a chronic condition. Unsurprisingly
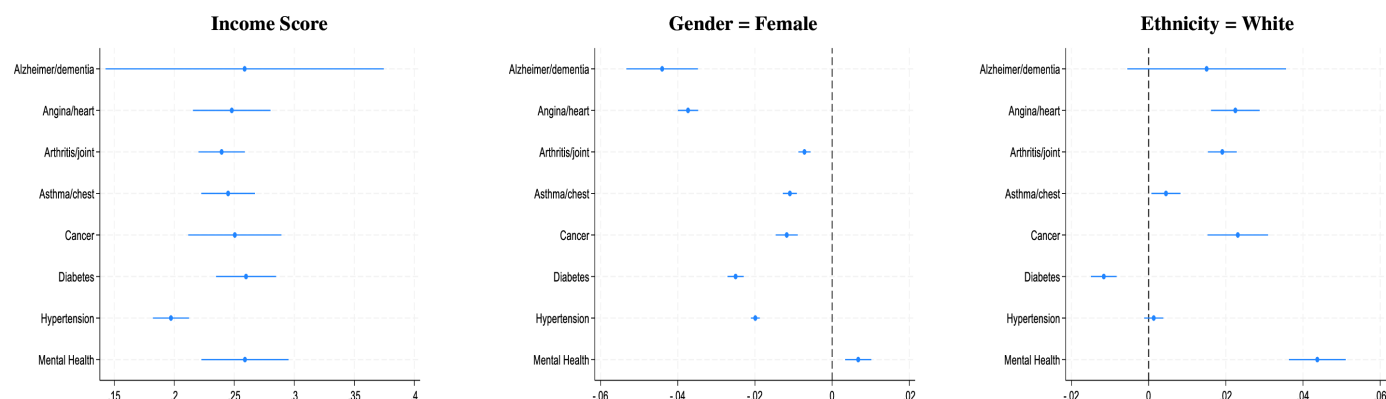
Figure 1 Plot of estimated coefficients and 95% CIs from ordinary least-square health models estimated on eight health condition-specific subsamples with a full set of covariates. Income score: average marginal effect of quadratic function of IMD income score. Full results are provided in online supplemental appendix 5. IMD, Indices of Multiple Deprivation.

mean health is lower in the older age samples. Average incomes for patients with a chronic condition are smaller than for respondents in the pooled sample, except for patients with cancer. In general, mean income is higher in older age groups.

Patients with a mental health problem had the greatest income-related health inequality on all three concentration indices. The measures agree that income-related health inequality is smallest among patients with Alzheimer/dementia, though the sample with this condition is much smaller than for any of the other conditions.

The northwest region has the largest income-related health inequality on all three measures, and the northeast region has the second largest on two of the three. London has the lowest inequality on SCI and ACI, and fourth lowest on the PCI. All three measures increase with age up to the 55–64 age group.

SCI and ACI are strongly and positively correlated across the region-specific, condition-specific and age-specific samples, but have much weaker correlations with the PCI in the region-specific sample (online supplemental appendix 9).

DISCUSSION
Principal findings
Our results suggest that the associations of individuals' health with their characteristics and the characteristics of their general practices vary across samples defined by chronic condition, region and age. This is particularly noticeable for the association of health and ethnicity. White patients have worse health than other ethnic groups in the pooled sample and in all regions except for London and Yorkshire and Humber. But they have better health in five of the eight condition-specific samples, and in the age-specific samples white ethnicity is associated with worse health up to age 64 and better health from age 65.

Practice clinical quality, as measured by the performance on the QOF, is positively associated with individual health in the full sample but only in 4 of the 23 (region-specific, age-specific, condition-specific) subsamples.

Health is positively associated with income in the pooled sample and in all the subsamples, but the strength of association varies considerably across the samples. Summary measures of income-related
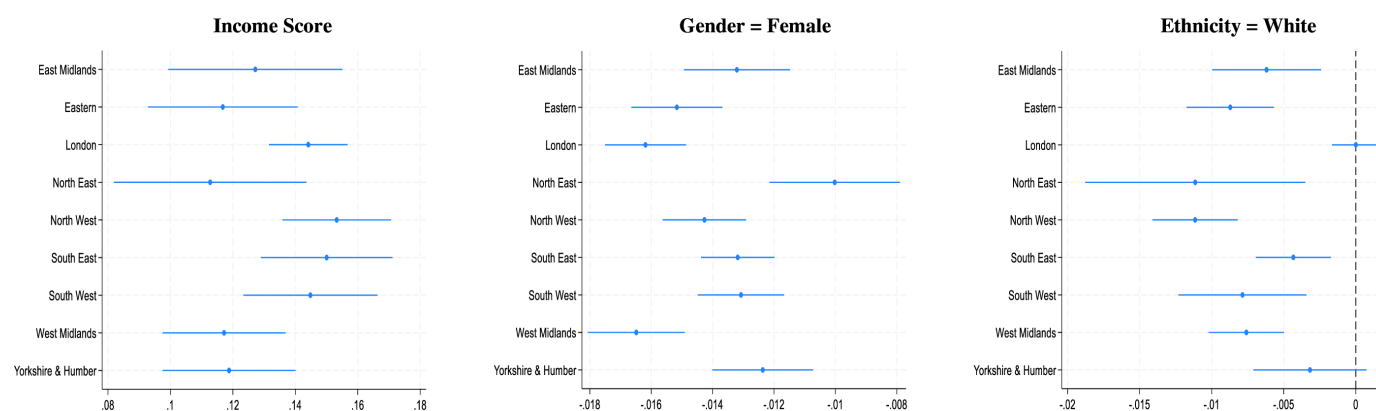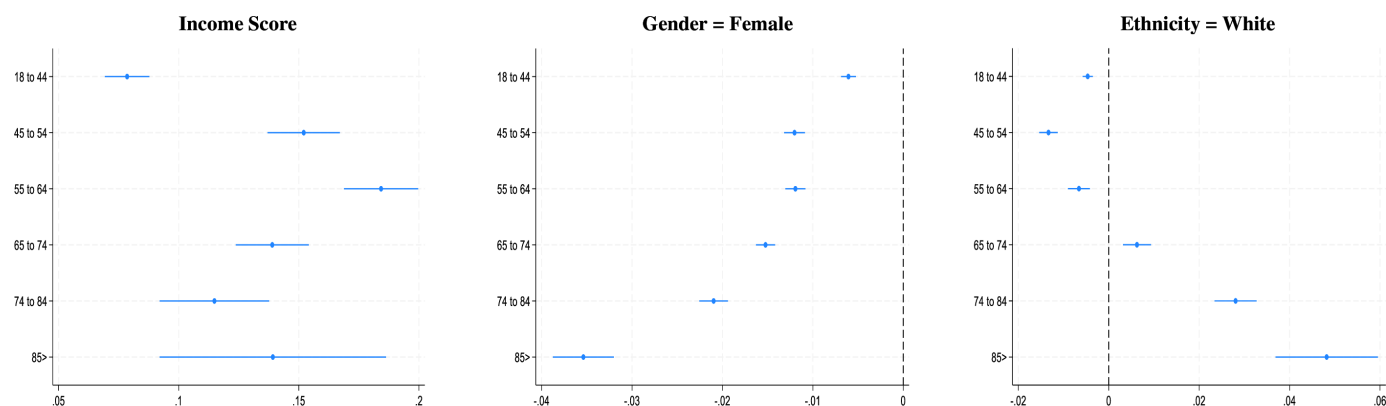

Figure 2 Plot of estimated coefficients and 95% CIs from ordinary least-square health models estimated on nine regional subsamples with a full set of covariates. Income score: average marginal effect of quadratic function of IMD income score. Full results are provided in online supplemental appendix 6. IMD, Indices of Multiple Deprivation.

Figure 3 Plot of estimated coefficients and 95% CIs from ordinary least-square health models estimated on six age band subsamples with a full set of covariates. Income score: average marginal effect of quadratic function of IMD income score. Full results are provided in online supplemental appendix 7. IMD, Indices of Multiple Deprivation.

health inequality are, in general, greater in the condition-specific groups compared with the general population, being highest for patients with a mental health problem on all three measures. The regions with the highest income-related health inequality are the northeast and northwest. According to all three summary measures of income-related health inequality, the inequality increased with age and peaked at the 55–64 age band.

Table 3 Concentration indices of health against income deprivation by condition, region and age group

Sample	Mean health	Mean income	Standard concentration indices	Absolute concentration indices	Partial concentration indices	Observations
Pooled	0.871	0.860	0.0142	0.0124	0.0077	1 089 398
Alzheimer/dementia	0.643	0.852	0.0145	0.0093	0.0109	7822
Angina/heart	0.776	0.853	0.0289	0.0224	0.0143	68 237
Arthritis joint	0.730	0.848	0.0292	0.0213	0.0144	176 803
Asthma/chest	0.813	0.847	0.0304	0.0247	0.0145	118 987
Cancer	0.815	0.869	0.0237	0.0193	0.0120	48 126
Diabetes	0.801	0.839	0.0222	0.0178	0.0147	103 975
Hypertension	0.829	0.855	0.0204	0.0169	0.0116	261 799
Mental health	0.700	0.828	0.0456	0.0319	0.0178	49 243
London	0.882	0.837	0.0104	0.0092	0.0061	159 319
East Midlands	0.871	0.872	0.0117	0.0102	0.0051	86 793
Eastern	0.880	0.888	0.0107	0.0094	0.0079	115 231
Northeast	0.850	0.827	0.0187	0.0159	0.0059	62 695
Northwest	0.856	0.832	0.0191	0.0164	0.0110	167 431
Southeast	0.883	0.900	0.0114	0.0101	0.0061	163 080
South West	0.878	0.889	0.0106	0.0093	0.0057	116 811
West Midlands	0.863	0.838	0.0142	0.0122	0.0066	118 423
Yorkshire and Humber	0.864	0.844	0.0161	0.0139	0.0070	99 615
18–44 years	0.918	0.845	0.0088	0.0081	0.0049	281 193
45–54 years	0.885	0.858	0.0189	0.0167	0.0094	193 463
55–64 years	0.865	0.862	0.0229	0.0198	0.0113	224 393
65–74 years	0.856	0.870	0.0180	0.0154	0.0081	231 210
75–84 years	0.817	0.870	0.0159	0.0130	0.0070	122 265
85 years and above	0.743	0.869	0.0147	0.0109	0.0092	36 874

Health is measured by EQ-5D-5L rescaled to $\in [0,1]$.

Strengths and limitations

A strength of this study is the sample of over 1 million patients in English general practices with data on a rich set of patient and practice characteristics. The sample is large enough to enable analysis of condition-specific, regional-specific and age band-specific subsamples to investigate whether, and how, the association of health with policy salient patient characteristics such as gender, ethnicity and income differs across patients with different medical conditions, or living in different regions, or of different ages.

Our choice of health measure (EQ-5D-5L) is widely used internationally as a generic measure of health status in population health surveys, observational studies and clinical trials. We use three summary measures of income-related health inequality (concentration indices) to investigate whether the value judgements they embody affect comparisons of inequality across subgroups of patients defined by region, condition or age.

Our data are a pooled set of cross-sectional patient samples, rather than a panel, so that we cannot use patient-level fixed- or random-effect methods to control for unobserved persistent patient characteristics correlated with health. Allowing for practice or CCG random effects made very little difference to our results. Observational cross-sectional data cannot be used to test for causal relationships though it can suggest hypotheses to be tested in stronger designs. It can also warn about the possibility that the relationship between health and possible causes, such as income, or ethnicity, may depend on other factors such as age, as we demonstrated in the subsample cross-sectional regressions.

The deprivation measures we use are for the small (1500 population) areas and so measure individual deprivation with error, possibly resulting in the estimated coefficient on deprivation being biased to zero. Finally, the GPPS stopped including the EQ-5D-5L instrument after the financial year 2016/2017. We are therefore unable to include more recent data in this study.

Interpretation of results and comparison with other studies

Many of our results about the relationship between patient characteristics and health are in line with previous literature: patients report better health if they live in less-deprived small areas, are younger, have no chronic health conditions and do not smoke. Women report worse health than men in all the samples, although not for patients with a mental health condition.

In the pooled sample and the regional-specific samples, we find that white patients report worse health than other ethnic groups, though they have worse health only in one of the condition-specific samples (diabetes). Watkinson *et al*[21] also used a sample from GPPS, covering 2014/2015 to 2016/2017, and found that white British patients had *better* EQ-5D-5L health than 15 of the other 17 ethnic groups. Our definition of white ethnicity includes patients classifying themselves as white British, Irish, Gypsy or Irish Traveller or other white background. Gypsy or Irish Travellers have substantially worse health than other ethnic groups but are a tiny proportion of our GPPS sample and do not account for the worse health of those we classified as white ethnicity. The reason for our finding that patients we classified as white ethnicity had worse health than those of non-white ethnicity is that the association of health with ethnicity varies with age and our sample has a wider age range (18 and over) compared with Watkinson *et al*[21] whose sample was those aged 55 and over. A cross-tabulation of ethnicity (white vs non-white) and health by age in our sample shows that white ethnicity is associated with worse health in younger age groups and better health in older age groups (online supplemental appendix 10). When we estimate OLS regressions (1) on the six separate age-specific samples (online supplemental appendix 7), controlling for patient covariates in addition to age and for practice covariates, the coefficient on white ethnicity is *negative* for age groups 18–44, 45–54, 55–64, and *positive* and increasing for age groups 65–74, 75–84 and over 85 (figure 3). One possible explanation is that the ethnic composition of the non-white patients varies across the age groups[22] and that there are differences in health across different non-white groups.[21]

Previous studies report that a better QOF clinical performance was associated with better health for some chronic conditions[23 24] but there is mixed evidence for mortality.[25–27] In a study with GPPS practice-level panel data for 2012/2013 to 2016/2017, results from practice fixed-effects models suggested that practice-level EQ-5D was positively associated with QOF clinical quality.[28] In this study, we find that practice clinical quality, as measured by the performance on the QOF, was positively associated with individual health in the full sample. However, condition-specific QOF indicators were not associated with better health for patients with those conditions, except for diabetes. Although previous work[7] has suggested that increased supply of GPs in an area is associated with better health, we find no association of patient health with the number of GPs per patient. An Italian feasibility study[29] suggested that specially trained primary care nurses had positive impacts on disease management for patients with cardiovascular disease (CVD), diabetes, heart failure or CVD risk. However, we found that the nurse/patient ratio was negatively associated with patient health in the pooled sample and in patients with angina or heart problems.

The finding (table 2, figures 1–3) that health is better for individuals in small areas with less deprivation is in line with the previous literature.[1 30 31] We also find (table 3) that *regions* with higher income generally had better health, though London, with the second highest mean health but third lowest mean income, was a noticeable exception.

We compared three summary measures of income-related health inequality, all based on the covariance of health and income rank but embodying different attitudes to absolute and relative differences in health (table 3). On two of the three measures, income-related

health inequality was greatest in the northeast and northwest regions, which also had the worst health and lowest income compared with other regions. In contrast to studies[1 4] which found that the health gradient with respect to social class was greater in London than other regions, we find that income-related health inequality was smaller in London than in all other regions on two out of the three inequality measures and in the middle of the ranking on the third measure. As in Schneider *et al*,[5] we find that income-related health inequality increases up to age 55–64 and then decreases.

There appear to be no previous papers comparing income-related health inequality for patients with different conditions. Table 3 suggests that, with the exception of patients with Alzheimer/dementia (who were a much smaller proportion of the sample than the other condition-specific groups), in general there was greater income-related health inequality among the condition-specific groups compared with the general population. Income-related health inequality was greatest among patient with a mental health problem on all three measures.

CONCLUSION

Income-related health inequality and the associations of health with patient and practice characteristics are heterogeneous by patients' chronic condition, age and region.

ORCID iD
Yan Feng http://orcid.org/0000-0003-4436-1126

REFERENCES

1. Marmot M. Health equity in England: the Marmot review 10 years on. *BMJ* 2020;368:m693.
2. Chetty R, Stepner M, Abraham S, *et al*. The Association Between Income and Life Expectancy in the United States, 2001-2014. *JAMA* 2016;315:1750–66.
3. Mackenbach JP, Stirbu I, Roskam A-JR, *et al*. Socioeconomic inequalities in health in 22 European countries. *N Engl J Med* 2008;358:2468–81.
4. Doran T, Drever F, Whitehead M. Is there a north-south divide in social class inequalities in health in Great Britain? Cross sectional study using data from the 2001 census. *BMJ* 2004;328:1043–5.
5. Schneider P, Love-Koh J, McNamara S, *et al*. Socioeconomic inequalities in HRQoL in England: an age-sex stratified analysis. *Health Qual Life Outcomes* 2022;20:121.
6. Starfield B, Shi L, Macinko J. Contribution of primary care to health systems and health. *Milbank Q* 2005;83:457–502.
7. Gravelle H, Morris S, Sutton M. Are family physicians good for your health? Endogenous supply and health. *Health Serv Res* 2008;43:1128–44.
8. Asaria M, Cookson R, Fleetcroft R, *et al*. Unequal socioeconomic distribution of the primary care workforce: whole-population small area longitudinal study. *BMJ Open* 2016;6:e008783.
9. Ipsos MORIb. Section 1 methodology changes in gp patient survey change log. 2016. Available: https://www.gp-patient.co.uk/surveysandreports-10-16
10. van Hout B, Janssen MF, Feng Y-S, *et al*. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health* 2012;15:708–15.
11. Cutler D, Lleras-Muney A. Education and health: evaluating theories and evidence. NBER Working Paper; 2006. Available: http://www.nber.org/papers/w12352
12. Davies NM, Dickson M, Davey Smith G, *et al*. The Causal Effects of Education on Health Outcomes in the UK Biobank. *Nat Hum Behav* 2018;2:117–25.
13. Devlin N, Parkin D, Janssen B. *Methods for Analysing and Reporting EQ-5D Data*. Cham (CH): Springer, 2020.
14. Pullenayegum EM, Tarride J-E, Xie F, *et al*. Analysis of health utility data when some subjects attain the upper bound of 1: are Tobit and CLAD models appropriate? *Value Health* 2010;13:487–94.
15. Rubin DB. Multiple Imputation After 18+ Years. *J Am Stat Assoc* 1996;91:473.
16. O'Donnell O, O'Neill S, Van Ourti T, *et al*. conindex: Estimation of concentration indices. *Stata J* 2016;16:112–38.
17. Allanson P, Petrie D. Understanding the vertical equity judgements underpinning health inequality measures. *Health Econ* 2014;23:1390–6.
18. Kjellsson G, Gerdtham U-G, Petrie D. Lies, Damned Lies, and Health Inequality Measurements. *Epidemiology (Sunnyvale)* 2015;26:673–80.
19. Heckley G, Gerdtham UG, Kjellsson G. A general method for decomposing the causes of socioeconomic inequality in health. *J Health Econ* 2016;48:89–106.
20. Gravelle H. Measuring income related inequality in health: standardisation and the partial concentration index. *Health Econ* 2003;12:803–19.
21. Watkinson R, Sutton M, Turner A. Ethnic inequalities in health-related quality of life among older adults in England: secondary analysis of a national cross-sectional survey. *Lancet Public Health* 2021;6:e145–154.
22. Office for National Statistics. Ethnic group by age and sex, england and wales: census 2021. 2021. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/ethnicity/articles/

ethnicgroupbyageandsexenglandandwales/census2021 [Accessed 19 Oct 2024].

23 Gutacker N, Mason AR, Kendrick T, *et al*. Does the quality and outcomes framework reduce psychiatric admissions in people with serious mental illness? A regression analysis. *BMJ Open* 2015;5:e007342.

24 Harrison MJ, Dusheiko M, Sutton M, *et al*. Effect of a national primary care pay for performance scheme on emergency hospital admissions for ambulatory care sensitive conditions: controlled longitudinal study. *BMJ* 2014;349:g6423.

25 Dusheiko M, Gravelle H, Martin S, *et al*. Quality of Disease Management and Risk of Mortality in English Primary Care Practices. *Health Serv Res* 2015;50:1452–71.

26 Forbes L, Marchand C, Doran T, *et al*. The role of the Quality and Outcomes Framework in the care of long-term conditions: as systematic review. *Br J Gen Pract* 2017;67:e775–84.

27 Ryan AM, Krinsky S, Kontopantelis E, *et al*. Long-term evidence for the effect of pay-for-performance in primary care on mortality in the UK: a population study. *Lancet* 2016;388:268–74.

28 Feng Y, Gravelle H. Patient self-reported health, clinical quality, and patient satisfaction in English primary care: practice-level longitudinal observational study. *Value Health* 2021;24:1660–6.

29 Ciccone MM, Aquilino A, Cortese F, *et al*. Feasibility and effectiveness of a disease and care management model in the primary health care system for patients with heart failure and diabetes (Project Leonardo). *Vasc Health Risk Manag* 2010;6:297–305.

30 Marmot M, Allen J, Goldblatt P, *et al*. Strategic review of health inequalities in england post. Fair Society, Health Lives; 2010. Available: https://www.parliament.uk/globalassets/documents/fair-society-healthy-lives-full-report.pdf

31 Maheswaran H, Kupek E, Petrou S. Self-reported health and socio-economic inequalities in England, 1996-2009: Repeated national cross-sectional study. *Soc Sci Med* 2015;136–137:135–46.