PracEQ1b 27/11/00 17:13 6775

Measuring income related inequality in health at practice level: some preliminary results

Mark Dusheiko[*] Hugh Gravelle[*] Stephen Campbell[**]

Abstract. Estimates of income related inequality in health as measured by concentration indices are constructed from a sample of patients from 60 practices. We find unsurprisingly that the poor have a disproportionately large share of ill health. Allowing for the correlation of income with other factors affecting health reduces the extent of inequity but does not eliminate it. Most of the crude income related inequality arises from the direct effect of income on health rather than from health affecting variables correlated with income. We find that allowing for practice effects makes little difference to the estimate of concentration for the whole sample or to the level of inequity within practice populations. Practice inequality scores were significantly different across HAs and were affected by whether the practice had a diabetes clinic and a PMS contract.

[*] National Primary Care Research and Development Centre, Centre for Health Economics, Alcuin Block D, University of York, York YO10 5DD; mad@york.ac.uk; hg8@york.ac.uk. Funding from the Department of Health to NPCRDC is acknowledged. The views expressed are those of the authors and are not necessarily those of the DH.
[**] National Primary Care Research and Development Centre, University of Manchester.

PRELIMINARY DRAFT: NOT FOR QUOTATION OR CITATION AT THIS STAGE

1. Introduction

Government policy in the NHS is strongly focused on reducing inequality in health. Since general practices are the source of most contacts with the NHS and act as gatekeepers to the rest of the system it is of interest to know if practice policies and organisation have any effect on inequality. In this paper we report some highly preliminary results of an investigation of the extent of inequality in health within practices and differences between practices.

There is horizontal inequity in the distribution of health across the population when persons with different incomes but otherwise similar characteristics have different health. It is possible to test for the existence of inequity by regressions of health on factors including those such as income which it is felt ought not to affect health. However, for policy analysis it is necessary to have a measure of the amount of inequity so that the effects of alternative policies can be compared.

We attempt to measure the extent of horizontal inequity at practice level using the methods developed by Wagstaff and van Doorslaer (1999) determine whether there is statistically significant difference in the health of individuals with different income levels after controlling for other factors which may affect health. We address a number of issues

- Does taking account of practice variables make any difference to the extent of income related inequality in health in the population as a whole?
- Can indices of inequality within practices be sensibly calculated from information from relatively small samples of patients within practices?
- Are there significant differences between practice level inequality scores?
- Are there characteristics of practices which explain the differences between practice level inequality scores?

We start by summarising the Wagstaff and van Doorslaer (1999) methodology (section 2) and its application to the detection of practice level differences in inequity. Section 3 describes the data and section 4 presents the results. Section 5 concludes.

2. Methods

Individual health is related to individual characteristics (age, sex, income, etc.) and possibly to the way in which their practice delivers health care. Some of these characteristics (age, sex) may be felt to be justifiable sources of variation and others may not. The most obvious candidate as an unjustified source of differences in health is differences in income. The extent of unjustified differences in health with respect to income may vary across practices and may be related to aspects of practices that are amenable to policy. We want to be able to construct a summary measure of the inequity in the distribution of health: a measure of the amount of variation which is due to variations in income rather than to variations in justifiable determinants of health such as age and sex. Such a summary measure is essential for assessing the impact of policies directed at reducing inequity.

2.1 Concentration indices

One commonly used inequality measure is the concentration index, which is a generalisation of the Gini coefficient. To measure income related inequality in health we plot the concentration curve $L(s)$ which graphs the cumulative proportion of ill health against the cumulative proportion of the population ranked by income (see Figure 1). If there is no income related inequality in health the poor will be, other things equal, as unhealthy as the rich and the poorest k% of the population will have k% of total population sickness. The concentration curve will then coincide with the 45^o line. If poorer people are less unhealthy than the rich, then poorest k% will have less than k% of the total ill health. Hence the concentration curve will lie below the 45^o line. If ill health is negatively related to income then the concentration curve will lie above the 45^o line. In Figure 1 the poor have a disproportionately large share of illness and $L(s)$ lies above the 45^o line.

The concentration index C summarises the total amount of income related inequality in illness and is analogous to the Gini coefficient. It is defined as twice the area between the illness concentration curve $L(s)$ and the diagonal:

$$C = 1 - 2\int_0^1 L(s)ds$$

When the poor have a disproportionately large share of ill health the concentration curve $L(s)$ lies above the diagonal and C is negative. When ill health is positively associated with income, so the poor have a disproportionately small share of ill health, the curve $L(s)$ lies beneath the diagonal and C is positive. We will refer to the former situation as inequity, so factors increasing the value of C (making it less negative by pulling $L(s)$ down towards the diagonal) are reducing inequity

It is possible for the concentration curve to cross the diagonal, so that there may be one type of inequality amongst the population with lower incomes and the opposite type of inequality within those with higher incomes. In such a case C could have a value close to zero disguising the fact that ill health was strongly associated with income, albeit in a highly non-linear way.

With individual level data the concentration index C can be calculated as

$$C = \frac{2}{nm}\sum_{i=1}^{n} y_i R_i - 1$$

where y_i is ill health for the ith individual (i = 1, …, n) where individuals are ordered from lowest to highest income. $R_i = (2i-1)/2n$ is the relative rank of the ith person and m the mean level of ill health. If person i's ill health increases by Δy and person j's falls by same amount, the change in C is given by

$$\Delta C = \frac{2}{nm}\Delta y\left(R_i - R_j\right)$$

which is negative if i has a lower income than j. The poor now have a greater share of illness and the C has become more negative (or less positive).

The concentration index can be easily obtained by OLS estimation of the regression

$$2s_R^2\left[y_i / m\right] = a_2 + b_2 R_i + u_i$$

where s_R^2 is the variance of the relative ranks. To see this notice that the mean of R_i is ½ and that the estimated coefficient on R_i is

$$\hat{b} = \frac{\operatorname{cov}(R_i, 2s_R^2 y_i / m)}{s_R^2} = \frac{2s_R^2 / m \operatorname{cov}(R_i, y_i)}{s_R^2} = \frac{2}{nm}\sum_{i=1}^{T}\left(y_i - m\right)\left(R_i - \tfrac{1}{2}\right)$$

Some manipulations show that the last expression is equal to C.

The approach assumes implicitly either that income is the only factor which affects health or, slightly more plausibly, that other factors which affect health are not correlated with income. But suppose that women are both poorer and healthier than men. Suppose also that the true effect of income on ill health is negative: rich women are healthier than poor and rich men are healthier than poor, but less healthy than rich women. The average amount of ill health of the rich will be increased relative to the average ill health of the poor by the fact that the proportion of healthier women is smaller in the rich. The simple bivariate association between income and ill health will be contaminated by the systematic variation of the sex ratio with income. Hence cumulating ill health by income without allowing for the effects of other factors which may be associated with health and income will give a misleading impression of the amount of inequality in health, which is due to income. The cumulation will show both inequality due to income affecting health and to income being correlated with other factors affecting health.

2.2 Indirect standardisation

Wagstaff and van Doorslaer (1999) have emphasised the importance of controlling for other factors. They refer to the inequality arising from the association of health with these other factors as “unavoidable” inequality and suggest that it should be removed from the calculation of income related inequality. The point is well taken though the label is perhaps unfortunate since it implies that the inequality which is directly related to income is “avoidable” and can be altered by policy measures to change the distribution of income or the way in which income affects health.

Wagstaff and van Doorslaer (1999) suggest indirect standardisation as a simple and convenient method of removing the confounding effects of health affecting variables which are correlated with income. With indirect standardisation we estimate the amount of ill health each individual would have had allowing only for the effect of non-income factors affecting health. Ill health is regressed on the non-income variables and the predicted level of ill health for each individual is cumulated with respect to income to yield the indirectly standardised concentration curve $L^*(s)$. If the non-income factors are categorical variables such as sex or age range or ethnicity then the regression using dummy variables is equivalent to calculating age/sex/ethnicity specific illness rates for the population as whole.

If income affects the risk of ill health, ceteris paribus, but is correlated with the non-income variables used to predict ill health, then the standardised concentration curve and the unstandardised or raw concentration curve will not coincide. The area between the two curves, therefore, can be used to measure the extent of income related inequality purged of the effects of the correlation of income with other health affecting variables. Thus, suppose that ill health risk increases with income and with age and that income is negatively correlated with age. Then the raw concentration curve $L(s)$ will lie above the standardised concentration curve $L^*(s)$ which in turn will lie above the 45^o line. The difference between the two curves is shows the income related inequality in health after allowing for the fact that age affects health and is correlated with income.

The inequality index is twice the area between the standardised and raw concentration curves:

$$I^* = 2\int_0^1 \left[L^*(s) - L(s)\right]ds = 2\int_0^1 \left[1 - L(s)\right]ds - 2\int_0^1 \left[1 - L^*(s)\right]ds = C - C^*$$

C^* is the concentration index computed by replacing the actual illness score y_i by the predicted score for *i* taking account of their non-income characteristics. Thus the inequality index is the difference between the raw and standardised concentration indices. Note (a) that if income was the only factor affecting ill health or (b) if income was not correlated with the other factors affecting income then $L^*(s)$ would be the 45^o line, C^* would be zero and I^* would equal the raw concentration index.

Estimates of I^* can be obtained from a two step procedure. First regress y_i on a vector of non-income variables x_i hypothesised to affect ill health. Then use the predicted values of ill health $y_i^* = \hat{b}_2 x_i$ from the first stage regression in the second stage OLS regression

$$2s_R^2 \left[\frac{y_i}{m} - \frac{y_i^*}{m*}\right] = a_2 + b_3 R_i + u_i$$

The estimated coefficient b_3 equals I^*.

2.3 Statistical inference

Since inequity indices are normally computed from survey data it is necessary to estimate standard errors to perform significance tests and calculate confidence intervals for the indices. The calculation of concentration indices from OLS regressions also provides an estimate of the standard error. Kakwani *et al.* (1997) showed that it is important to account for serially correlated errors ($E[u_i u_{i-n}] \neq 0$), and derived a formula for estimating the errors allowing for autocorrelation.

Their procedure assumes that the errors are homoscedastic. This may not be a valid assumption given that ill health is a function of other factors whose variance may differ systematically across income groups (different factors affect the probability of ill health amongst the rich and poor in different ways). Further, since our data is from a sample it may also be important to control for the effect of the sampling design. Sampling weights and clustering due to stratified sampling may have to be allowed for in estimating standard errors. Thus instead of using the Kakwani et al approach it may be better to use the Huber/White corrected standard errors and to allow for clustering within sampling units (practces in our case.)

2.4 Decomposing inequality

An alternative to using I^* to attempt to isolate the inequality arising from the direct effect of income is to decompose the overall concentration index C into components attributable to the other variables which affect health. Drawing on results from the decomposition of the

Gini coefficient, Wagstaff (1999) has suggested the following procedure. (We follow the account in van Doorslaer and Koolman, 2000).

Suppose that ill health is linearly related to income (or some suitable transformation of income) and to other variables:

$$y_i = \sum_j b_j x_{ij} + e_i$$

where one of the x_{ij} variables is individual i's income. The equation should be thought of as a reduced form where the explanatory variables are exogenous. (We thus ignore for example any feed back effects from health to income.) The overall concentration index for ill health with respect to income C can then be expressed as the weighted average of the concentration indices of the determinants of health

$$C = \sum_j (b_j \bar{x}_j / \bar{y}) C_j + (\bar{e} / \bar{y}) C_e$$

where C_j is the concentration index of variable j with respect to income and the weights $b_j \bar{x}_j / \bar{y}$ are the average shares of variable j in y. Note that the concentration index of income with respect to income is just the Gini coefficient.

To calculate the decomposition we estimate the reduced form model of ill health to get the estimates of the $\hat{b}_j$ coefficients and $\bar{y} = \sum_n \hat{b}_i \bar{x}_i$. We then regress health and each of the explanatory variables for ill health (including income) on relative rank R_i to get the overall ill health concentration index C and the concentration indices C_j for the explanatory variables.

The objective of the decomposition is to shed light on the two ways in which a factor can contribute to explaining income related inequalities in health:

1. through its average effect on health $\hat{b}_j \bar{x}_j$
2. its association with income (C_j)

For instance, if increasing age increases the likelihood of ill health, the average age in the sample is high, and older individuals have much lower incomes than younger individuals, then we should find that age would be a major contributing factor to the observed income related inequality in health.

3. Data

The General Practice Assessment Survey (GPAS) was sent to a sample of patients from 60 practices. GPAS asks patients about use of their general practice, their views on its accessibility and quality of care. We used an augmented version of the basic instrument, which had additional socio-economic questions, including income and employment status and various aspects of health. The health variable we use is limiting long standing illness.

The survey design was complex, with practices selected non randomly due to stratification by Regional (RHA) and local health authority (HA) areas. Initially, 3 RHAs were selected from 8 possible areas, and then 2 HAs were chosen from each of the 3 RHAs so that the proportion of practices receiving Jarman (deprivation) payments and rural practice payments were similar to the proportions of practices receiving these payments nationally. The sampling of HAs was purposeful, hence the use of sampling weights to adjust for differential probabilities of HA selection may not be possible. Instead we controlled for sampling effects by estimating models with fixed effects for each HA.

From each of the 6 selected HAs, a stratified random sample of 10 practices was taken. The stratification ensured that the sampled practices were nationally representative according to their number of partners, proportion receiving deprivation payments, and proportion having training status. If a sampled practice refused to take part, then the next practice of that 'type' or near enough was selected.

Within each practice, approximately 200 adult patients were randomly selected from the practice list. Practices varied in population list size, so that the probability of an individual being selected depended on the size of the practice. Hence since the sample is not self-weighting, practice level sampling weights are required.

Questionnaires were sent to 12104 patients in 60 practices. 4462 completed questionnaires were returned, an overall response rate of 37%. There was some item non-response, especially for the income question with valid responses obtained for only 3090 individuals. In this paper we limited our sample to patients who responded to the income question.

4. Results

We examine the overall level of income related inequality in ill health across the whole sample (section 4.1). The level of inequality within practices is estimated in section 4.2 and practice factors associated with practice inequality scores in section 4.3. Section 4.4 considers the decomposition of overall inequality. Table 1 describes the variables used in the analysis. The measure of ill health is binary: whether an individual reports limiting long standing illness.

4.1 Extent of income related inequality

Concentration indices were estimated using the convenient regression approaches outlined in the methods section. Both crude and indirectly standardised estimates incorporated sampling weights, which are the inverse of the probability that an individual is selected from their practice. As suggested by Cowell and Jenkins (2000) the individual data were weighted before performing any transformation required for inequality estimation. For the indirect standardisation probability weights were included in the standardising regression and HA dummy variables were used to control for the effect of stratification .

Income was reported in ranges and for the indirect standardisation procedure we generated a continuous income variable by interval regression. Interval regression utilises the information on the location of an individual's income within a range and on other variables to make a point estimate of income. It can cope with the fact that income was censored at 40,000 and bounded below by zero. Point estimates were made only for those individuals who were willing to report the band in which they were located. This reduced our sample by nearly 25%, and may also introduce selection bias if these individuals differ systematically from those who did not report income. We will investigate the problem in future work.

4.1.1 Crude income related illness inequality

Estimates of inequality were obtained using the regression approach outlined in Section 2. Standard errors were estimated by the Huber/White method to control for heteroscedasticity and errors were clustered by practice to allow for the possibility that they may not be independent within practices. The use of weights may also affect the sampling distribution of the errors, though allowing for clustering within practices may overcome this.

The estimated raw unstandardised concentration ratio for illness (C) is equal to the estimated coefficient on income rank: -0.2017. The t-ratio was significant at the 1% level, implying that we can reject the null hypothesis of no significant income related inequality in ill health. The negative value of (C) means that the poor have a disproportionately large share of illness.

4.1.2 Indirectly standardised illness inequality

A better reflection of the extent of income related inequity in ill health is derived by standardising the index for non-income factors that can be considered exogenous determinants of ill health inequity. In this analysis, reported long-standing limiting illness was standardised for the age, gender, ethnicity, and marital status of the individual. Health Authority dummy variables were included to account for any affect of stratification in sample selection.

The selection of non-income variables to include in the first stage standardisation regression is problematic. The construction of a measure of income related inequality implies that a value judgement has been made that there are certain characteristics of individuals which ought not to be related to their health and whose relationship with health are amenable to policy manipulation. Income is one such characteristic. But there may be other variables with these properties, for example education. Should we include such variables in the standardisation regression when constructing a measure of income related inequality? If the variable is not correlated with income they will have no effect on the indirectly standardised concentration index. But such variables are likely to be correlated with income (education being a case in point). Omitting them from the standardisation regression will then bias the estimated amount of income related inequality because the index will in part reflect the correlation of income and the omitted variable. This suggests that we should include in the standardisation equation all non-income variables which are hypothesised to affect health, whether or not we believe that they ought to do so. In our particular example one may feel that ethnicity ought not to affect health. The justification for including it in the standardisation equation is that by doing so we can purge the income related inequality measure of any indirect effects of ethnicity arising from its correlation with income.

A two step procedure was employed. A probit regression was used to estimate the probability of long standing limiting ill health, conditional on non-income characteristics deemed to affect their health state. The relative difference between actual and expected health status was then calculated and used to estimate the indirectly standardised concentration index I^*. I^* is equal to the coefficient on income rank in Table 3 and is estimated to be -0.0977 , and significant at the 1% level. The result suggests that after controlling for non-income factors thought to affect health, there was significant income related inequality with the poor having a disproportionately high share of illness.

4.2 Practices and inequality

4.2.1 Practice effects and overall inequality

In this section we report some initial attempts to test whether allowing for differences between practices has implications for the measurement of inequality across the whole sample. This could arise if there are characteristics of practices which affect health and are correlated with income. Practice effects can arise from environmental factors, unobserved socio economic individual characteristics whose distribution varies across practice populations, and attributes of practice organisation and resources.

Initially we have considered the effects of including practice dummies in the standardisation equation in addition to age, sex, ethnicity, and marital status. The dependent variable in Table 4 is the predicted probability of ill health and the coefficient on the income rank variable gives the estimated standardised concentration index $I^* = -0.08504$.

Comparison with Table 3 shows that controlling for practice effects reduces the amount of income related inequality by about 13% and also reduces the significance of the estimate somewhat. Practice characteristics appear to have some effect on ill health and to be correlated with income. On average these characteristics are either negatively related to income and positively related to health or positively associated with income and negatively related to health. However either the relationship with income or with health is weak since the reduction in I^* is not significant.

4.2.2 Practice level inequality

As a further initial test for the effects of practice characteristics on income related inequality in illness we estimated standardised inequality indices I^* for the 60 practice populations. We first estimate a standardisation equation for illness using the whole sample of patients from all practices. We then take the sample of patients from a practice and calculate their relative income rank within the practice and use this as the income rank variable in the second stage of the estimation of I^* for the practice.

One problem with estimating inequality at practice level is the small number of patients from some practices. One practice had only 8 patients and was dropped from the analysis. For the remaining 59 practices the sample size ranged from 12 - 70 observations, with a median 47.

Two different approaches to indirect standardisation were taken and the resulting practice inequality indices are in Table 5. The first column of Table 5 is derived from a standardising regression without practice level dummies. It is based on the same standardisation equation as the estimates of I^* for the whole sample in Table 3 in section 4.1.2. The implicit assumption is that individuals in different practices but with the same age, sex, ethnicity and marital status have the same predicted probability of illness. Differences in inequality within practices are therefore assumed to be due solely to differences in the distributions of income or to differences in the impact of income on health.

Column (2) of Table 5 is derived from a standardisation regression including practice dummies. The assumption here is that individuals with the same age, sex, ethnicity and marital status but in different practices may have different predicted illness probabilities. All individuals in the same practice have their estimated illness probability increased or decreased by the same amount. Even if the inclusion of practice dummies had no effect on the estimated value of the other coefficients in the standardisation equation, the practice inequality scores will be changed since the concentration index is invariant to proportionate but not additive changes in the distribution of illness.

Figures 2 and 3 show the frequency distributions of practice inequality scores in columns 1 and 2 in Table 5. In both cases the bulk of the practices have negative inequality scores indicating that the poor have disproportionately high levels of illness after standardising for age, gender, ethnicity, and marital status, irrespective of whether practice effects are allowed for. The concentration index is distributed fairly normally around the mean, with one extreme outlier in the left tail. The correlation between the scores is 0.9963.

When practice effects are not included in the standardisation regression (Figure 2) the mean level of practice inequality was -0.17, with a range from -0.92 to 0.36, and an inter-quartile range of -0.26 to -0.8. Only 6 practices had statistically significant inequality, ranging from -0.47 to -0.21 with a mean of -0.33.

Allowing for fixed practice effects in the standardisation (Figure 3) had little effect on the mean inequality score and reduced the range of estimated inequality scores slightly (-0.90 to 0.30). The inter-quartile range shifted towards zero (-0.25 - 0.7). Seven practices have inequality scores significantly different from zero. The same six practices plus another were found to have significant negative scores ranging between -0.46 and -0.23 with a mean of -0.33.

4.3 Determinants of practice inequality

We have a rich set of data on practices and will use it in future to investigate the relationship between practice level inequality and practice characteristics. Here in Table 6 we report the results of some initial regressions of practice inequality against a limited set of practice characteristics. (Similar results were obtained using the inequality indices obtained without practice dummies in the standardisation equation.) In addition to significant HA effects, inequality was affected significantly by whether the practice had a PMS contract or the standard GMS contract. Remember that a negative inequality index indicates that the poor have a disproportionately high share of illness. Not being a PMS practice reduces the poor's share of illness whereas not having a diabetes clinic increases it. Other variables were not significant though we note that in larger practices (more GP) and where GPs have bigger lists, the poor's share of illness is higher.

4.4 Decomposition of inequality

To assess how the various characteristics of income related inequality in health contribute to observed inequality we apply Wagstaff's inequality decomposition to self-reported long standing illness. A linear probability model with the binary dependent variable being whether the individual reported limiting long standing illness was estimated. The explanatory variables included the equivalised income. Table 4 shows (a) the income related concentration indices for each variable C_i ; (b) the estimated coefficients and t statistics from the linear probability model; (c) the average proportionate contribution of each variable to ill health ($\hat{b}_i \bar{x}_i / \bar{y}$); and (d) its contribution to the overall concentration of ill health $(\hat{b}_i \bar{x}_i / \bar{y})C_i$ and its percentage share of overall concentration ($(\hat{b}_i \bar{x}_i / \bar{y})(C_i / C)$). See Table 1 for variable definitions. Practice dummy variables were included in the regression but are not shown in the table to save space.

The proportionate contribution indicates the percentage increase or decrease in inequality that would occur if, holding all else constant, that factor was equally distributed across the income range or had no effect on the health. It should be noted that the contribution is relative to the reference category where the variable is categorical. The constant term captures the reference individual (when all variables are set equal to zero). It has a concentration index of zero and so serves as an appropriate baseline to compare the effects of explanatory variables in the decompostion.

The crude concentration index of limiting long standing illness was estimated to be -0.202, indicating that the poor have a disproportionately large share of illness. Tables 7 and 8 report the details of variables, in order of importance, which reduce or increase overall inequality by more than 1%.

Whether a variable contributes to the overall level of income related inequality depends on its effect on health and its concentration with respect to income. Variables which are more closely correlated with income will have higher absolute concentration ratios.

The results show that income related inequity in ill health is driven predominantly by its association with income, employment status and socio-economic conditions. Holding other things equal, differences in income levels explain nearly a third of the observed crude inequality in ill health. This is predominantly due to its strong negative relationship with ill health, which increases expected illness as income decreases, and the degree of income inequality. Not surprisingly, individuals who are economically inactive through retirement and long term illness contribute around 45% of inequality.

Looking after the household is also associated with poorer health and more income inequality, as van Doorslaer and Koolman (2000) found. Proxies for socio-economic deprivation, specifically local authority or residential housing tenure and car ownership also worsen inequity by a total of 20%. The effect of car ownership appears to influence inequality because of its prevalence in the population, whereas housing tenure acts through its relationship with income. Non smokers are less likely to report illness and are more likely to be wealthy. Members of the Bangladeshi community (Iethnic_7) tend to contribute positively to the observed income related inequality in poor health, mainly due to the fact that they have low incomes.

Some of the practice dummies contribute more than 1% of total inequality. Controlling for other factors, and relative to practice 1, individuals from these practices are expected to report less long term illness and to have higher incomes.

Relatively few variables improve the distribution of illness. Unemployed individuals still seeking work have less limiting long standing illness and lower income. Surprisingly, holding all else constant, we find that age reduces inequity in ill health, because older individuals are wealthier but also have more ill health. Members of the Pakistani community have less illness and lower incomes. Patients in practice 33 have significantly less illness relative to practice 1 and a lower income. The overall net practice effect is to reduce inequity (increase I^*) very slightly (0.28%).

5. Conclusions

Since we are reporting on work in progress the following conclusions are necessarily highly tentative:

- unsurprisingly the poor have a disproportionately large share of ill health
- allowing for the correlation of income with other factors affecting health does reduces the extent of inequity but does not eliminate it
- most of the crude concentration in inequality arises from the direct effect of income on health
- allowing for practice effects makes little difference to the estimate of concentration for the whole sample or to the level of inequity within practice populations
- practice inequality scores were significantly different across HAs and were affected by whether the practice had a diabetes clinic and a PMS contract.

The work reported is still in its early stages and we plan to investigate

- alternative standardisation and decomposition equations. Some of the variables in the decomposition equation may be considered to be endogenous and others to be jointly determined with income. We need to consider the underlying model of the determination of health and its implications for the definition and measurement of inequality.
- additional practice characteristics to investigate the determinants of practice inequality scores
- inequality in patient satisfaction and in use of practices.

References

Cowell F.A. and S. P. Jenkins (2000). Estimating Welfare Indices: Household Weights and Sample Design. Colchester: University of Essex, Institute for Social and Economic Research (ISER), ISER Working Papers: No. 2000-23

van Doorslaer, E. and X. Koolman (2000). Income-Related Inequalities in Health in Europe: Evidence from the European Community Household Panel. Erasmus University: 1-34.

Kakwani, N., A. Wagstaff, et al. (1997). "Socioeconomic inequalities in health: Measurement, computation, and statistical inference." Journal of Econometrics 77: 87 - 103.

Wagstaff, A. (1999). Inequalities in Child Mortality in the Developing World: How large are they? How can they be reduced? The World Bank: 34.

Wagstaff, A. and E. v. Doorslaer (1999). "Measuring and Testing for Inequality in the Delivery of Health Care." mimeo.

Table 1. Variables

Variable	Mean	Description	Variable	Mean	Description
Iillness_1	0.248575	Limiting long standing illness	Ichild_5	0.012923	4 children in the household
ISEX_2	0.578867	Female	Ichild_6	0.003421	5 children or more in the household
age*	**48.85899**	Age	Iaccom_1 (ref)		Owner occupied/mortgaged
Istatu_1	0.124287	Single	**Iaccom_2***	**0.131889**	Rented local authority/ housing association
Istatu_2 (ref)		Married/cohabiting	Iaccom_3	0.055492	Rented from private landlord
Istatu_3	0.085519	Separated	Iaccom_4	0.017104	Other arrangement
Istatu_4	0.067655	Widowed	**eqlinc**	**9.739116**	Log of equivalised income
Iethni_1 (ref)		White	Ioccu_1 (ref)		Working full time
Iethni_2	0.010642	Black - Caribbean	Ioccu_2	0.13835	Working part time
Iethni_3	0.007982	Black - African	Ioccu_3	0.018624	Self Employed (employing others)
Iethni_4	0.005321	Black - Other	Ioccu_4*	0.04371	Self Employed (not employing others)
Iethni_5	0.007982	Indian	Ioccu_5	0.00076	Starting a job
Iethni_6	**0.009882**	Pakistani	Ioccu_6	0.002661	Government or employment training
Iethni_7	**0.003801**	Bangladeshi	**Ioccu_7**	**0.020144**	Unemployed/ looking for a job
Iethni_8	0.002281	Chinese	Ioccu_8	0.011022	At school/full time education
Iethni_9	0.015203	Other	**Ioccu_9***	**0.053592**	Unable to work because of sickness
Ichild_1 (ref)		No children in the household	**Ioccu_10***	**0.223489**	Retired from paid work
Ichild_2	0.139491	1 children in the household	**Ioccu_11**	**0.08856**	Looking after home and family
Ichild_3*	0.145572	2 children in the household	**Iownca_1***	**0.808818**	Access to car
Ichild_4*	0.053972	3 children in the household	**Ismoke_2***	**0.492969**	Never smoked for > 1 year

Table 2. Estimation of crude unstandardised concentration index

```
Regression with robust standard errors                 Number of obs =    2631
                                                       F(  1,    59) =   48.78
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.0253
Number of clusters (prac_id) = 60                      Root MSE      =  .36359

------------------------------------------------------------------------------
         |               Robust
depwxih0 |      Coef.   Std. Err.       t     P>|t|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
  incrank|  -.2017259   .0288816     -6.985   0.000      -.2595178    -.143934
   _cons |   .2641603   .0259808     10.168   0.000       .2121727    .3161478
------------------------------------------------------------------------------
```

Table 3. Estimation of indirectly standardised concentration index

(Standardised for age, gender, ethnicity, marital status. Sample weights used, and health authority effects allowed for but not practice level effects.)

```
Regression with robust standard errors                 Number of obs =    2631
                                                       F(  1,    59) =   11.38
                                                       Prob > F      =  0.0013
                                                       R-squared     =  0.0065
Number of clusters (prac_id) = 60                      Root MSE      =   .3512

------------------------------------------------------------------------------
         |               Robust
 wdepIh1 |      Coef.   Std. Err.       t     P>|t|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
  incrank|  -.0977062   .0289584     -3.374   0.001      -.1556518   -.0397606
   _cons |   .0462518   .0253269      1.826   0.073      -.0044273    .0969309
------------------------------------------------------------------------------
```

Table 4. Estimation of standardised concentration index controlling for practice effects

(Standardised for age, gender, ethnicity, marital status and practice fixed effects. Sample weights used.)

```
Regression with robust standard errors                 Number of obs =    2631
                                                       F(  1,    59) =    8.50
                                                       Prob > F      =  0.0050
                                                       R-squared     =  0.0049
Number of clusters (prac_id) = 60                      Root MSE      =  .35032

------------------------------------------------------------------------------
         |               Robust
 wdepIh2 |      Coef.   Std. Err.       t     P>|t|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
 incrank |  -.0850364   .0291703     -2.915   0.005      -.1434061   -.0266667
   _cons |   .0402542   .0251058      1.603   0.114      -.0099825    .0904909
------------------------------------------------------------------------------
```

Table 5: Comparison of relative inequality indices

Prac. id	(1)	(2)	Prac. id	(1)	(2)
1	-0.16632	-0.17858	31	-0.27477	-0.27544
2	-0.31224	-0.31042	32	-0.08712	-0.08737
3	-0.18953	-0.18721	33	-0.24332	-0.24877
4	-0.13338	-0.12784	34	-0.01534	-0.01305
5	-0.07497	-0.06398	35	-0.26293	-0.28589
6	-0.14886	-0.15045	36	0.03781	0.054613
7	.	.	37	-0.10474	-0.1122
8	-0.04691	-0.04681	38	-0.43637	-0.43939
9	-0.21758	-0.2036	39	-0.11621	-0.11214
10	-0.06589	-0.06817	40	-0.0989	-0.08066
11	-0.15635	-0.14546	41	-0.29206	-0.29267
12	-0.91545	-0.89799	42	-0.45428	-0.4493
13	-0.29265	-0.27467	43	-0.21287	-0.23156
14	-0.00033	0.053585	44	-0.17015	-0.16842
15	-0.15614	-0.16311	45	-0.11206	-0.12687
16	-0.17306	-0.16914	46	-0.2597	-0.26277
17	0.090208	0.095399	47	-0.3142	-0.30986
18	-0.13127	-0.13048	48	-0.03043	-0.03307
19	-0.00935	-0.02483	49	-0.16089	-0.15608
20	-0.06795	-0.05883	50	-0.30672	-0.30406
21	-0.47016	-0.45475	51	-0.26437	-0.24849
22	-0.1288	-0.12893	52	-0.18093	-0.18707
23	-0.26614	-0.24829	53	-0.04763	-0.05199
24	-0.00714	-0.0106	54	-0.1619	-0.16075
25	-0.16341	-0.17052	55	-0.1821	-0.1749
26	-0.04815	-0.04849	56	-0.28166	-0.23699
27	-0.15922	-0.15705	57	0.364522	0.305701
28	0.063476	0.052021	58	-0.21302	-0.21223
29	-0.29936	-0.29977	59	-0.22953	-0.22626
30	-0.12862	-0.13294	60	-0.12716	-0.12107

Column (1) derived from standardisation regression with no practice dummies. Column (2) standardisation regression uses practice dummies.

Table 6. Simple model of practice inequality

(Dependent variable is practice inequality standardised with practice effects)

```
Regression with robust standard errors                Number of obs =      57
                                                      F(  4,     5) =   25.68
                                                      Prob > F      =  0.0016
                                                      AR2           =  0.035
                                                      R-squared     =  0.2249
Number of clusters (HA_ID) = 6                        Root MSE      =  .16453
------------------------------------------------------------------------------
         |               Robust
     in2 |      Coef.   Std. Err.       t     P>|t|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
IHA_ID_2 |  -.0145459   .0280101     -0.519   0.626      -.0865481    .0574564
IHA_ID_3 |  -.0682305   .0227466     -3.000   0.030      -.1267025   -.0097586
IHA_ID_4 |  -.0807367   .0515058     -1.568   0.178      -.2131366    .0516632
IHA_ID_5 |  -.1022812   .0144399     -7.083   0.001      -.1394001   -.0651622
IHA_ID_6 |    -.03398   .0129612     -2.622   0.047      -.0672979   -.0006622
NotPMS   |   .0698602   .0179133      3.900   0.011       .0238125    .1159078
NotTRAIN |  -.0539822    .037288     -1.448   0.207       -.149834    .0418696
Deprvpay |  -.0493142   .0707885     -0.697   0.517      -.2312818    .1326533
   WTEGP |  -.0242382   .0144723     -1.675   0.155      -.0614404     .012964
LISTWTEG |  -.0000393   .0000328     -1.198   0.284      -.0001235     .000045
NoDiabcl |  -.1464453    .033735     -4.341   0.007      -.2331638   -.0597268
```

```
   _cons |   .1264775   .1269857      0.996   0.365      -.1999498    .4529047
------------------------------------------------------------------------------
```

Table 7. Decomposition of income related inequality in self reported long standing illness

Variable	Concentration index	Regression coefficient	t-stat	(*B***Xi*_bar/ *Y*_bar)	(*B***Xi*_bar/ *Y*_bar)**C*	Share (%)	Share by variable group
Illness	-0.2017						
ISEX_2	0.017837	-0.02914	-1.791	-0.06785	-0.00121	0.478962	**0.478962**
age*	**0.004167**	**0.004852**	**4.356**	**0.953751**	**0.003974**	**-1.57297**	**-1.57297**
Istatu_1	0.139289	0.023904	0.735	0.011952	0.001665	-0.65887	**-0.22766**
Istatu_3	-0.0855	0.014908	0.323	0.005129	-0.00044	0.173553	
Istatu_4	-0.32612	0.007335	0.128	0.001996	-0.00065	0.257657	
Iethni_2	0.096541	0.115014	0.934	0.004924	0.000475	-0.18814	**-0.33448**
Iethni_3	-0.33571	-0.12887	-1.827	-0.00414	0.001389	-0.5498	
Iethni_4	-0.02213	-0.05839	-0.787	-0.00125	2.77E-05	-0.01095	
Iethni_5	0.037404	0.030959	0.309	0.000994	3.72E-05	-0.01472	
Iethni_6	**-0.40718**	**-0.2234**	**-1.386**	**-0.00888**	**0.003616**	**-1.43125**	
Iethni_7	**-0.50962**	**0.404691**	**1.853**	**0.006188**	**-0.00315**	**1.248079**	
Iethni_8	0.340159	-0.1538	-1.505	-0.00141	-0.00048	0.189966	
Iethni_9	-0.34957	0.04991	0.827	0.003053	-0.00107	0.422331	
Ichild_2	0.094899	-0.04414	-1.547	-0.02477	-0.00235	0.930297	**0.14393**
Ichild_3*	-0.03155	-0.06517	-2.383	-0.03816	0.001204	-0.4765	
Ichild_4*	-0.03601	-0.08501	-3.494	-0.01846	0.000665	-0.2631	
Ichild_5	-0.37593	-0.08222	-1.625	-0.00427	0.001607	-0.63594	
Ichild_6	-0.77046	0.140399	0.904	0.001932	-0.00149	0.589166	
Iaccom_2*	**-0.49954**	**0.103943**	**3.147**	**0.05515**	**-0.02755**	**10.90361**	**11.03049**
Iaccom_3	-0.0205	0.029643	0.704	0.006617	-0.00014	0.05369	
Iaccom_4	-0.06204	0.043316	0.651	0.00298	-0.00018	0.073187	
eqlinc	**0.098038**	**-0.0206**	**-1.813**	**-0.80691**	**-0.07911**	**31.30911**	**31.30911**
Ioccu_2	-0.00314	0.033147	1.185	0.018449	-5.8E-05	0.022893	**45.51256**
Ioccu_3	0.456025	0.024526	0.337	0.001838	0.000838	-0.33165	
Ioccu_4*	0.155316	0.086155	2.799	0.01515	0.002353	-0.93126	
Ioccu_5	-0.63913	0.386319	1.62	0.001181	-0.00076	0.298852	
Ioccu_6	0.038274	0.023288	0.155	0.000249	9.54E-06	-0.00378	
Ioccu_7	**-0.47889**	**-0.10263**	**-1.576**	**-0.00832**	**0.003983**	**-1.57641**	
Ioccu_8	-0.25665	0.070026	0.985	0.003105	-0.0008	0.315404	
Ioccu_9*	**-0.50001**	**0.6884**	**20.705**	**0.148417**	**-0.07421**	**29.37058**	
Ioccu_10*	**-0.28434**	**0.167247**	**4.583**	**0.150369**	**-0.04276**	**16.92187**	
Ioccu_11	**-0.15688**	**0.064467**	**1.881**	**0.022968**	**-0.0036**	**1.426064**	
Iownca_1*	**0.137711**	**-0.05461**	**-2.08**	**-0.17769**	**-0.02447**	**9.68463**	**9.68463**
Ismoke_2*	**0.139438**	**-0.03896**	**-2.451**	**-0.07726**	**-0.01077**	**4.263461**	**4.263461**
Iprac_1-60*							**-0.28804**
				C_hat	-0.25267		

Table 8. Variables worsening the distribution of illness by more than 1%

Variable	***B***	***C***	**(*B***Xi*_bar)**	**(*B***Xi*_bar /*Y*_bar)**	**(*B***Xi*_bar/ *Y*_bar)**C***	**Share (%)**
eqlinc	**-0.0206**	**0.098038**	**-0.20058**	**-0.80691**	**-0.07911**	**31.30911**
Ioccu_9*	**0.6884**	**-0.50001**	**0.036893**	**0.148417**	**-0.07421**	**29.37058**
Ioccu_10*	**0.167247**	**-0.28434**	**0.037378**	**0.150369**	**-0.04276**	**16.92187**
Iaccom_2*	**0.103943**	**-0.49954**	**0.013709**	**0.05515**	**-0.02755**	**10.90361**
Iownca_1*	**-0.05461**	**0.137711**	**-0.04417**	**-0.17769**	**-0.02447**	**9.68463**
Ismoke_2*	**-0.03896**	**0.139438**	**-0.0192**	**-0.07726**	**-0.01077**	**4.263461**
Iprac_12*	**-0.22302**	**0.425465**	**-0.00305**	**-0.01228**	**-0.00522**	**2.067216**
Iprac_49*	**-0.10326**	**0.373361**	**-0.00255**	**-0.01026**	**-0.00383**	**1.516514**

lprac_42*	-0.21002	0.220518	-0.00423	-0.01702	-0.00375	1.485396
loccu_11	0.064467	-0.15688	0.005709	0.022968	-0.0036	1.426064
lprac_26*	-0.10288	0.319613	-0.00274	-0.01101	-0.00352	1.392932
lprac_47*	-0.16313	0.20543	-0.00415	-0.01671	-0.00343	1.358771
lethni_7	0.404691	-0.50962	0.001538	0.006188	-0.00315	1.248079
lprac_32*	-0.16388	0.232283	-0.0033	-0.01328	-0.00308	1.220905

Table 9: Factors improving the distribution of illness by more than 1%

Variable	*B*	*C*	(*B***Xi*_bar)	(*B***Xi*_bar/*Y*_bar)	(*B***Xi*_bar/*Y*_bar)**C*	Share (%)
loccu_7	-0.10263	-0.47889	-0.00207	-0.00832	0.003983	-1.57641
age*	0.004852	0.004167	0.237078	0.953751	0.003974	-1.57297
lethni_6	-0.2234	-0.40718	-0.00221	-0.00888	0.003616	-1.43125
lprac_33*	-0.20673	-0.23292	-0.00369	-0.01486	0.00346	-1.36953

Figure 1. Concentration curves and inequality indices

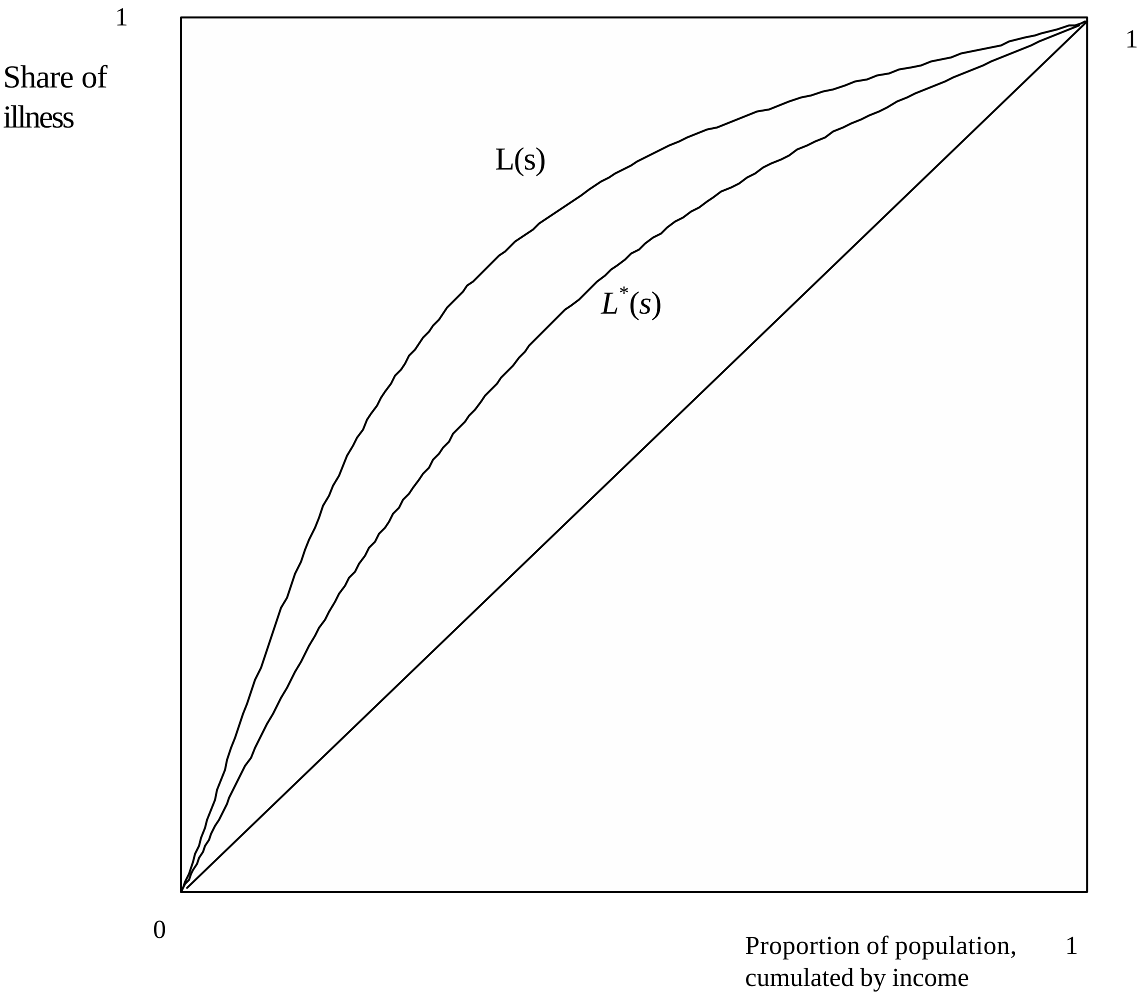

Figure 2. Distribution of practice inequity without allowing for practice fixed effects

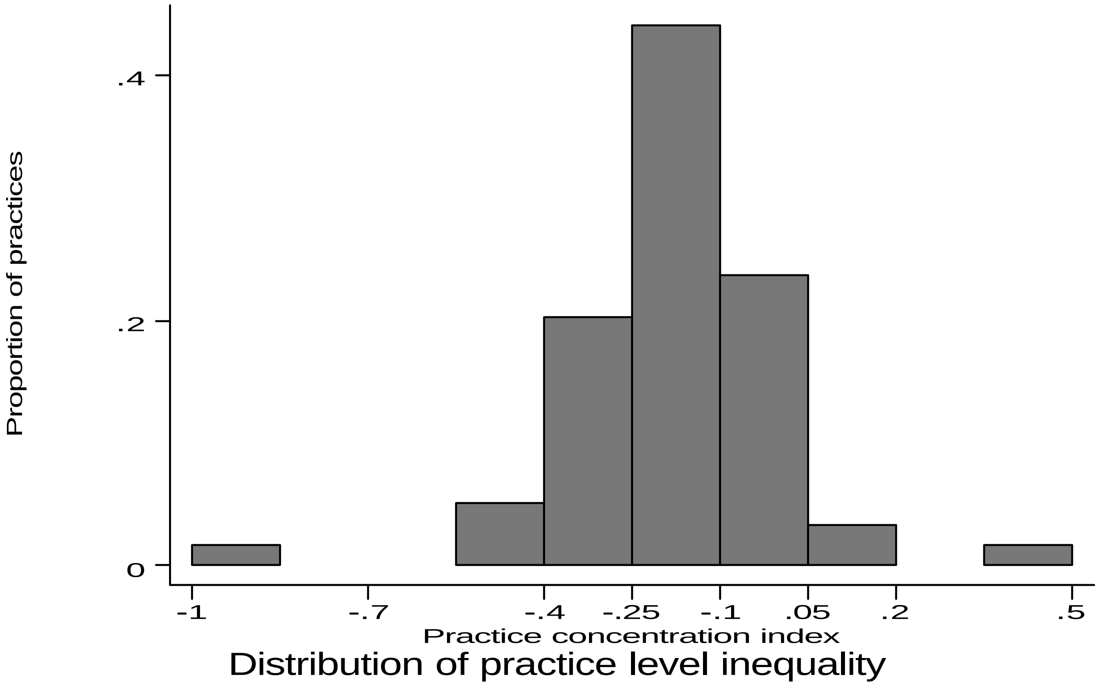

Figure 3: Distribution of relative practice inequity, allowing for practice fixed effects