Putting general practitioners where they are needed: an overview of strategies to correct maldistribution
Putting general practitioners where they are needed: an overview of strategies to correct maldistribution

Professor Bonnie Sibbald
National Primary Care Research and Development Centre
University of Manchester
October 2005

Summary
Inequalities in general practitioner (GP) distribution in the UK have persisted for many decades. Domestic strategies for improving equity have placed greatest reliance on restricting GP entry to over served areas and financially compensating those who locate to deprived areas. Greater use might be made of the wider range of strategies that have been deployed in other countries. These include: educational and other programmes dedicated to the selection, training and professional support of doctors willing to work in underserved areas; additional non-pecuniary benefits for doctors working in underserved areas such as the provision of excellent clinical facilities, higher professional training and a wider range of employment benefits; scholarships or educational loan repayment in return for service in underserved areas; or restricting immigrant doctors to service in designated areas as a condition of licensure. The comparative cost-effectiveness of different incentives is largely unknown. However it is clear that no one strategy is fully effective and that a blend of approaches, incorporating both pecuniary and non-pecuniary incentives, is needed.

Inequalities in distribution
Inequality in the geographical distribution of general practitioners (GPs) has been a persistent policy concern in the UK. Gravelle and Sutton (2001) found that inequality in GP distribution across Family Health Service Authorities (FHSAs) in England and Wales decreased between the mid 1970s and mid 1980s but was at best constant to the mid 1990s. Hann and Gravelle (2004) found that inequalities have persisted to 2003. Despite increases in overall workforce size, areas that were relatively under-provided in 1974 were still under-provided in 2003.

Most unrestricted GP principals are owners of general practices that in turn employ nurses and other GPs. Inequality in the distribution of principals therefore exacerbates inequalities in the distribution of practice employed staff. Hann and Gravelle (2004) found that areas with above-average provision of principals also attracted a disproportionate share of non-principal GPs. Hirst et al. (1998) found that the number of GPs was the most significant predictor of inequalities in the distribution of practice employed nurses across FHSAs during 1988-95.

The geographical areas that experience the greatest problems with GP recruitment and retention tend to be urban and deprived. Gosden and colleagues (2000) showed that the most important influence on GPs’ choice of practice was aversion to location in an area
of high deprivation. Levels of GP turnover and migration are positively correlated with population deprivation, with the most ‘needy’ health authorities sustaining the greatest net loss in GPs (Taylor & Leese 1997). Urban deprived areas tend to have the most unfilled vacancies after one year and the most longstanding vacancies (Department of Health 2004). The situation has been made worse in some areas by the high proportion of GPs nearing retirement age (Mathie & McKinlay 1999; Taylor & Esmail 1999).

Deprived areas pose a number of disadvantages for GPs (Beardow et al. 1993; Gosden et al. 2000; Jarman 1984; Jebb 1991). The workload generated by economically and socially disadvantaged patients is higher than that of comparable populations in more affluent areas. Practise premises and equipment tend to be poor. Partnership sizes are small - often single-handed - offering little opportunity for workload sharing or mutual support. Community amenities such as housing and schools are often poor. There may be few satisfactory employment opportunities for spouses. High crime rates and the threat of violence make the obligation to provide home visits distasteful to doctors, particularly women. Women GPs are less likely than men to locate in urban deprived areas (Beardow et al. 1993) but form an increasingly high proportion of the workforce (Taylor & Leese 1997).

These disadvantages occur against the background of more general disincentives to work as a GP. GP recruitment and retention has been curtailed by the lack of work time flexibility, labour mobility and wage flexibility associated with traditional models of partnership working (Young & Leese 1999; Young et al. 2002). What GPs prefer are practices with larger clinical teams, opportunities to develop outside interests, shorter working hours and smaller list sizes (Gosden et al. 2000; Wordsworth et al. 2004). Higher income is an important determinant of practice choice (Gosden et al. 2000; Wordsworth et al. 2004) but tends to be less important than other factors in predicting job satisfaction (Whalley et al. in press), motivation (Pearson et al. 2004) or intentions to quit (Sibbald et al. 2003). This suggests that policies aimed at influencing GPs’ location choices should take account of both non-pecuniary and pecuniary factors (Scott 2001).

**Strategies for redistribution**

Policies for improving staffing levels in underserved areas need to alter the balance between un-modifiable disincentives to working in deprived areas (e.g. needy patients) and those which are susceptible to intervention (e.g. working hours; remuneration). Strategies may be characterised as normative, utilitarian or coercive (Crandall et al. 1990).

- Normative or missionary strategies aim to encourage a sense of responsibility to serve in places where needs are greatest - for example through tailored education and training schemes - and rely heavily on symbolic rewards based on personal values or prestige.

- Utilitarian or mercantile strategies aim to compensate doctors financially for the additional costs of serving deprived populations, and offset other disadvantages through the provision of excellent facilities, good job structure and employment benefits.
Coercive or military strategies aim to deploy doctors for defined periods in designated underserved areas as a condition of training, financial support or licensure.

Figure 1 gives examples of normative, utilitarian and coercive strategies as applied to individual clinicians, provider organisations and state-level systems concerned with medical training and licensing.

**Figure 1: Redistribution strategies**

<table>
<thead>
<tr>
<th>Target level</th>
<th>Strategic orientation</th>
<th>Normative</th>
<th>Utilitarian</th>
<th>Coercive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual</td>
<td>Professional support and development schemes.</td>
<td>One-off payments to attract or retain doctors.</td>
<td>Increased income</td>
<td>Educational loan repayment or scholarship in return for obligated service.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Increased job flexibility and/or benefits</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Provider organisation</td>
<td></td>
<td>Low cost building and development loans.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>State system</td>
<td>Educational schemes designed to select and prepare doctors for service in underserved areas.</td>
<td></td>
<td></td>
<td>Obligated service in designated area for all graduate doctors. License or visa restrictions limiting immigrant doctors to work in underserved areas.</td>
</tr>
</tbody>
</table>

**UK policies**

Many of the strategies outlined in Figure 1 have been trialled in the UK but few have been subject to thorough evaluation.

**Normative**

Normative schemes are based on the observation that many established GPs feel better fulfilled when they combine general practice with other activities such as research, teaching, management, or hospital medicine (Handysides 1994 a&b). Linking such career development opportunities to service in deprived areas might then enhance recruitment and retention. The single largest scheme of this kind in the UK was the London Implementation Zone Education Initiative (LIZEI) in which £60.9 million was invested in range of “recruitment, retention and refreshment” initiatives during 1995-98 (Hayward & Modell 1995). More than 70% of eligible GPs participated in LIZEI showing there was a high level of demand (North Thames Executive 1997). Evidence in terms of improved
GP recruitment and retention was, however, modest and achieved at high cost (Pringle et al. 2000).

**Coercive**

The principal regulatory control on GP distribution during 1948-2002 was the national Medical Practices Committee (MPC). Local primary care organisations (FHSAs or equivalent bodies) required the permission of the MPC to replenish or increase their GP workforce (Medical Practices Committee 2002). The MPC designated geographic areas as ‘open’, ‘intermediate’, or ‘restricted’ according to the average list size per GP after adjustment for list inflation and indicators of population need. GP recruitment to ‘open’ areas, where average list size was markedly above average, was permitted while GP appointment to other areas was restricted or prohibited.

From April 2002, responsibility for GP recruitment passed to local Primary Care Trusts (PCTs), overseen by Strategic Health Authorities (Department of Health 2002a). PCTs were set targets for expansion in GP numbers based on normative standards (average number of whole time equivalent GPs per 100,000 ‘weighted’ population). PCTs below the average were encouraged to recruit 6% more GPs; those above the average were encouraged to recruit 3% more GPs. There are no specific rewards for meeting targets or penalties for failing to do so.

Although the MPC is credited with achieving greater equity in distribution than has been achieved by other countries such as the USA (Taylor 1998), it has not obviously reduced inequalities in the UK. It remains unclear whether the recent devolution of control over GP recruitment from the MPC to PCTs will promote greater equity. PCT control could generate more locally appropriate and responsive strategies; but gains might be offset by the loss of central control over GP distribution across PCT boundaries.

**Utilitarian**

Financial

A key limitation of the strategy embodied by the MPC was its assumption that doctors could be pushed into underserved areas by restricting access to over served areas. Without additional compensation, however, the net advantage to GPs of locating in underserved areas remained weak (Powell 1990). This limitation was partly addressed by including payments intended to compensate GPs serving deprived populations in the national General Medical Services (GMS) contract.

The GMS contract of 1990 introduced extra capitated payments for each patient on a GP’s list who lived in an electoral ward designated as deprived by the Jarman Underprivileged Area index (UPA8). The UPA8 measured the additional workload incurred by GPs serving patients from deprived communities. Higher levels of deprivation on the UPA8 triggered higher levels of capitated payment. The UPA8 has been criticised for the variables used to construct the index, the way they were combined
statistically to develop an indicator, and the reliance on Census data that is updated only once every 10 years (Carr-Hill & Sheldon 1991).

The new GMS contract of 2004 aimed to overcome some of these technical limitations. Practices receive a ‘global sum’ which is related to list size adjusted for inflation multiplied by a practice ‘weighting’ factor. Practice weights have been developed to reflect variations in GP workload associated with population age, sex, morbidity, mortality and list turnover. Area level adjustments are made to reflect the higher costs of practice in rural areas plus a Market Forces Factor (MFF) representing geographic variation in staff costs (British Medical Association 2003). The 2004 contract, however, also provided a ‘minimum practice income guarantee’ in its first year of operation that will have dampened any effect of income redistribution to deprived areas.

If the introduction of deprivation payments under the 1990 GMS contract had a positive impact on equity of distribution, practices in underserved areas should preferentially have undergone a reduction in average list size per whole time equivalent GP. Whether average list size has declined more in deprived than other areas since 1990 has yet to be investigated. However, there has been no apparent reduction in equity of distribution at health authority level from 1990 onwards (Hann & Gravelle 2004).

Specific area-related financial incentives have also been introduced. The NHS ‘Golden Hello’, introduced in April 2001, offered £5,000 (later increased to £7,000) to a GP taking up an eligible post in an underserved area (Department of Health 2002b). This one-off payment is small relative to the lifetime income of GPs and unlikely to result in sustained reductions in inequity if it is not reinforced by continuing inducements.

Practices located in underserved areas have access to low cost loans to improve their premises. Under the NHS Local Improvement Financial Trust (LIFT), introduced in 2000, up to £1 billion was set aside to upgrade general practice premises, particularly in urban deprived areas where practice infrastructure most needs attention. (Department of Health 2000).

Non-financial

Experiments in job redesign have been carried out in Personal Medical Services (PMS) pilots. Here salaried GP contracts were used to create more attractive and flexible jobs and so enhance recruitment to deprived urban areas. A national evaluation showed the main attractions of salaried contracts to be reduced hours and freedom from out-of-hours and administrative responsibilities (Williams et al. 2001; Gosden et al. 2002). The disadvantage of lower income was offset by these non-monetary rewards as well as the stability of income and enhanced employment benefits. For many GPs, salaried employment led to reduced stress and enhanced job satisfaction. The impact on GP recruitment to deprived areas was modest but positive.

The risk in moving to salaried service is that GPs will have little incentive to sustain their work effort or enhance their quality of care. However, a controlled before-after-study of
GP practices that moved to salaried status suggested these problems may not be realised in practice (Gosden et al. 2003).

PMS sites were preferentially targeted to deprived and underserved areas when they were first introduced (PMS National Evaluation Team 2002) but were rapidly rolled out to other localities across England. By 2004, 37% of GPs (Department of Health 2005) and 36% of practices (Mark Hann, personal communication) held PMS contracts. Research conducted at NPCRDC suggests that, as PMS contracts have become more widespread, their targeting to areas of deprivation has become weaker (Mark Hann & Hugh Gravelle, personal communication). This will in turn have diminished their ability to reduce inequalities in GP distribution.

Comment

Collectively these strategies have failed to produce measurable improvements in the equity of GP distribution at health authority level over recent decades. Some initiatives, such as the GMS contract of 2004 and one-off payments on locating to underserved areas have been introduced too recently for their full effects to become apparent. Current policy relies mainly on PCTs’ ability to introduce new practices to areas of need coupled to financial compensation for practices located in such areas. However, present financial and governance structures do not allow PCTs to remove GMS funds from existing practices to fund new ones, and do not reward PCTs when they successfully increase their workforce in accordance with government targets. PCTs could and should make more use of alternative contractual arrangements, such as PMS or APMS (Alternative Provider Medical Services), to create new practices in areas of need.

Policy alternatives in other countries

Normative

The USA, Australia and Norway have attempted to improve recruitment and retention to underserved rural areas by establishing rural medical schools, preferentially admitting students from rural backgrounds, and/or offering clinical training in rural settings. Evaluation of such programmes in the USA has shown that rural training increases doctors’ preparedness for rural practice and small-town living and predicts longer retention in such areas (Pathman et al. 1999). However there is a lack of convincing evidence that these associations are causal in that the doctors recruited to such schemes are more likely to have opted for rural service in any case (Pathman et al. 1994; Alexandersen et al. 2004).

Utilitarian

Australia and Canada have introduced comprehensive personal and professional support systems for doctors willing to serve in shortage areas. Subsumed under these strategies are assistance in finding housing, financial support for relocation, funding for continuing
medical education, locum provision, and the establishment of rural practice networks. The provision of salaried posts in state-sponsored clinics, as an alternative to private practice, is an additional asset, offering security of income and freedom from financial obligation for practice expenses. Such programmes are credited with maintaining a viable workforce in rural areas but cost-effectiveness has not been rigorously evaluated (Rourke & Strasser 1996).

Canada and USA have introduced schemes to boost GP income in underserved areas. These include: higher fee-for-service payments and recruitment and retention bonuses. In some areas these are coupled to financial disincentives for doctors working in over served areas, including: billing caps; income caps; and discounted fees-for-service. Most schemes have not been rigorously evaluated. However, a review of available research concluded that these strategies were only partially effective and greater attention should be given to non-pecuniary incentives (Barer et al. 1999).

**Coercive**

The USA, Australia and Canada offer money to fund the university education (scholarship scheme) or repay the educational loans (loan repayment scheme) of primary care doctors and nurses in return for service in designated areas of need. Obligated service is for a period of 4 years and there are harsh financial penalties for breach of contract to encourage compliance. Cost-effectiveness has been well evaluated in the USA where the schemes have been in operation since 1970. Both scholarship and loan repayment schemes have proved effective in staffing rural and underserved areas (Langwell et al. 1986a&b), but need to be operated continuously as 60% or more of providers move from their assigned site within one year of completing their obligated service (Pathman et al. 1992). The cost to the federal government of buying one year of obligated service is higher for the scholarship than loan repayment programmes without measurable differences in outcome (General Accounting Office 1995). This is because the longer period of time between recruitment and obligated service for scholars increases the likelihood that their personal and professional circumstances will have changed leading to lower retention rates and higher breach of contract rates.

Increasing levels of indebtedness among graduate health professionals promotes participation in obligated service initiatives (Pathman et al. 2000). As student indebtedness rises in the UK, such initiatives may become a viable option here. The potential impact on equity of introducing a loan repayment scheme in the UK was assessed by Sibbald et al. (2002) A discrete choice experiment, conducted in a convenience sample of 100 newly qualified GPs and nurses in Wales, showed that a one year loan repayment scheme made the rural and urban deprived locations 1.6 and 1.4 times more likely to be chosen respectively than a well served suburban location.

Restricting immigrant doctors to service in designated underserved areas is a second coercive strategy adopted by the USA, Australia and Canada. In the USA, immigrant doctors admitted to study for a period no longer than seven years may be given the right to stay indefinitely if they agree to work in a designated underserved area (the J-1 visa
waiver). Although there is evidence to suggest that immigrant doctors form a higher percentage of the total workforce in rural underserved areas (Baer et al. 1998), most show location preferences similar to those of domestically trained doctors and eventually move to well served areas (Mullan 1995).

Australia has a more robust approach. Medical licensure is contingent on service in designated workforce shortage areas and withdrawn if doctors move to alternate locations. Workforce statistics suggest the policy is successful in targeting immigrants to areas of need (DIMA 2000), but the programme is understandably unattractive to international medical migrants (Young et al. 2003) and little information is available on longer term retention.

Canada operates a similar system to Australia except that restrictions on location operate only for the first few years after immigration after which doctors may locate where they wish. Evidence suggests that many doctors exercise their right to buy out of obligated service or move on completion to the major population centres that are already well served (Barer & Stoddard 1999).

In sum, coercive strategies appear to be successful in targeting doctors to areas of need. However, operational costs are high and poor long term retention means that incentives cannot be provided on a one-off basis.

**General operational problems**

Defining geographic areas of under-service is a major operational challenge.

The choice of geographic unit may produce radically different resource allocation patterns. For example, at high levels of aggregation (health authorities) the north of England is much less well served than the southeast but at low levels of aggregation (enumeration districts) parts of London appear to have the lowest density of GPs (Taylor 1998).

Population size must be adjusted to reflect variations in GP workload associated with patient age, sex, morbidity and mortality among other factors (Gravelle et al. 2003). These data need to be readily available and easily updated. As noted above, deprivation payments under the 1990 GP contract were criticised for their dependence on census data that were updated only once every ten years.

Estimates of workforce supply should be adjusted for time commitment and take labour substitutability into account. Contract commitment (whether full time or part time) can be a poor guide to actual hours worked; but workload surveys are costly to conduct and may be updated infrequently. In the UK, the last formal GP workload survey was conducted in the late 1990s (DDRB 1998). In the modern general practice workforce, nurses increasingly substitute for GPs (Laurant et al. 2005) but are not generally included alongside GPs in measures of workforce density (General Accounting Office 1995).
The national average number of whole time equivalent health professionals per need adjusted population is the usual benchmark against which underserved areas are identified. Area designations need to be graduated to reflect the degree of under (or over) service. This is because, if both marginal and markedly underserved areas are included without distinction in an incentive scheme, the main effect will be to move doctors from the worst to the marginal areas.

Only a proportion of all new recruits to an underserved area will be attracted by an incentive; but all will be eligible to receive it. This means that the additional cost of recruiting one extra GP will be lower, the higher the level of the incentive. Knowing the comparative effectiveness of different strategies for inducing GPs to locate to underserved areas is therefore of great importance to policy-makers. However, as the above review suggests, few strategies have been thoroughly evaluated and there is almost no information on comparative performance. It is clear, however, that no one strategy is fully effective and a blend of approaches, incorporating both pecuniary and non-pecuniary incentives, is needed.

References


Barer ML, Wood L, Schneider DG. (1999) Towards improved access to medical services for relatively underserved populations. Canadian approaches, foreign lessons. Vancouver, Canada: Centre for health services and policy research. (As reported to the annual conference of the International Medical Workforce Group in 2002)


