The equity implications of health system change in the UK

Citation for published version:

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Health Inequalities: Critical Perspectives

Publisher Rights Statement:
This is the Accepted Manuscript version of chapter 11 in the book 'Health Inequalities: Critical Perspectives' published by Oxford University press, edited by Katherine Smith, Sarah Hill & C. Bambra, published on 26/11/2015.

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
The Equity Implications of Health System Change in the UK
By Mark Hellowell

Abstract
This chapter considers the role of healthcare in health inequalities debates. Drawing on international literature concerning equity in health systems, it considers the challenge to equity posed by recent healthcare reforms and the emergence of severe funding constraints. It argues that, while the former challenge has dominated scholarly debate in recent years, the latter represents the greater threat to equity. While acknowledging that healthcare is only one of an array of factors that influence health inequalities, the chapter challenges the tendency of researchers in this area to regard a focus on healthcare as a distraction from a more appropriate emphasis on the social determinants of health, making the case for stronger links between research in healthcare, health inequalities and the social determinants of health.

Word count: 5,738

1. Introduction: healthcare and health inequalities
The National Health Service (NHS) was created in 1948 with the aim of ensuring that access to healthcare would depend on need and not ability to pay. “The essence of a satisfactory health service” wrote the Minister of Health Aneurin Bevan, “is that the rich and the poor are treated alike, that poverty is not a disability, and wealth is not advantaged” (Webster 2002). Since then, the NHS has provided a globally prominent example of how socioeconomic inequalities in access to healthcare can be reduced through progressive tax-based financing and government stewardship (van Doorslaer and Wagstaff 1992; Wagstaff et al. 1999; Dixon et al. 2003).

Of course, healthcare – the delivery of goods, services and activities that maintain or improve health - is only one of a complex array of factors that determine health. For this reason, some health inequalities researchers (particularly those working in contexts in which healthcare is free-at-the-point-of-need, such as the UK) have tended to regard a focus on healthcare as a distraction from a more appropriate emphasis on the social determinants of health (Marmot 2010). However, studies of the relationship between institutions and
cultural values show that institutions that embody values of fairness and social justice can help to elicit and sustain such values in wider society (Svalfors 2010). The commitment to social justice embodied in the NHS is likely to play a central role in marshalling the societal efforts required to address health inequalities in the UK.

Recent changes in the structure of the NHS have raised questions about its ability to continue to meet social justice objectives. Some of these relate to deliberate adjustments by policymakers to the way the service is configured, including reforms to modes of governance and resource allocation, particularly in England.\(^1\) Others relate to external factors (political, demographic and macroeconomic) that are present to some degree across the UK. To a large extent the marginalisation of healthcare in the inequalities literature has arisen in a context in which the socioeconomic gradient in healthcare access has been limited due to state intervention (van Doorslaer and Wagstaff 1992). Literature concerning health inequalities in contexts where this is less true, such as the USA and many low and middle income settings, have tended to place a much stronger emphasis on healthcare, as Chapters 4 and 5 note. If, as seems likely, inequalities in health status are combined with an increasing degree of inequity in coverage and access in England, and perhaps in the rest of the UK, the healthcare system is likely to become a more salient topic for health inequalities researchers in future.

This chapter draws on the international health systems literature to assess the significance of the two most prominent categories of change in the NHS policy context from an equity perspective. To provide a foundation for subsequent analysis and discussion of these changes, section 2 reviews how the principle of equity in healthcare is conceptualised by scholars and operationalised by policymakers. Section 3 provides an analysis of the current phase of structural reform in the NHS. This has been the focus of substantial policy debate and academic criticism - with one article describing the reforms as “paving the way for the

---

\(^1\) Since devolution in 1999, there have been four separate healthcare administrations in the UK – one for each of England, Scotland, Wales and Northern Ireland. There has also been some divergence in the policies of these administrations (especially between England and the rest) in terms of the emphasis placed on competition between healthcare providers as a means of improving the efficiency and quality of services.
introduction of a US-style health system by eroding entitlement to equality of healthcare provision” (Pollock et al. 2012, p.387). However, we explain in section 4 how the international health systems literature directs our attention to the central importance of funding constraints. We argue that this latter category of change poses a major threat to the mechanisms of progressive revenue collection and risk pooling – institutions that, to paraphrase Aneurin Bevan, comprise the central pillars of an equitable health service.

2. Equity in the Health System: An international perspective
This section provides an overview of the principles of equity in healthcare systems as a foundation for subsequent analysis and discussion of systemic change.

2.1 Equity in healthcare
In the healthcare system, equity concerns the fair and socially just distribution of health services. Markets in healthcare will fail to meet this requirement, as they allocate goods and services on the basis of their demand for these things - a person’s ability and willingness to pay – and not on their need for them. Low-income members of society lack ‘demand’ because they are unable to afford the market price of healthcare. This observation underlies the well-known inverse care law, which states that the availability of good medical care tends to vary inversely with the need for it in the population served. This inverse care law operates more completely where medical care is most exposed to market forces, and less so where such exposure is reduced (Tudor-Hart 1971).

How the equity requirement should be put into operation through policy action is the subject of intense and persistent debate within normative economic analysis. Olsen and Rogers (1991) suggest the appropriate objective of policy is to ensure equality of access, in which case everyone in society is equally able to obtain or make use of healthcare. Others point out that it is not access to healthcare itself that is of concern, but equal access to needed services – i.e. services that are effective (and perhaps cost-effective) in attaining a higher state of health (Culyer 1989). In contrast, Culyer and Wagstaff (1993) observe that the normative basis for equity in healthcare resides in its ability to promote health, and that the appropriate equity goal is therefore to generate equality of health.
Different notions of equity are also evident at the policy level in terms of how the commitment to equity is put into operation through the allocation of resources to different geographies (Bevan 2009). For example, different population-based resource allocation models are employed by each of the four NHS jurisdictions in the UK (Carr-Hill and Dixon 2006). We also observe important differences over time within the four NHS jurisdictions (Barr et al. 2014). The traditional NHS approach has been “to secure through resource allocation equal opportunity of access for people at equal risk” (Department of Health and Social Security 1976, p.7), which is roughly consistent with Culyer’s (1989) principle of equal access to needed services, and remains at the core of the capitation approaches in Scotland, Wales and Northern Ireland.

However, the Labour government of 1997-2010 sought to use the capitation formula to contribute to a reduction in avoidable health inequalities in England by targeting more resources at deprived areas. This may be regarded as a move towards an ‘extra-welfarist’ approach and a focus is on the equal distribution of health.\(^2\) Notwithstanding this diversity in normative/policy frameworks, it is notable that in each case there is an emphasis on distribution of services according to need and health status and not ability to pay, such that there is (broadly) equal access to needed care by all individuals in society.

2.2 Safeguarding equity in the health system

Unpredictability in the need for medical care generates an important role for insurance – either public or private - in the healthcare system (Arrow 1963). Where insurance institutions are effective, individuals are able to pool the financial risks associated with healthcare, and those who fall ill are able to obtain the healthcare they need. Unlike most goods and services, the production of insurance is an inherently collective activity. It is impossible for a single individual to produce insurance, except in the limited sense of self-insurance: maintaining a fund to pay for future healthcare costs through saving (Ehrlich and Becker 1972). An individual can produce insurance only by joining together with others, even if this is through market-based transactions, to form a risk pool. In a well-functioning

\(^2\) After a review of allocations policy in 2013, NHS England implemented a new formula that gives less weight to deprived areas. This has been the subject of criticism by health inequalities scholars (Barr et al., 2014).
market system, it is possible to equalise risk between the sick and the well in this way. However, in order to additionally equalise risk between different socioeconomic strata, government intervention is necessary.

A system of private insurance markets is unable to provide the broadly equal access to needed care that is called for in the previous section. As noted, markets allocate goods and services on the basis of demand - a person’s ability and willingness to pay. Low-income members of society lack ‘demand’ because they are unable to afford the market price of private insurance.\(^3\) This is particularly so given the inverse relationship between socioeconomic status and health status, which make premiums in a private market higher for those with fewer economic resources (and normally greater need).

When insurance markets form risk pools by voluntary enrolment, inequities are exacerbated by information asymmetries between providers and purchasers of insurance (Hurley 2001). Risk-selection arises when suppliers of insurance selectively enrol low-risk individuals (cream-skimming) or when high-risk consumers seek out more generous insurance (adverse selection). The former type of risk-selection leads directly to inequities as high risk members of society are unable to buy insurance. The latter type of risk-selection also reduces the breadth of access since even people who are willing to buy insurance at a premium that reflects their risk status may be unable to do so.

For these reasons, significant public financing is a necessary condition for achieving distributional equity in terms of coverage and access. This will work most effectively in equity terms when the level of coverage is perceived to be adequate by better-off groups, thus ameliorating demand for private health insurance (which contributes to inequities in coverage and access).

A sufficient condition is that the organisation and delivery of services support this. The extent to which these conditions are achieved has been termed effective coverage (Kutzin

\(^3\) Due to the high cost of medical care, the actuarially fair price of insurance is likely to be too high for low-income members. Due to various categories of market failure (including, at a minimum, adverse selection, moral hazard, monopoly, transaction costs), the market price is also likely to be higher than this fair price.
No country succeeds in completely eliminating shortfalls in effective coverage; gaps exist because not all individuals in a society can be aware of all of their needs for services. In addition, new and more expensive diagnostic and therapeutic technologies continuously emerge, which generate new need (in the sense that there are individuals in society who may secure health benefits from them) that cannot be immediately satisfied. In addition, empirical evidence has shown that there is also a socioeconomic dimension to this shortfall in coverage, even in countries such as the UK that have largely been successful in attaining universal financial protection (Cookson et al. 2012).

It may therefore be more useful to think of equitable access and coverage as a direction rather than a destination. The international health systems literature indicates that the following policy actions are likely to move health systems away from achieving equity goals (van Doorslaer et al. 2000): (i) a reduction in the proportion of revenues for health that come from compulsory prepaid funds (Wagstaff et al. 1999); (ii) a reduction in the extent to which the budgets of healthcare providers and purchasing agencies reflect the risk status of the populations they serve (Kutzin 2001); and (iii) a reduction in pooling which compromises the redistributive capacity of prepaid funds (Kutzin 2013), thereby enabling a lower level of financial protection and equity in the distribution of resources from a given level of resources.

In addition, the pattern of income redistribution associated with public financing through progressive taxation is likely to serve wider equity objectives. Private insurance with risk-rated premiums provides no ex ante redistribution of income (though, ex post, there is redistribution from the well to the sick within the pool). In contrast, public financing embodies ex ante redistribution from the wealthy to the poor (who have on average worse health) as well as ex post redistribution from the healthy to the sick (van Doorslaer and Wagstaff 1992).

3. The UK Conservative-Liberal Democratic Coalition Government’s reforms in context

The NHS has seen regular structural reorganisations since the Thatcher era. An important change occurred in 1991 when integrated NHS organisations were disaggregated into
geographically-defined entities (district health authorities, primary care groups, primary care trusts) which provided primary healthcare to the local population and contracted with local hospital organisations (NHS Trusts and, later, Foundation Trusts) to deliver acute care services. That structure has, broadly, been maintained in England, Wales and Northern Ireland, whereas Scottish NHS services were re-integrated on a regional basis in 2003. Additionally, since 2003, a small proportion of NHS acute services have, in England and Scotland, been delivered by private sector providers, often in Independent Sector Treatment Centres, which provide diagnostic and elective healthcare under contracts with government.

The nature of structural reform in England has been consistently market-oriented since 2002. One important component of reform was a change in the basis of payments to hospital organisations under contracts with local healthcare commissioners. Prior to 2002, payments were made to hospitals on the basis of prospective global budgets, which were set according to the estimated costs of the treatments to be provided. In 2002, a prospective case-based system - ‘Payment by Results’ - was introduced, in which providers are paid a fixed price per ‘finished treatment episode’, with this price designed to approximate the average cost of delivering the treatment across the NHS (Hellowell and Pollock 2007). As referral decisions are to some degree made according to the choices that patients (and GPs) make, hospitals have an incentive to attract patients from competitors, at least where the associated treatments are likely to generate surpluses.

Since the prices in this NHS market are fixed under the Payment by Results regime, the surpluses of provider organisations can be maximised by: (i) reducing the cost of treatments (i.e. increasing technical efficiency); (ii) increasing the demand for services by enhancing patients’ perceptions of service quality (i.e. enhancing clinical quality and/or other aspects of quality such as hotel services); and/or (iii) favouring access for patients for whom the cost of treatment is lower than the price (i.e. increasing inequity through ‘cream-skimming’) (Propper et al. 2006).

Under the 2012 Act, all 152 primary care trusts in England were abolished and replaced by groups of commissioning organisations led by GPs and other clinicians, who are in turn
accountable to a central NHS commissioning board. This works to amplify the emphasis on choice and competition, especially in relation to hospital care, with patients and GPs able to choose to be treated by any accredited provider in the public or private sector. Providers are subject to EU competition law and the scrutiny of an economic regulator, Monitor – which was previously the financial regulator of Foundation Trusts, but now has powers similar to those held by regulators of the privatised utilities such as water, energy and telecoms. Under the Competition Act, Monitor has powers to prevent “anti-competitive behaviour in the provision of healthcare services for the purposes of the NHS which is against the interests of people who use such services” (Department of Health 2012, p.36).

The central motif of these reforms is *marketisation* – i.e. a shift from bureaucratic planning and cost-based resource allocation to a model in which funding is allocated to healthcare providers according to their ability to attract patients within a competitive market. They also involve a degree of *privatisation* – i.e. a change in the ownership of healthcare assets as non-state actors expand their production of healthcare. This is currently of some significance; it is estimated that about 5% of NHS expenditure in England is spent on healthcare services supplied by private companies and voluntary organisations, about twice that in Scotland (Laing and Buisson 2013).

However, it is important also to highlight what the reforms do *not* do. They do not involve a change in the way money is raised to pay for care; and they leave in place the principle that resources should continue to be allocated on the basis of population need. Concerns have been raised that the market-oriented reforms may increase inequity, as providers select patients according to the cost of treatment (which may be associated with socioeconomic factors) (Hunter 2009). However, evidence suggests this has not happened in practice (Cooper *et al.* 2009, Cookson *et al.* 2012). Although the combination of markets, private ownership, progressive tax-financing and needs-based planning certainly make the ‘new’ NHS in England highly distinctive in a global context, the components of the system that are crucial to achieve equity have been retained.

The capacity of any health system to provide comprehensive healthcare free at the point of use depends on the level of public funding it receives. This is limited by a country’s macroeconomic performance, a government’s fiscal policy decisions, and how a government allocates its budget to different areas of expenditure. For reasons explored in more detail below, the demand for NHS services has risen over time despite consistent increases in real prices. As a result, public expenditure on the NHS increased in real terms by an average of 4% annually between 1950 and 2011 (Office of Health Economics 2012). NHS spending also increased as a share of GDP, from 3.5% in 1950 to 7.7% in 2011.

Historically, periods of fiscal contraction have led to a decrease in access to healthcare – i.e. an increase in the extent to which supply is rationed. For example, in the five years from 1950/51, the NHS experienced an average real terms budget cut of 2.4% per year, resulting in the introduction of charges for prescriptions, dental services and spectacles (and one high-profile resignation – that of Aneurin Bevan as Minister of Health). In 1975/76 to 1979/80, the NHS budget grew by an average of just 1.3% a year in real terms. In the context of rising demand and higher relative prices, this prompted the onset of a long period in which waiting times for hospital services increased, a trend that intensified through the period of Conservative Government between 1979 and 1997 (in which real-terms NHS spending grew below the trend, at 3.3%) and was not reversed until 2003 (Appleby et al. 2009).

During the 2000s, NHS spending grew at its fastest ever rate, at an average of 6.6% per year between 2000 and 2008. Along with the reduction in waiting lists and decreased waiting times, these additional resources led to higher levels of activity in hospitals and primary care and better health outcomes, including improved survival and improvements in the control of chronic conditions (Bojke et al. 2013).

Given the historic record, what do current and projected levels of NHS spending tell us about the likely future level of supply? The 2010 Spending Review outlined cash spending plans for the period 2011/12 to 2014/15 that were just sufficient to freeze NHS spending in real terms (Office for National Statistics 2014). The current NHS England planning framework assumes that public sector health expenditure in 2020/21 will fall, as a
proportion of GDP, from a peak of 7.7% in 2010/11 to 6% in 2020/21 – equivalent to the level last seen in 2003 (NHS England 2013).4

Assuming these plans are implemented, they will require a level of cost containment that has neither historical precedent nor international parallel. Already, there is growing evidence of financial pressures building in the NHS and an emphasis on rationing-through-waiting as a means of managing demand (Appleby et al. 2014). Of course, the level of NHS spending is ultimately a result of political choices which are impossible to predict. But the choice is constrained by the fiscal planning framework which all three political parties currently support (at least in terms of current spending). It is estimated that this requires an average annual real reduction in departmental spending of 3.7% to the end of 2018/19 (Institute for Fiscal Studies 2014). It therefore remains unclear if the relative protection from cuts that the NHS has experienced since 2010 can or will be sustained over this period. The Labour Party’s plan for a “zero-based review” of public budgets that covers “all areas of public spending, including those that have been protected in the current Spending Review such as health” (Labour Party 2014) suggests that a change in government is unlikely to ease the current budgetary constraints on the healthcare system.

However budgets are restricted, demand for healthcare is certain to grow. Estimates of the impact of the various drivers of health spending vary (Newhouse 1993; Cutler 1995; Oliveira et al. 2006), but there is a general view that real growth of 3% to 6% a year is required to allow the NHS to meet the growth in demand. On this basis, NHS England has attempted to model the level of productivity improvements required to close the funding gap in the context of a real-terms freeze in funding. This calls for productivity growth of £20bn between 2011/12 and 2014/15 – equivalent to average annual productivity growth of around 5% a year (NHS England 2013). Reflecting this, if the NHS continues with the expected levels of funding, savings of £20bn will leave a funding gap of £30bn between 2013/14 and 2020/21; and the gap will increase from then onwards (Department of Health 2013).

4 Although these estimates relate to England only, they are relevant for all four healthcare jurisdictions as these determine the NHS component of the block grant paid by HM Treasury to the devolved administrations.
Setting aside the uncertainties over projections of healthcare demand, there are reasons to believe that the expected shortfall is an underestimate. As the only major purchaser of healthcare in the UK, the NHS has considerable market power relative to providers (healthcare organisations and the healthcare workforce). Current progress towards the productivity target has been driven by means of a sector-wide pay freeze that has been made possible by the monopsony status of the NHS. In the long-term, productivity improvements will be more difficult to achieve. All developed world healthcare systems have been shown to be subject to what the economist William Baumol has called the “cost disease” (Baumol et al. 2012). Because health services rely on a significant ‘handicraft’ component, it is hard to replace labour with capital, and the rate of productivity growth is therefore lower than the average in the economy. It follows that prices will rise faster in the health sector compared to the economy as a whole. Between 1974/5 and 2007/8, pay and prices in the NHS in England rose by around 1,000% - nearly twice the increase in the GDP deflator, the main measure of inflation in the economy (Appleby et al. 2014). Higher relative prices imply that health spending must grow relative to inflation if demand for care is to be met.

4.1 How will the NHS seek to address the funding gap, and what are the implications for equity? The expansion of rationing by waiting
As noted above, the traditional response to funding constraints has been to intensify the degree of rationing. In a market system, price is the factor that balances supply and demand, providing signals to production and consumption. A higher price rations demand and stimulates supply so that in equilibrium balance is achieved. Where public financing and provision of healthcare remove the monetary price to the consumer, other balancing factors are needed. In relation to the NHS, Klein describes the existence of an implicit “bargain between the State and the medical profession”, in which “Politicians in the Cabinet made the decisions about how much to spend; doctors made the decisions about which patient should get what kind of treatment” (Klein 2010, p. 61).
The prominent problem of rationing by waiting has been part of the management of demand in the NHS since 1948 (Bevan 2009). As NHS finances have become increasingly constrained, waiting times have once again begun to grow. In April 2014, the number of people waiting more than 18 weeks for an operation reached three million, the highest number for six years (NHS England 2014). In May 2014, the emblematic cancer treatment target was missed for the first time since it was introduced in 2009 (Department of Health 2014).

4.2 An expanded role for private insurance

It seems reasonable to assume that the lower the capacity of the NHS to address healthcare needs, the greater the scope for a private healthcare sector to exist. The econometric evidence highlights the significance of waiting times in particular as a factor in determining the demand for private healthcare (and, consequently, private health insurance) in the UK. Besley and colleagues (1999) matched survey data with administrative data at health authority level in England over the period 1986-1991. They found that an increase by one person per 1000 in the proportion of patients that had to wait more than 12 months increased the probability of buying private health insurance by 2%. Similarly, King and Mossialos (2005) used survey data over the period 1997-2000 and found a significant positive effect of inpatient and outpatient waiting times on the demand for private insurance. It is therefore likely that the role of private insurance will increase in the coming years due to the constraints on NHS funding and supply and, for the reasons outlined in section 2, a health system in which private insurance plays a more prominent role will lead to inequities, with comprehensive coverage for members of society with the ability and willingness to pay, and more restricted coverage for low-income members and individuals with lower health status (and greater need). This is a context in which the healthcare system itself is likely to become a more salient topic for health inequalities research in the future.

Although, during the 2000s, there were two key developments that changed the ways in which the NHS in England rations care. First, the Labour Government’s overriding objective of the increases in funding the NHS in England in the 2000s was to end crude rationing by waiting. The government’s targets for the time patients had to wait from seeing a GP to being admitted to hospital for an elective operation were reduced from over 2 years (in 2000) to 18 weeks (in 2005) (Thorlby and Maybin, 2010). This was achieved by regimes of performance management (based on command and control strategies of ‘targets and terror’). These approaches were not introduced in the NHSS in Scotland, Wales and Northern Ireland where, although each too had similar increases in funding, there was less progress in reducing waiting times (Connolly et al., 2010).
5. **Conclusion: Why healthcare is a priority for health inequalities research**

There is a clear asymmetry in the degree of scholarly attention afforded to the reform of the health system in England relative to the issues related to the broader (and UK-wide) contextual factors. To a degree, this is unsurprising: deliberate policy actions often receive greater public attention than external factors, even when the latter may have far greater influence on social and economic outcomes (Easterly 2014). However, policy inaction is all that is required to create a financial crisis in the NHS that will weaken its ability to provide comprehensive free healthcare in the near future. As Richard Smith, a former editor of the *BMJ*, observed (1999): “Most institutions on the scale of the NHS end not with a bang but with a whimper ... one possible endgame is that the middle classes lose confidence in the service and begin to make other arrangements.” This chapter argues that this possibility is worthy of greater study and reflection by advocates of equity in health. If their voices are not heard in debates around how to respond to the growing financial crisis in the NHS, others - with far less interest in equity – are likely to dominate.

**REFERENCES**


