Allyson M. Pollock and David Price

The break up of the English NHS: the new market bureaucracy needs information systems based on members and not geographic populations


Michael Quarterly 2011;8:460–75.

Universal health systems are characterized by administrative functions that focus on needs assessment and resource allocation across geographic populations so as to ensure comprehensive coverage that is country and system-wide. This is because in universal systems financial risk is pooled at government level. In market systems, however, financial risks are allocated (spread) across different parts of the system through market contracting. Administrative functions in these systems focus on risk pricing and segmentation among providers, members or enrollees. Using the example of the English National Health Service (NHS), we argue that the information systems required to enable risk segmentation are different from those underpinning comprehensive care. Administration is non-geographic in market bureaucracies and geographic in comprehensive systems. This leads to a shift in information requirements that poses difficulties for resource allocation and service planning, and therefore for universalist goals.

Introduction

Since 1948, the NHS has assumed the responsibility for the risks and costs of health care for all its citizens. However, a succession of statutory changes dating back to 1990 and culminating in the Health and Social Care Bill currently before parliament (1) have increasingly undermined this national responsibility. If the latest bill becomes law, risks and costs will be spread among government funders, local authorities, insurers, providers, the public and ultimately patients.

The Bill introduces new structures taken from the American health maintenance organisation (HMO) industry in which commercial organiza-
tions insure and provide care for a selected membership rather than geographic populations (2). The introduction of these structures to the NHS is accompanied by a major overhaul of administrative structures involving the substitution of market for public bureaucracies. The new market bureaucracies will give discretion to commissioners to define the scope of NHS services and to some extent to select the population for which they purchase care. They will also allow providers to select the publicly financed services for which they tender. Deregulation of this type facilitates the emergence of business strategies based on risk selection.

Such has been the strength of public opposition to the Bill that in April 2011 the prime minister suspended the legislative process for three months and instigated a special forum to hear and respond to public and professional concerns. However, as the Bill now continues its passage through parliament, and in spite of three hundred significant amendments by the Government, criticism continues. The government proposes to abolish the Minister’s duty to provide comprehensive care throughout England and to create an independent body, the NHS Commissioning Board, to oversee commissioning. The fundamental concerns is that abolition will result in a loss of the universal and comprehensive character of the NHS and pave the way for a switch from taxed-based to insurance-based financing and user charges (2).

In this paper we focus on the new administrative bureaucracy which is necessary to support market reform of this type. We show that information is key to our understanding, specifically the change in the unit of analysis associated with a shift in responsibility from area-based populations to membership-based systems akin to insurance pools. This analytical change also facilitates a transfer of risk from government to patients.

Market bureaucracies
For the last 25 years, in common with many health systems, NHS reform has been profoundly influenced by market theory (3–9). Among the most widely pursued policies internationally are the substitution on economic grounds of competing, commercial providers for publicly administered government units and the introduction or extension of competition among third party payers and insurers of health care.

The introduction of these markets has implications for public health bureaucracies in terms of their control over the various components of health systems. Whilst the impact of markets on public health functions such as workforce planning and financial incentives has attracted considerable research (10–13), less attention has been given to their impact on
population data. However, market systems require a different analytical framework from publicly integrated ones and this in turn affects the data systems and the information that is used.

Universal health systems are characterized by administrative functions that focus on needs assessment and resource allocation across geographic populations and system-wide. This is because in universal systems financial risk is pooled at government level and the whole population must be comprehensively covered. The NHS is a prime example of the geographic focus that has grown up around a virtually monopolistic provider of health care and encouraged extensive data collection and analysis on population health, health inequalities and access to health care by social class and ethnicity.

These are the data that the market analytic threatens. In market systems financial risks are allocated across different parts of the system through market contracting and there is no duty to provide services on a comprehensive basis or to collect data on a geographic basis. Instead, administrative functions in market systems focus on risk pricing and segmentation among providers, members or enrollees. In advocating market or “strategic purchasing” as an important tool for improving health system cost efficiency, the authors of the World Health Report 2000 (14) largely overlooked the implications for information of this different analysis.

The impact of markets on some public health functions is relatively well understood. For example, market-related, unmanaged outflows of health workers are known to damage health systems, undermine planning projections and erode the skills base, according to the WHO (10). The workforce crisis among trained personnel in resource-poor countries has recently been recognised as an issue in international aid (11–13). Following the World Health Report 2006 (12) which estimated a staffing shortfall of 4 million in developing countries, the first global forum on workforce issues was held in Kampala in 2008.

Another example are the lack of routine data in marketised health systems which have under-developed population-based information systems because public authorities do not have planning powers or resources to justify their collection (15) or because large proportions of the population are not covered. The efficiency of planning systems depends to a large extent on routine data collection and the power of public bodies to require data returns from providers. Loss of this capacity can be seen very directly in the UK’s long-term sector where in 1991 the private sector and local authorities were allowed to pass risks back from the NHS to the public via new eligibility criteria involving a means test. The planning, resourcing and provision of care of older people to this day is left to the market and it is now impos-
sible to assess on a population basis the scale and distribution of resources devoted to the various elements of care because data that would allow these assessments are no longer available.

The effect of privatization on data availability for long term care suggests a wider question, namely the role that data can play when governments seek to withdraw from universal policies. In long-term care, planning data did not merely cease to be available; it was no longer officially needed because the government had largely relinquished responsibility and thus financial risk for care. Normative change of this sort is undoubtedly assisted by absence of evidence of unmet need and hardship.

The shift from integrated public to market health care systems requires innovations in the identification of risk which in turn change the basis of health information. For example, markets require incorporation of providers so that risk can be allocated through commercial contracting; they involve changes to revenue and capital accounting so that commercial loans can be substituted for government grants; they typically include revised reimbursement mechanisms based on price signals in which case- or activity-based payments to providers are substituted for block grants (16–19); and to facilitate ‘consumer choice’ or incentivize providers they are commonly associated with new performance management frameworks with a focus on firms or providers rather than whole populations living in contiguous areas (20, 21).

How the integrated approach of risk pooling is achieved in a universal public bureaucracy
Throughout most of its sixty-three year existence the NHS has had a duty to employ administrative structures that promote equity and redistribution through resource allocation, service planning and needs assessment. The issue of how government deals with catastrophic risks and costs on behalf of its citizens was the normative problem addressed by the introduction of the welfare state in the UK and the systems of benefits which underpinned it (22). Accordingly, the structures of the NHS were originally based on contiguous or seamless geographic tiers of administration designed to ensure universal coverage. Service providers were directly administered and integrated into the organisation with responsibility for meeting needs and planning services. There was no billing, invoicing or contracting and, crucially, no selection or denial of care on basis of place of residence or ability to pay. Instead, resource allocation methods and service planning dealt with universal populations. The denominator was always all the citizens living within a geographic area drawn from census or census estimates; the numerators comprised the sub-
groups of specific interest which always related back to the whole population. These methods proved to be highly efficient and the NHS was one of the lowest cost universal systems in the world (23).

The risk allocation approach of a market
Market and business strategies focus on individuals or groups of individuals as customers or members of insurance pools. The denominators here are members or enrollees, a provider’s customer list or an insurer’s list of potential claimants. Under the new Bill, it will be patient registrations belonging to the lists of general practitioners (family doctors). To maximize income and profit, market actors must now engage in complicated risk selection strategies that enable them to avoid contracting for high risk patients and treatments or markets with low profitability. Examples of risk selection methods include the differential premiums charged by insurance companies or the range of tiered benefit plans offered to consumers: a minimum package, or higher benefits price with risk sharing or coinsurance; time limits on care or an annual cap on attendances for example. Risks are identified in market contracts which function as the legal means by which risks are allocated and paid for. There is little or no empirical verification of the cost efficiency claims made for market structures (24–26) whilst their effect on equity is largely unevaluated (27, 28).

The Risk shift
Under the NHS Bill risk analyses will become central to both the public and private side of contracting. Private firms require services and patient lists to be unbundled so that they have the ability to select on commercial grounds. Meanwhile, in an attempt to counter risk selection of this type, or its consequences, the government will seek to risk adjust provider incentives (for example, through adjustments to diagnosis related group (DRG) reimbursement) or to equalize risk among different payer organizations with disparate, and not necessarily equally high risk, memberships (29, 30).

A shift to market bureaucracies requires a change in the information which underpins the methods for funding and financing In the first three of five examples we show how the information requirement changes and in so doing the unit of analysis from comprehensive area based populations to members or providers which do not provide comprehensive care to all. In the last two examples we show the implications of this change in information for performance measurement and comparisons of health systems.
1. Changes in system of resource allocation: from areas to members

(i) The current system of resource allocation

Since 1948, area-based funding has been the method of allocating resources across England for the NHS. Area-based formulae have been used since the 1970s to distribute resources fairly among the “local populations”, “catchments” or “resident populations” of PCTs (31):

“Primary care trusts (PCT)s are responsible for funding NHS hospitals, GPs and other health care services for their local populations.

The Government, through the Department of Health, provides the money to all of the 151 PCTs across the country to fund these health services.

The Department sets PCTs’ budgets in advance, mainly on the basis of a formula to calculate each PCT’s fair share of the total available budget for England.

In 2011-12, the total health budget for PCTs was £89 billion and the Department has to find a way for allocating this between PCTs in a fair way.”

Since its inception, the NHS has been based on the principle of equal access for equal need. This principle is embodied in two longstanding objectives for resource allocation from the centre to local health services:

i) To distribute resources based on the relative need of each area for health services. Currently, this objective is to enable PCTs to commission the same levels of health services for populations with similar needs;

ii) In addition, to contribute to the reduction in avoidable health inequalities (32, 33).

The resource allocation formulae devised to meet these objectives includes

- the age profile of the population (localities with more elderly populations have higher needs, all else being equal);
- additional need over and above that relating to age (localities with less healthy populations and higher levels of deprivation have higher needs, all else being equal); and
- unavoidable geographical differences in the cost of providing services - the Market Forces Factor (MFF) (it costs more to provide the same level of services in high cost areas such as London and the South East).

The formulae have been the subject of revision for many decades.

(ii) How the resource allocation system will change under the Bill

Under the Health and Social Care Bill 2011 the Secretary of State’s principal duty remains unchanged and under Clause 10[2] the duty to “arrange for the provision of services to such extent as it considers necessary… to
meet …. reasonable requirements” (1) is transferred to commissioning groups. Commissioning groups however will not be contiguous geographic area based administrative structures, they will be membership organisations.

Moreover, the Department of Health Memorandum to the Delegated Powers and Regulatory Reform Committee (34: paragraphs 53 and 57) states that the intention of Clauses 10 and 11 is to allow commissioning groups discretion with respect to the selection of patients and services and that this discretion will be limited by regulation.

Thus responsibility for provision will no longer be to all persons in an area but only “persons for whom it [the CCG] has responsibility” (1); nor will it involve purchasing all services deemed part of a comprehensive health system.

The Bill has been anticipated and so too have the new structures for quite some time. For the last few years behind the scenes the civil services has been at work to effect that transformation. In 2010, and well in advance of the legislation the Secretary of State instructed the Advisory Committee on Resource Allocation (ACRA) from 2013 to switch from PCT and area-based populations to GP registrations in deriving its new formulae (35). This is in recognition that CCGs will not have the same geographical basis as PCTs, because they are able to include patients registered with GP practices anywhere in England. They are only required to have “a sufficient geographic focus” to be able to take responsibility for agreeing and monitoring contracts for locality-based services (such as urgent care services), to have responsibility for commissioning services for people who are not registered with a GP practice, and to commission services jointly with local authorities” (36, p. 29). However, “sufficient geographical basis” is not otherwise defined.

‘Fair allocation’ objectives need to be embodied in the new NHS system, but it is made almost impossible by the loss of responsibility for a defined geographic population. It is unclear, for example, if and how a measure such as DFLE could be derived for clinical commissioning groups, given their irregular, non-geographic overlapping mosaic of footprints.

Since CCGs will no longer be geographically focused, the loss of area-based population responsibilities has serious implications for the stability and accuracy of measurements of need and the equity of resource allocation and funding. In the absence of area-based planning information, CCGs will be able manage their risks and costs in different ways.
2. Capital planning protocols – from area-based needs to provider’s affordability under resource accounting and the private finance initiative

i) From capital grants to debts and private finance

The problems of fairness of funding highlighted in the Bill are not new to the NHS. They have already been played out in capital allocations for new hospitals under the private finance initiative (PFI), a policy introduced in 1991 according to which new capital is distributed among hospitals on the basis of provider finances rather than area needs criteria. Before PFI NHS capital budgets were allocated as block grants on the basis of need and regional development plans. Financial reforms in 1990 effectively transformed grants into loans that local service providers were responsible for repaying to the Treasury. PFI exploited this development by transforming debts to the Treasury into debts to private consortiums with local service providers still responsible for repayment. The result was that budgetary allocations for new building were no longer awarded to areas on behalf of the whole population on the basis of need but to providers to finance loan repayment (16–19) on the basis of what commissioners thought they could afford to pay from their revenue budgets.

NHS hospitals award PFI contracts to the private sector to design, build, finance and operate new facilities. Because investment and operating costs are paid for out of hospital operating budgets, PFI required before its introduction the creation of a special revenue stream that could be directed to pay for capital. This in turn required the government to change its financing method for hospital building from capital grants to loans, and it required hospitals and services to put their accounting on a commercial basis in order to reflect the new element of capital charging in their financial management.

ii) Shifting the risk

Crucially the impact of the PFI was to shift risk of capital budgets from national and regional level to local providers, making them responsible for the affordability of capital.

This type of risk devolution impairs the geographic focus underpinning service planning because it makes providers’ debts and debt servicing a higher priority than funding capital health care needs (17). In early PFIs planning took place outside public health as teams of management consultants were brought in to model bed closures using implausible productivity targets in order that the clinical budget could be redirected to pay for capital at the expense of clinical care. Thus PFI planning and allocations...
for capital turns on the question of what a hospital can afford to pay (or rather borrow), and not on the health care needs of a geographic population for capital (16–19).

PFI is also based on a form of revenue unbundling necessary for risk allocation. A typical PFI contract involves separating out revenue streams for capital, maintenance and some ancillary services in order that risks can be transferred or allocated in different proportions and to different parties. This is because the PFI contract involves specifying and pricing risks that the private sector is paid to undertake instead of the public sector.

PFI is a good example of the new market analytic at work. It shows how the information requirements changed from capital needs to the affordability of debt repayment at provider level and displace the role of area-based needs data in capital allocation. It also highlights difficulties in provider-level calculations of cost efficiency arising from the arbitrary or contentious nature of risk pricing and the way in which risk was shifted.

3. Commissioning and contracting with private providers:
the case of the ISTC
(i) risk segmentation through patient selection
The UK’s 4 billion pound independent sector treatment centre (ISTC) programme provides a third example of how a non comprehensive provider focus erodes area-based population data, in this case as a result of selection bias in the allocation of patients to and non-recording of data by the commercial sector.

In 2000, the UK government announced a plan to contract-out elective surgery under the ISTC programme (37). The opportunity to purchase care from the commercial sector under this scheme represented a major departure from the original model of the NHS as virtual monopoly provider of publicly-financed hospital services. In order to facilitate commercial participation two important risk management measures were adopted. In the first place, elective surgery was unbundled (or cases differentiated according to criteria of complexity such as age and co-morbidity) so that more complex and potentially higher cost cases could be left with NHS facilities and commercial providers could concentrate on high volume, low risk operations. The separation was achieved by making it mandatory to risk select patients through treatment protocols. Secondly, demand risk was retained by the state via the adoption of “take or pay” contracts according to which contractors were to be paid for a set number of operations whether or not this number of patients actually materialized.
(ii) loss of population data
With providers not bound by the same data collection duties as the NHS and standards of collection largely unmonitored or poor, the intrusion of risk selection was to have a profound effect on public health data. Mason and colleagues (38) at the University of York have shown that ISTCs do not collect good timely data, but that such data as are collected show that ISTCS recruit and select from a healthier patient population. This selection will of course affect hospital league tables and general performance data (see below). It also undermines routine data collection, which ceases to be comprehensive.

Evaluation of the ISTC programme was itself a victim of these new weaknesses for the data were not available to carry out a reliable assessment (37). Strikingly, evaluation in Scotland was carried out by the management consultants that had helped the government set up the programme.

This example shows that data collection is fundamentally affected when patient selection is allowed. In the first place, data integrity and continuity are undermined. Secondly, aggregate output or outcome data ceases to have meaning because it reflects risk selection not relative performance.

4. How international comparison of health system performance obscure risk selection and ignore differences in population – the case of Kaiser Permanente
It is possible to see risk selection play out in comparisons of health system performance both at provider level and internationally. In 2002 Richard Feachem and colleagues (39) published a paper purporting to compare the cost efficiency of the NHS with that of Kaiser Permanente, an American HMO. In fact it illustrates the way in which risk selection undermines health system comparison. HMOs are of course premised on patient and service selection, they combine insurer and provider functions and their populations comprise members or enrollees. The populations are unstable due to high enrollment and disenrollment rates, coverage is not universal nor does it serve contiguous areas and risk selection is rife.

Accordingly, aggregate performance data will reflect relative success in selection not relative success in attaining comprehensive cover within areas, which is the goal of the NHS.

Health system comparison has become an important tool in the political management of market reform. Frequently used to evaluate performance among systems at different stages of marketisation (40), international comparison are increasingly used to assess the relative effectiveness from a cost containment perspective of different structural elements. Comparisons of
this type are fraught with methodological difficulties because health care systems can be more or less selective. Comparisons between universal and non-universal systems is deeply problematical because like is not being compared with like.

Measurement and reporting issues can also confound results because of differences in the definition of health spending, under-reporting and over-reporting and variation in methods for measuring the size of the informal sector. Price comparison is also problematic. Prices can be compared by converting into US$ dollar equivalents using current exchange rates or by estimating ‘purchasing power parities’ (PPP), also referred to as the ‘international dollar rate’. PPPs involve contestable assumptions about market prices and the costs they reflect.

These problems were apparent in the paper by Feachem and colleagues published in the British Medical Journal in 2002 shortly after the NHS Plan announced the introduction of provider market (39). The paper’s authors purported to show that Kaiser Permanente was more cost efficient than the NHS and that risk selection and other problems of comparison had been taken into account. However, within a week of publication the BMJ had 170 responses mostly critical of data, methods and assumptions (41).

Crucially, the populations served by Kaiser and the NHS are very different and so too are the range of benefits. Whereas the NHS provides universal, comprehensive cover to all citizens in the UK, around 60 million people largely free of charge Kaiser recruits healthier, wealthier and younger patients as members to plans with restricted benefits and numerous additional charges and does not provide comprehensive care (41). No amount of risk adjustment can adjust for these differences; Kaiser’s population can never be risk adjusted to become like the NHS as it was never designed to be comprehensive.

This case study shows a studied attempt to promote as more efficient a system in which competing insurers and providers can optimize their risks by selection. However, the more an insurer and provider can risk select the more cost efficient it is likely to appear. In fact, the paper rested on erroneous methodology and false claims.

5. How hospital league tables engender risk selection
The switch to providers and members as part of risk selection strategies of market oriented systems is played out in performance measures locally. Performance league tables, widely promoted as a means of privatizing provision (10), also reflect a shift from universality to risk individualization by
substituting provider based performance for measures of access and equity at area level. They are products of economic theory that predicts markets will work imperfectly in circumstances where purchasers do not have full information. This problem, known among economists as “information asymmetry” (42), is particularly acute in health care where providers know far more than purchasers and are easily able to trade off cost against quality because the latter is so difficult to measure. League tables are intended to overcome this tendency by making comparisons of outcomes publicly available.

Hospital based mortality rates are often favoured as the measure of hospital quality; the reality is that area based mortality should be the focus. Jacobson and colleagues (20) and more recently Bottle and colleagues (21) have shown weaknesses in hospital mortality league tables. Among the problems identified are:

1. quality of data and coding and gaming
2. the numerator problem of deaths: patients may die on their way to hospital die because of delays in referral, or they may stay in hospital for different periods of time. Some may die out of hospital and others may be admitted but discharged and then readmitted to other hospitals before death and here the response is to apportion a death to all of the hospitals - the patient dies not once but several times over in several different hospitals!
3. denominator problems: the population is unstable and comprises those that can get access, but how are they counted as admissions, first admissions or FCEs and which population do you count?
4. case mix adjustment: different hospitals have different mixes of patients and services
5. institutional issues, small hospitals. poor quality and lack of data: coding incompleteness of data
6. inability to adjust for institutional differences

If they are to convey reliable information, league tables must at least adjust for differences in the risks present in different populations. However, risk adjustment methodologies are often spurious and involve unreliable data and complex modeling that conceals bias. In a landmark critique of hospital league tables in the Journal of the Royal Statistical Society, 1996, Goldstein and Spiegelhalter (43) conclude: ‘No amount of fancy statistical footwork will overcome basic inadequacies in either the appropriateness or the integrity of the data collected’. 
The journals are bursting with critiques of league tables but public authorities are impervious to them and use and publish them for marketing and recruitment purposes. A more profound problem arises when this type of data are used as the basis for pricing and cost comparisons and determination of whether an NHS hospital goes into deficit and whether it stays open or closes. More recently economists have tried to argue that hospital concentration and competition saves lives based on erroneous assumptions and data of this type (44). However such performance league tables can discourage clinicians and providers from treating patients with high morbidity and high costs.

And yet the geographic population focus of a public health frame is far simpler. For many decades public health has analysed variations in treatment, service use and access over time between and across populations, districts or PCTs and by social class. These data have been and used to conduct sensible detailed audits, confidential enquiries, surveys and investigations into the whole patient pathway including referrals from and access to primary care.

Conclusion
It has long been recognised that health care planning and equitable resource allocation cannot be left to the invisible hand of the market. As Abel-Smith put it in 1976, markets mean that there is “no single organization pledged to provide the best health service possible out of a limited budget (45, p. 154).” Indeed, the NHS was originally conceived as a non-market model in order to optimize resource use. However, our case studies show that a comprehensive service requires an administrative bureaucracy underpinned by information consistent with that function. Public bureaucracies are not destroyed overnight but rather through a succession of technical incremental changes to the information systems which inform the systems for resource allocation, capital allocation and coverage and provision. Where risk segmentation is the goal then there is an interplay between information and resource allocation systems which change the data requirements in fundamental ways so that the focus shifts from area-based populations to aggregates of individuals or enrollees for the purpose of risk management. This shift fundamentally affects the availability, production and analysis of area-based comprehensive data. Finally, once introduced risk-based data can be used to undermine universal systems by allowing providers and commissioners to be selective about coverage and care. This switch in data systems is therefore both a cause and a consequence of market fragmentation and enables the shift from national risk pooling to risk selection.
Central to all this is the unit of analysis. When the focus ceases to be comprehensive health care to the whole population living within an area and becomes instead risk allocation to individuals, members, enrollees or providers, then universality is no longer attainable because the necessary data underpinning it are no longer available and providers have the opportunity to risk select. As we show in our examination of the latest NHS legislative proposals, this change of focus is a way of managing the normative transition from universality to selection (2). That is why we have concluded that the information requirements which underpin risk selection contribute to the abolition of the NHS as a universal system of health care.

**Literature**


38. Mason A, Street A, Verzulli R. Private sector treatment centres are treating less complex patients than the NHS. JRSM 2010;103 :322-331


Allyson M. Pollock
Professor of public health research and policy at Queen Mary, University of London
allyson.pollock@gmail.com