Gwyn Bevan and Lawrence D. Brown
The political economy of rationing health care in England and the US: the ‘accidental logics’ of political settlements

Article (Published version) (Refereed)

Original citation:

© 2014 Cambridge University Press
This version available at: http://eprints.lse.ac.uk/57129/
Available in LSE Research Online: August 2014

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LSE Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain. You may freely distribute the URL (http://eprints.lse.ac.uk) of the LSE Research Online website.
The political economy of rationing health care in England and the US: the ‘accidental logics’ of political settlements

Gwyn Bevan and Lawrence D. Brown

Health Economics, Policy and Law / FirstView Article / April 2014, pp 1 - 22
DOI: 10.1017/S1744133114000127, Published online: 23 April 2014

Link to this article: http://journals.cambridge.org/abstract_S1744133114000127

How to cite this article:

Request Permissions : Click here
The political economy of rationing health care in England and the US: the ‘accidental logics’ of political settlements

GWYN BEVAN*
Department of Management, London School of Economics, UK
LAWRENCE D. BROWN
Department of Health Policy and Management, Mailman School of Public Health, Columbia University, USA

Abstract: This article considers how the ‘accidental logics’ of political settlements for the English National Health Service (NHS) and the Medicare and Medicaid programmes in the United States have resulted in different institutional arrangements and different implicit social contracts for rationing, which we define to be the denial of health care that is beneficial but is deemed to be too costly. This article argues that rationing is designed into the English NHS and designed out of US Medicare; and compares rationing for the elderly in the United States and in England for acute care, care at the end of life, and chronic care.

Submitted 1 September 2013; revised 28 February 2014; accepted 5 March 2014

Introduction

This article compares institutional arrangements for, and consequences of, the rationing of health care in England and the United States, which we define to be the denial of health care that is beneficial but is deemed to be too costly. In this introduction we outline the argument we develop in the rest of the paper. We first clarify complexities of nomenclature. There is no single UK National Health Service (NHS); and, following devolution in 1999, there has been increasing divergence of policies between the NHSs of England and Scotland, Wales and Northern Ireland (Connolly et al., 2011). In this article we refer to England where policies for its NHS differ from the other NHSs, and use the United Kingdom where the four NHSs have common features and where the United Kingdom has been used for international comparisons. Within the English NHS, as there have been regular structural reorganisations since 1991 (Klein, 2010), we use the term ‘purchaser’ to cover their successive organisational forms (district health authorities, primary care groups, two types of primary care trusts and clinical commissioning groups from 2013).

*Corresponding to: Gwyn Bevan, Department of Management, London School of Economics and Political Science, Houghton Street, WC2A 2AE London. Email: R.G.Bevan@lse.ac.uk
This article begins by considering the differences in the power of the executive in each country, which we argue shape legislative reforms of health care. In the United States, legislative process, the checks and balances on the power of the executive mean current vested interests are crucial actors in any political settlement that seeks to reform health care. The concessions deemed necessary to those interests to enact legislation mean that they continue to be key players after legislation has been enacted. In England, the government can largely ignore the vested interests that will cease to matter after legislative reform of health care is enacted, and so can focus on the political settlement with those who will be key players in a future system. These differences have been powerfully demonstrated by the comparative difficulty in the United States and ease in England of implementing two current controversial reforms. The Affordable Care Act (ACA) (Obamacare) was central to the policies of two successful Presidential election campaigns, but, despite the process being crafted to enable important concessions to be made in framing the legislation, there was a struggle for this legislation to be enacted, and after being enacted, it was subject to, in 2012, challenges in the Supreme Court, and, in 2013, to Republican threats to shut down the Federal Government if the Act were not repealed or delayed (Jost, 2014). The English policies of the Secretary of State for Health (2010) did not feature in the 2010 campaign of the Conservative Party, contradicted the commitment of the Coalition Government (Cabinet Office, 2010) to ‘stop the top-down reorganizations of the NHS that have got in the way of patient care’ (p. 24), but have been in a process of implementation since the 2010 election (Timmins, 2012). We outline two landmark political settlements that have shaped the institutional arrangements for health care in each country, the creation of the Medicare and Medicaid programmes in the United States and the NHS in England. Each has had what Tuohy (1999) describes as ‘accidental logics’, which she describes as follows: ‘episodes of policy change have required a confluence of factors in the broader political arena; the resultant systems have been shaped by the climate of ideas and the constellation of interests that exist at the time that such a confluence occurs. Once established, the institutional mix and structural balance of these systems intersect to generate a distinctive logic that governs the behaviour of participants and the ongoing dynamics of change’ (p. 7). Hence the different ‘accidental logics’ have path-dependent consequences for the different institutional arrangements for health care, which we argue imply different social contracts for rationing health care. We explain how, in England, explicit and implicit rationing is designed into the NHS; but, in the United States, rationing is deemed to be unacceptable within the systems of health care and hence mainly applies through user charges and in variations in, and lack of, insurance. We explain why the principal challenge of rationing in both nations is over decisions on the capacity of new technology to do more for the elderly. Hence we focus our comparisons on how different institutional arrangements ration the care of the elderly, in the United States and England for acute and end of life (EOL) care, where technology matters, and chronic care, where technology tends to be less important.
Political settlements and their institutional consequences

As Marmor and McKissick (2012) argue, the US Medicare programme was designed to offer financial protection from the risk of financial disaster due to the escalating costs of health care facing the elderly, and its direct correlation of rising elderly poverty rates, by transferring the responsibility for payment to the government (i.e. to social insurance contributors and to taxpayers, including younger Americans); ‘but to do so without interfering significantly with the traditional organisation of American medicine’ (p. 59, emphasis in original).

Partly to seek peace with organised medicine and partly to avoid impressions that Medicare was somehow ‘second tier’ coverage, the programme’s designers also adopted payment arrangements that largely mimicked those to which they were accustomed in private Blue Cross and Blue Shield insurance plans: payment based on usual and customary charges. (Hospitals received retrospective payment of their ‘actual’ costs for Medicare patients.) Its payment methods were quickly derided as a ‘blank check for inflation’ and so they proved to be. Tuohy (1999: 60) observes that ‘In the design of the Medicare program, American policy makers made even greater concessions to the interests of providers and private insurers than was the case in Canada’. As Brown (2011, 2012a), argues, these concessions may have been necessary to create the coalition required in the United States for the political settlements that extended coverage.

The outcome of legislative reforms of health care in the United States is that the institutional arrangements are those of a mixed private-public system with a rhetorical bias towards confining government to filling gaps in private coverage. The irony is that the government provides insurance for approximately one-third of Americans, including those over 65, veterans of the armed forces, active military servicemen and women as well as many Americans with disabilities, Native Americans and the poor through Medicaid. The private health insurance industry covers many employed Americans and their families, but historically, tens of millions have been uninsured. In 2011, the United States was one of the three countries where less than half the total costs of health care were publicly financed; the other two were Chile and Mexico (OECD, 2013a). US health care has been delivered by various organisational forms, but is generally characterised by paying hospitals negotiated prospective sums per clinical episode and physicians fees for services delivered, some consumer cost sharing and a mix of for-profit and not-for-profit hospitals, with patients often having direct access to specialists. In the United States, acute and EOL care are covered by the Medicare programme; but long-term chronic care is not (apart from treatment for renal failure, short-term skilled nursing services directly following a hospital stay, and home health care services for beneficiaries unable to care for themselves).

The objectives of the creation of the NHS were to achieve equity of access by need, which justified the principal elements of ‘socialised medicine’: universal coverage, access free at the point of delivery, finance by taxation, regulation of the
distribution of general practitioners (GPs), and the nationalisation of voluntary and local government hospitals. These principal elements of ‘socialised medicine’ still apply to each NHS in the United Kingdom. Following the creation of the NHS, private medical care became of marginal importance: it is typically a perk of employment and generally directed at elective care for those of working age. In 2011, the United Kingdom was one of the three countries where over 80% of the total costs of health care were publicly financed; the other two were Denmark and Sweden (OECD, 2013a). The power of the executive in the British Parliamentary system meant that despite the impacts of the radical legislative changes on insurers and hospitals in the creation of the NHS in the 1940s, they barely feature in accounts of the political settlements, which focus on hospital specialists and GPs (Webster, 1991; Klein, 2010). These settlements delivered effective systems of cost control: specialists are paid salaries; GPs’ contracts are based on payment by capitation; and access to specialists is via GPs who act as gatekeepers. This settlement also resulted in rationing being designed into the NHS though what Klein (2010) described as 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’ (p. 61). And the commitment to equity resulted in the development of policies that aim to distribute that budget geographically in relation to the estimated relative need of populations (Bevan, 2009). A paradoxical outcome has been that, despite some fears of the medical profession at the time of the creation of the NHS that it would result in them being subject to state control (see e.g. Webster, 1988), this arrangement has allowed them considerable clinical autonomy as compared with the mixed public/private system in the United States.

**Institutional arrangements and their implicit social contracts**

In England, the accidental logic of the political settlement that created the NHS was a set of institutional arrangements for rationing health care (including the three types of care we examine later) that create an implicit Universal Social Contract (USC), for which Daniels (1990, 2003) and Dworkin (1992, 2000) provide different philosophical justifications, which may be summarised as follows. Because we are all behind a ‘veil of ignorance’ about our future health and incomes, the best form of insurance against an uncertain future (from both severe illness and poverty) is collective. Indeed Arrow (1963: 947) argued that, although such a USC looks to be justified by a preference for redistribution, it ‘can be reinterpreted as a desire for insurance’. In this USC insurance is required (and hence constrains choice), independent of ability to pay (and not based on individual risk) and thus in effect is a system of taxation (Evans, 1987). Aneurin Bevan (1978: 100), the Minister who created the NHS emphasised the redistributive nature of its USC: ‘Society becomes more wholesome, more serene and spiritually healthier, if it knows that its citizens have at the back of their
consciousness the knowledge that not only themselves, but all their fellows have access when ill, to the best that medical skill can provide. Daniels and Dworkin argue that such a USC provides a justification for the principle of ‘lifetime rationing’ that, ceterus paribus, the young ought to have priority over the old; although other differences between the young and the old may mean that priority may be given to treat the young on grounds other than age, such as the higher risks from treatment for the old and better outcomes for the young. Dworkin (2003: 314) illustrates the principle of lifetime rationing through his thought experiment in an idealised state, in which there is equity of wealth and income, no information asymmetry on health care for patients, insurers are denied information on enrolees health, and health care decisions are left to ‘as free a market as we can imagine’ (p 312). Individuals have to decide, in this idealised state, what they are prepared to pay, over their lifetimes, for insurance cover for health care. Dworkin argues that: ‘Most young people on reflection would not think it prudent to buy insurance that could keep them alive by expensive medical intervention, for four or five months at the most if they had already lived into old age. They would think it wiser to spend what that insurance would cost on better health care earlier, or on education, or training or investment that would, provide greater benefit or more important security’ (p 314). Daniels emphasises that the principle of lifetime rationing avoids the utilitarian problem of interpersonal comparisons because it operates ‘within a life and not between lives’ (Daniels (1990: 96, emphasis in original) and thus justifies a fundamental element of cost-effectiveness analysis, which gives priority to services directed at the young simply because of their greater potential life years. Daniels also argues that the principle of lifetime rationing implies entitlement to a ‘fair innings’ so that resources are allocated to achieve greater equity in life expectancy, which Williams (1997) argued ought to take account of quality of life.

The accidental logic of the political settlement that created Medicare and Medicaid has striking consequential differences from those of the NHS. Medicare, as Marmor argued, breaches ‘one version of social solidarity' by separating the beneficiaries, who are entitled to coverage because they are ‘retired’ (on grounds of age, disability or renal failure), from those who are in employment, and are left ‘flailing in the sea of either uncertain coverage or added constraints on their choices within insurance coverage’ for themselves and their families (Marmor, 2012: 47). This follows from Medicare being based on an implicit Intergenerational Categorical Social Contract (ICSC), on the principle that its beneficiaries have earned their privileged entitlement to health care when ‘retired’ through the financial contributions they made when they were in work: as Marmor points out, there is ‘no evidence that any substantial number of Americans’ would want to end Medicare’s relatively generous coverage’ (Marmor, 2012: 47). No such entitlement applies to Medicaid financed by transfer payments from taxpayers as welfare for the eligible poor. These various insurance arrangements give philosophical justification neither for the principles of lifetime rationing, nor using
the criteria of cost-effectiveness, nor a fair innings; and have resulted, as Daniels (1990) observed, as lifetime rationing in reverse: policymakers have been ‘more willing to impose stricter eligibility requirements and budget ceilings in Medicaid, most of whose recipients are young women and children, than alter our practices for the dying elderly’. Indeed about one-third of Medicaid spending goes to long-term care, and the programme ‘spends more than five times as much on each senior in long-term care as it does on each poor child’ (Bernstein 2012: A1).

In practice Medicare is financed by complex arrangements from a Pay As You Go (PAYG) scheme and its trust fund. As the costs of Medicare increase with an ageing population, there is confusion over the implications of the ‘insolvency’ of its trust fund. The implications for the problem of insolvency of pensions financed by a PAYG scheme with an aging population are, in principle, quite straightforward: which are to extend the age of retirement (Barr and Diamond, 2006). As we explain below, however, the main driver of increases in costs of health care with an aging population are not caused simply by increases in life expectancy only, but because developments in technology mean that more can be done for those who live longer. Medicare’s ICSC is based on the promise that the generation who are retired are entitled to whatever new technology has been proven to be effective, regardless of cost. This promise is designed to escalate the costs of Medicare as developments in technology mean that expenditures on current beneficiaries will exceed their contributions when in work (which financed care for the previous generation of retirees with ‘old’ technology), and hence requires increasing financial contributions from each generation in work.

Institutional arrangements for determining expenditures on health care and the distribution of resources

We compare the outcomes of the different institutional arrangements for determining expenditures on health care in the United States and the United Kingdom/England and the distribution of resources to the NHS in England and the Medicare programme in the United States.

Figure 1 shows total costs of health care as a percentage of GDP in the United States and the United Kingdom over 50 years (1960–2010). In the United Kingdom, the outcome of the top-down process of cost control of the NHS has been that constrained growth in total costs of health care has been the norm. Indeed, the two ‘financial crises’ of England’s NHS in the past 30 years, in the winters of 1988–1989 and 1999–2000, were ones of ‘underfunding’ after two decades of limited growth (in the 1980s and 1990s) (Bevan and Robinson, 2006). In each case these crises were followed by increases in NHS funding and the decade of the 2000s was one of sustained real growth.

The disparate components of US health care largely generate costs from the bottom up with the outcome that the dominant pattern of the total costs of US health care is one of continued escalation. Various empirical studies have
investigated and rejected common perceived causes of why the United States is an outlier in international comparisons because of its exceptional high and increasing costs of health care. These include wasteful administrative costs, more defensive medicine, expensive care for the terminally ill, the spread of health insurance, and its ageing population (Newhouse, 1993; Reinhardt et al., 2004; Anderson et al., 2005; Smith et al., 2009). One reason for the United States being an outlier is the high prices paid for health care (Anderson et al., 2003). And, although studies have consistently shown that applying current costs by age group has limited effects on future costs (Barer et al., 1987; Zweifel et al. 1999), Smith et al. showed that a second explanation is the combination of aging and technology, which accounted for over 60% of the increase in costs in the United States from 1960 to 2007: i.e. developments in technology meant increasingly more has been done for ageing populations. The next three sections of this article that follow are our examination of how different institutional arrangements impact on the care of the elderly, in the United States and England for acute and EOL care, where technology matters, and chronic care, where technology is much less important. Before we do so, we make two comparisons in the way total spend on health care is determined and how it has been distributed.

In England the NHS budget is determined in relation to other pressures on the exchequer and since 1977, that budget has been distributed with the objective of securing ‘equal opportunity of access for those at equal risk’ (Bevan, 2009). Although that policy reduced the range of variations of actual spend per capita (adjusted for risk) from nearly twofold in the 1980s (Beech et al., 1990), to about 1.3-fold in the 2000s (Department of Health, 2007); this had been accompanied by widening inequalities in health outcomes. So, from 1999, Ministers added a second objective: to reduce ‘avoidable’ health inequalities, which is consistent with rationing on the equitable basis of a ‘fair innings’ (Bevan, 2009). These processes of resource allocation are intended to ensure that in different areas doctors’ rationing decisions are made within an equitable distribution of

![Figure 1. Total expenditure on health care as % GDP for the United States and the United Kingdom, 1960–2011. Source: OECD (2013b).](image-url)
NHS resources. Medicare itself enfranchised millions of previously uninsured older Americans, many of whom were poorer and minorities, emphasises an equality of benefits (and equal cost sharing) across its enrollees and has recently begun to focus on reducing inequalities in both health outcomes and access (Eichner and Vladeck, 2005). But, despite this commitment to equity within Medicare, wide variations in spend per capita have been consistently documented by the Dartmouth Atlas, after accounting for differences in cost allowances (Skinner, 2011), age, sex and race. Gottlieb et al. (2010) reported that, in 2006, adjusted total Medicare spend per capita varied threefold (ranging from about $15,000 in Miami to $5000 in Honolulu). The unintended opportunity costs of US arrangements include incomplete coverage and comparatively poor health outcomes: OECD (2013c) points out that in the United States, ‘Life expectancy is now more than a year below the OECD average of 80.1, compared to one year above the average in 1970’ (p. 1).

Acute care

Aaron and Schwartz (1984) compared acute care in the United States and ‘Britain’ (which they use to describe both England and the United Kingdom) in the 1980s, which was during a decade of limited growth for the NHS (see Figure 1). They found rationing in Britain not by reducing quality to increase numbers treated, but by doctors within the NHS using the principles of lifetime rationing and cost-effectiveness, from constraints on supply and by waiting lists. The evidence for lifetime rationing was that the United Kingdom had lower rates than the United States for haemodialysis, which largely affects the elderly (we discuss this further below), but not for haemophilia, which is usually discovered in childhood (p. 37). The evidence for rationing by cost-effectiveness (Williams, 1985) was that British rates of hip replacements and coronary artery surgery were 80% and 10% of US rates (p. 92). The evidence of rationing by constraints was for scanners in Britain, which they saw as causing ‘a substantial reduction in the quality of care’ (p. 72); but they suggested that use of diagnostic X-rays in the United States was excessive. The persistent problem of rationing by waiting (p. 14) has notoriously plagued the NHS since it began; its perceived absence in the United States is often used as an argument against ‘socialised medicine’.

Although the later comparative study by Aaron et al. (2005) in the 2000s (during the decade of sustained growth in NHS funding – see Figure 1), reported dramatic increases in rates of treatment in Britain for renal failure, hip replacements and coronary artery surgery, the British rates for renal failure and coronary artery surgery were still below those for the United States. The former reflected differences in what physicians believed to be appropriate care (pp. 43–44), and the latter large differences in supply: e.g., New Jersey, with about one-seventh of the UK’s population, had over 50% more cardiologists (p. 73). They interpreted the low British rates for coronary artery surgery as evidence of rationing, in
that patients who could benefit were denied treatment, but speculated that the US rates may have been too high (p. 76). They also found that the United Kingdom continued to have much lower rates of diagnostic services than the United States (in use of CT scanners and MRI equipment), which they saw as both eroding quality of care as compared with the United States and as a sound rationing response (p. 92). Figure 2 gives rates for the United Kingdom as a percentage of those for the United States for selective diagnostic and treatment activities for 2006 using the most recent comparative data (OECD, 2011), which do not report rates by age. These show the United Kingdom to have had much lower rates than the United States for treatment for renal failure, expensive diagnostic services, coronary artery surgery, and knee replacements; but rates of hip replacements were higher in the United Kingdom than in the United States.

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 over-riding 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 strong performance management (based on strategies of ‘naming and shaming’ and ‘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 long-waiting times (Bevan and Hood, 2006; Connolly et al., 2011).

Second, since 1999, explicit rationing of new technology has been the responsibility of the National Institute for Health and Clinical Excellence (NICE) (2012).
NICE does not approve for NHS coverage all drugs and devices that the UK’s medicines and healthcare products regulatory agency licenses, but recommends what should be available on the NHS in a deliberative process with reference to evidence from cost-effectiveness analysis. This was how the Labour Government intended to end geographical variations in access to care (‘postcode rationing’) (Klein, 2010).

NICE normally considered interventions costing less than £20k (~$34.4k at 2008 prices) per quality-adjusted life year (QALY) to be cost-effective; the least cost-effective intervention NICE approved cost £49k ($84k)/QALY (Steinbrook, 2008). Claxton et al. (2013) suggest that the threshold should be below £20k (at 2008 prices) otherwise the NHS by funding new technology could drive out current care that is more cost-effective. NICE provides a remarkable contrast with analogous institutions in the United States. As Marmor observed over 10 years ago, the ‘profound’ question whether Medicare ‘should finance all the medical care passing the test of efficacy’ (Marmor, 2000: 154) has remained unresolved. After decades of debate the United States has still not agreed to add cost to effectiveness in its formal decisions on the treatments Medicare covers. Neumann and Weinstein (2010) criticise the limits on the new Patient-Centered Outcomes Research Institute, created by the ACA, in conducting comparative-effectiveness research and specifically the prohibition from ‘developing or using cost-per-QALY thresholds’ on the grounds that this uses the principle of lifetime rationing, which ‘seems to reflect long-standing concerns. The worry is that the metric unfairly favours younger and healthier populations’ (p. 1496). They point out that this rejection of cost-effectiveness analysis leads into the funding of care, which Dworkin argued against, with unintended consequences:

Taken literally, it means that spending resources to extend by a month the life of a 100-year-old person who is in a vegetative state cannot be valued differently from spending resources to extend the life of a child by many healthy years. Though the ACA may be seeking to avert discrimination, it instead helps to perpetuate the current system of implicit rationing and hidden biases (p. 1496).

A recent joint report by Age UK and the Royal College of Surgeons (2012) pointed out that, until October 2012, NHS bodies were exempt from provisions of the Equality Act 2010 that outlawed discrimination on grounds of age; but now such discrimination is illegal and so decisions ought to be made on clinical grounds without reference to age. Furthermore this is now part of the NHS Constitution (Department of Health, 2012). The Report found evidence that rates of surgery consistently declined with age with indications of rationing by age for those over 85 in England, which had also been identified by Judge et al. (2010). This suggests that ending age discrimination would create considerable extra demand for services in England. The medical director of the NHS, Sir Bruce Keogh, in evidence to the Public Accounts Committee of the House of Commons reported being ‘deluged’ with letters of complaint over inconsistency in the
thresholds for surgery for which purchasers would pay for hip replacements and cataract operations in the winter of the first year of austerity for the English NHS (2010/11); and that the Royal College of Surgeons is working on value based commissioning of elective surgical care, to define appropriate access thresholds (UK House of Commons, Committee of Public Accounts, 2013, ev 25). The implications of the NHS constitution proscribing aged-based discrimination are to end the principles of ‘lifetime rationing’, a ‘fair innings’ and the use of QALYs as a measure of effectiveness to give priority to the young over the old; NICE still uses QALYs but recognises that this use of cost-effectiveness analysis is controversial (http://publications.nice.org.uk/guide-to-the-methods-of-technology-appraisal-2013-pmg9). We return to this point in our discussion.

Care at the EOL

Superficially, for EOL care, the first comparative impressions are of considerable common ground in the objectives of the two systems in England and the United States. For both systems the central problem of EOL care is defined as not when, but rather where and how one dies, with a corollary that better EOL care occurs not in hospitals but in homes or hospices, sites that are often both more comfortable for patients and families and less costly. Both systems increasingly endorse the development of alternative EOL settings (e.g. hospices), treatments (e.g. palliative care), and trajectories (e.g. evidence-based ‘pathways’). Both envisage the expansion of personnel trained to give a range of largely non-clinical services, including counselling and pain management. Both seek to make patients, families, and the public at large better acquainted with these less clinically aggressive options and their advantages. However, the very different institutional arrangements have resulted in different emphases in considering how to resource and organise EOL care.

The Department of Health grounds the case for new directions in EOL care primarily on the contention that less aggressive, intensive, and intrusive hospital care is better for, and preferred by, a great many patients. For example, between 56% and 74% of people surveyed prefer to get care and die in their own home, or a care (nursing) home instead of an acute hospital; but, in 2006, 58% of deaths took place in an acute hospital and only 35% in a home or care home. The overriding argument for reform is to meet individuals’ preferences, but doing so would also lead to impressive savings: e.g., about £4.5 million would be available annually for investment in community services if the 40% of patients who died in an acute hospital in Sheffield in October 2007 without ‘medical needs which required them to be admitted’ had died in a less intensive setting (House of Commons, Public Accounts Committee, 2009: 7, 9 and 10).

Since the late 1980s, for acute services in England, the emphasis has been on provider competition to replace a planned economy of health care (except from 1997 to 2002) (Bevan and Robinson, 2006; Klein, 2010). For EOL reform,
However, the emphasis has been on planning: that is publicly driven strategies that elaborate and incorporate patients and providers (both NHS clinicians and local social service agencies) into care plans, that aim to keep EOL patients at home, with hospital stays the exception. This line of policy gained prominence in 2008 when the NHS adopted its End of Life Care Strategy, festooned with ambitious objectives ensuring that ‘pain and suffering are kept to an absolute minimum’, that ‘caregivers are appropriately supported’, and that ‘services provide good value for the money to the taxpayer’ (Gray, 2011). In 2011, the NHS went further, issuing new evidence-based EOL guidance. With regards to EOL care, the arrival in office of the Conservative/Liberal Coalition in 2010 so far seems to have had little effect. Despite the NICE guidance, EOL care in England triggers controversy as shown by different purchasers being said to have followed different philosophies and protocols, raising again the spectre of ‘postcode rationing’. We mentioned above that aged-based discrimination is officially proscribed, but the aged may be short-changed nevertheless, because, as observed in 2000, ‘some services operate with upper age limits, which may not be openly publicised but which are real barriers nevertheless’ (King’s Fund, 2000: 4). Klein and Maybin wonder whether NICE’s willingness to attach special value to drugs for terminal cancer care in the last few months of life raises questions of consistency: ‘Should not the same extra value be attached to the last few months of life for older people in hospital?’ For that matter, are prevailing priority setting processes systemically biased to favour ‘interventions as distinct from focusing on the quality of care?’ (Klein and Maybin, 2012: ix, 43). The coordination of health care with social services, though essential to sound EOL planning, is itself tortuously difficult to plan, because the two sectors are organisationally and financially distinct. Nor does more humane and cost effective EOL care come free. In March 2012, e.g., Andrew Lansley, the then Secretary of State for Health, announced that the NHS would invest 1.8 million pounds (supplemented by 2.5 million pounds from the Marie Curie Cancer Care Organization) to support eight pilot sites at which to develop ‘a new patient funding system for all providers of palliative care’ (http://www.theyworkforyou.com/debates/?id=2012-03-27a.1317.6). It also remains uncertain whether plans launched amid major new infusions of cash into the NHS in the early and middle 2000s will survive the straitened circumstances of the English budget today and for the foreseeable future.

In the United States there appears to be a different emphasis on the continued scope for medical progress and its capacity to improve and prolong the lives of the aged, and a history of strong antagonism to any notion of lifetime rationing. There is continuing vigilance against cost-minded federal bureaucrats, portrayed by some politicians as perversely bent on concocting ‘death panels’ designed expressly to pull plugs on seniors who live too long and cost too much (Rutenberg and Calmes, 2009). In the United States, needless to say, there exists no analogue to England’s publicly directed strategic planning for changes in EOL policies or practices. Gray (2011: 2) notes that none of the recommendations in the Institute
of Medicine’s report of 1997, Approaching Death, was ‘directed at any specific entities – governmental or nongovernmental – that could be held accountable for implementing them’.

Reform in the US centres on expanding the range of (consumer) choice, on the consolidation of formal rights for patients and their families to exercise choices about EOL care, and on the cobbled together of funding sources to pay for palliative care and kindred services. These reforms – living wills, power of attorney provisions, assisted suicide (illegal in England but authorised in three US states and under debate in several more), growth of hospices, greater availability of palliative care, and so on – advance via legal authorities and program budgets at both the state and federal levels of government and draw encouragement from foundations (Robert Wood Johnson, for instance) (Brown, 2012b). Under their Mental Capacity Act of 2005 the English too honour advance statements and decisions as well as Lasting Powers of Attorney, but what distinctively legitimates these reforms in the United States is that they simply expand the rights of patients to fashion their own EOL fates, free of pressures or mandates by government and its bureaucratic legions. A US government-generated equivalent of the monthly Newsletter, issued by the National End of Life Care Programme of the National Health Service, which addressed in its issue of August 2012 such topics as ‘End of Life Diabetes Care’, would currently be unimaginable.

EOL reformers in the United States worry that even when conversations between patients and providers are framed by living wills and other legal fortifications for choice, the ingrained bias towards ‘more’ makes it hard to stop the clinical juggernaut, which often by definition is in a hospital setting, not a home-care based setting, where most Americans, like most English, say they prefer to die. (As Gawande (2010) notes, it is rarely the case that, clinically speaking, ‘nothing’ can be done at the EOL.) The empowering documents may not be filled out or filed properly, or may be ambiguous in the face of the patient’s complex and rapidly changing clinical conditions. The family may disagree among themselves about what the incapacitated patient would want done or forgone. Providers may fear legal repercussions if they cannot prove they did all they could to prolong the patient’s life. It may be unclear which specialist is in charge, with the authority to end aggressive treatment. In the United States and the United Kingdom both, EOL patients determined to see their rightful choices implemented may face a sharp uphill struggle.

These comparative EOL conundrums admit no evident endpoint. In terms of objectives, the two nations are gradually converging with aims to offer more options and choices for patients, more treatment alternatives and more efforts to accommodate patients and their families’ wishes of dying outside a high-tech hospital setting. However, institutional arrangements suggest that the EOL care that is provided puts England and the United States at opposite ends of the cross-national EOL continuum – a bias towards less is more in England and towards more is better in the United States.
Chronic care

One consequence of the original political bargains to pass Medicare in the United States was it was not designed to cover care for the chronically sick elderly (Marmor, 2012), although Medicare does now cover some elements of chronic care. Certainly, the chronic care challenge is significant and growing in both countries. In the United States, an estimated 65% of the Medicare population have at least two chronic conditions (Wolff et al., 2002) with more than 26% of the Medicare population suffering from diabetes alone (Fradkin, 2012). For the over 65s in England, over 60% report living with a chronic illness or disability (Office of National Statistics, 2011). In the United States, costs associated with chronic illnesses account for more than 75% of Medicare spend (Sochalski et al., 2009). The ACA depends on savings in Medicare, which accounted for 20% of total US health care spend in 2010 (Kaiser Family Foundation, 2011), to meet its health care budget targets. Similarly, in England, the government has touted restructuring chronic care as having the potential to improve care quality, constrain future overall costs by helping people stay healthier for longer and save on the government’s on-going costs through introducing greater cost sharing on individuals for care homes (‘care homes’ or ‘nursing homes’ are not under the purview of the NHS).

Whether or not reforms when translated into practice in either the United States or the United Kingdom will alter the ways in which chronic care is delivered and paid for remains to be seen. The NHS and Medicare currently both ration chronic care for the elderly, albeit generally in different ways. Below we offer three examples of such rationing, which illustrate all three possible outcomes: the first is where there is clear evidence of more severe rationing in England, the second where there is less rationing in England, and the third is where the principle of explicit rationing seems to be conspicuously similar but there are differences in practice, which reflect institutional arrangements.

The first example is that notorious difference historically between the NHS and the United States in the availability of costly treatment for renal failure. In the United States, Medicare began financing renal care failure treatment in the 1970s for people over and under 65; high rates of treatment resulted across all ages. In contrast, the English NHS has severely rationed treatment for renal failure; not surprisingly, recent UK rates of treatment per capita for the over 55s were 30% of the US rate. Specialists providing renal services in England, however, never were forced to turn anyone away: other doctors decided which patients suffering from renal failure ought to be referred to the specialists providing treatments for renal failure. In doing so they informally employed ideas of lifetime rationing and cost-effectiveness, well before NICE formally adopted similar approaches starting in 1999 (Aaron and Schwartz, 1984).

The second example is prescription drugs. In England, drugs prescribed by GPs are subject to a copayment (of nearly £8), but many people are exempt from this charge including anyone over 60 (5 years younger than the Medicare age
threshold) provided the drugs have been approved by NICE (NHS Choices, 2014). Despite public perceptions to the contrary, the English NHS must provide all medicines NICE has approved free-of-charge for those suffering from cancer, heart disease or other chronic conditions, or for people over 60. If a patient over 60 wants to take a drug not approved by NICE, he or she must pay an effective co-pay to cover the difference between the value NICE computed and the drug’s actual cost (Faden and Chalkidou, 2011). In 2011, NICE extended its approval for some drugs for the treatment Alzheimer’s disease, which had been restricted to those with moderate disease to mild disease (NICE, 2011). In the United States, Medicare Part D began providing drug coverage for Medicare beneficiaries in 2006 (before 2006, Medicare enrollees had to pay for all prescriptions outside the hospital setting out-of-pocket or through supplemental drug insurance). Any drug prescribed by the patient’s doctor is de facto covered, as there is no American NICE equivalent for pharmaceuticals. Unlike in the NHS, however, Medicare Part D plans include a deductible and a 25% co-payment up to an initial coverage limit, after which the Medicare enrollee pays the full cost of prescriptions until the full out-of-pocket threshold is reached. The ACA will eliminate that coverage gap, the so-called ‘donut hole’, by 2020. Until then, however, the United States clearly will continue rationing prescription drugs for its Medicare beneficiaries on the basis of ability to pay, which disproportionally impacts those suffering from chronic conditions. People 65 and older with two or more chronic conditions are about as likely to fill initial prescriptions in both the United States and England but older Americans vs their English counterparts are more likely to let prescriptions lapse because of cost concerns (Morgan and Kennedy, 2010).

The third example is obesity, which is one of the leading risk factors associated with multiple chronic diseases, including heart disease, diabetes and many cancers on both sides of the Atlantic. Medicare covers bariatric surgery for Medicare beneficiaries suffering from type 2 diabetes, hypertension, coronary artery disease or arthritis who are also morbidly obese, defined as having a body mass index (BMI) of at least 35. That is an example of a Medicare National Coverage Determination (NCD), an assessment by Medicare’s medical team that defines who is eligible for any new treatment: a positive NCD does not automatically follow an approval by the US Food and Drug Administration (FDA). Even though the NCD contains a public comment period, in which, by US law, agencies are required to solicit and consider the views and recommendations of the public, the NCD methodology is notably opaque, in contrast to NICE’s commitment to transparency at each step of its evaluative and decision processes. Subsequent to a positive NCD decision, Medicare assigns a price to each treatment, technology, procedure or drug, which it updates every year. Medicare no longer mandates that the bariatric surgeries it covers occur in certain high-volume hospital settings only. (Centers for Medicare and Medicaid Services, 2013). In England, NICE guidance says that anyone with a BMI of 40 or more can be considered for bariatric surgery, as can those who would also qualify under Medicare’s criteria: i.e., those with a
BMI of at least 35 with an obesity-related chronic condition. Both Medicare and the NHS have explicitly disqualified those with a BMI less than 35 from coverage, even though in 2011, the FDA approved bariatric surgery for people with a BMI between 30 and 35 based on recent scientific findings that people in that category experienced improved health and fewer diabetes-related complications post surgery (FDA, 2011).

Although the explicit rationing policies are the same for bariatric surgery in Medicare and the NHS, actual access varies. In Medicare, if someone meets the eligibility criteria for bariatric surgery, he or she will be able to get the surgery performed at an approved hospital (albeit often with a 2–3 months wait). As with all inpatient and outpatient services, however, bariatric surgery for Medicare enrollees carries a variety of co-payments, up to an out-of-pocket maximum threshold across all inpatient services. Even within Medicare, this is another example of rationing by ability to pay because of the cumulative impact of the requirement for co-payments on people with multiple chronic conditions. In the NHS in England, lack of resources has resulted in many purchasers and GPs denying patients what NICE deems to be cost-effective bariatric surgery on the grounds that it is too costly, with many restricting bariatric surgery only to those with a BMI of 50 or above, and even then often requiring waits of 6 months or more (Royal College of Surgeons, 2010). Owen-Smith et al. (2013) highlighted the scale of rationing: ‘less than 0.6% of those who are potentially eligible for bariatric surgery receive treatment on the NHS’. Sir Bruce Keogh is reported as saying that there was ‘probably quite a strong case for us changing the threshold at which we offer bariatric surgery’ to treat morbid obesity (UK House of Commons, Committee of Public Accounts, 2013, ev 25). Owen-Smith et al. report that the NHS atlas (NHS, 2010) identified 93-fold variation in rates of bariatric surgery in England across 150 purchasers in England (over the period 2007–2008 to 2008–2009 after standardising for age and sex); and which, as Owen-Smith et al. point out, cannot be explained by differences in the prevalence of simple obesity. They argue that purchasers will struggle to achieve fairness and equity of provision in the current period of austerity.

Discussion

This article has argued that the differing powers of the executive arms of government have shaped the political settlements of the landmark legislative reforms of health care in the United States and England, which have in turn created different institutional arrangements with their own ‘accidental logics’. The institutional arrangements in England are that the government decides the total budget for the NHS, how that is distributed and thus sets the context for the rationing of care; and, in the United States, the government’s role within health care is in the main limited to being an additional third party payer. The outcomes of these institutional arrangements are stark differences in terms of total costs of health
care, coverage, the geographical distribution of resources, and the ‘accidental logics’ of what types of rationing of health care are and are not acceptable: in England it is accepted that within the NHS some beneficial health care will not be provided because it is too costly and it is unacceptable that access would depend on ability to pay; and in the United States the converse applies. As the reason for the differences in total costs is likely to be from differences in the uptake of new developments in technology for the elderly, we have examined how different institutional arrangements impact on the care of the elderly, in the United States and England for acute and EOL care, where technology matters, and chronic care, where technology is much less important.

The two comparative studies of acute care in Britain and the United States in the 1980s (Aaron and Schwartz, 1984) and 2000s (Aaron et al., 2005) found stronger indications of lifetime rationing in the English NHS in the 1980s, a decade of severe resource constraints, than in the 2000s, a decade of increasing resources. But even after that substantial increase in NHS funding, the British rates for renal failure, coronary artery surgery, knee replacements, and use of CT scanners and MRI equipment were still below those for the United States. Recent studies suggest that, in the English NHS, the elderly are being denied beneficial surgery for hip and knee replacements, and cataracts. For EOL care, despite common objectives in both countries, the different institutional arrangements tend to mean more aggressive interventions in the United States and awareness of their opportunity costs in England. For chronic care: rates of treatment are higher in the United States than the United Kingdom for renal failure, which is funded by Medicare; the NHS has more generous coverage in providing prescription drugs without co-payment, although this ‘donut hole’ in the United States will eventually be filled by the ACA; and for bariatric surgery, rationing in England is by local purchasers and GPs as gatekeepers, which may or may not follow NICE guidance, and in the United States by individuals’ responses to arrangements for cost sharing. More generally, in the United States it is toxic even to broach a topic that could be interpreted as lifetime rationing of Medicare enrollees (as illustrated by the furor about ‘death panels’) with an unintended consequence of rationing care for many poor Americans.

We began by pointing out how the current reforms to health care demonstrate the comparative power of the executive in England to impose top-down changes on the NHS and the comparative vulnerability of the executive in the United States to current vested interests from insurers and providers. In England the latest attempt to make provider competition generate incentives for improved performance has been implemented in an era of severe cost constraints with questions over how rationing will apply to the joint cost pressures from an ageing population and developments in technology. The legislative and constitutional proscription of lifetime rationing within the NHS removes what Daniels and Dworkin argued provides a sound basis of responding to those pressures. In the absence of lifetime rationing, the alternative formal bases of rationing are: using a modified form of cost-effectiveness that
ignores differences in life expectancy, which raises formidable problems of interpersonal comparisons; or a lottery. In practice, a different kind of lottery is likely to develop which lacks the transparency of explicit deliberative approaches: by purchasers, with questions over ‘postcode rationing’; and implicitly by doctors, making the complex clinical judgements over the benefits of treatment for the old as compared with the young, similar to that of the 1980s, as described by Aaron and Schwartz (1984). For the United States, the problem and central question are different. The political settlement that delivered the ACA is one of expansion in coverage (though perhaps, after the Supreme Court’s ruling, less by means of Medicaid than was initially envisaged), with little real prospect of controlling escalation in costs. So, how will the system pay for the elderly and young alike, and in particular for new treatments for an ageing population? And will the elderly still have priority over those who are younger and poor?

Acknowledgements

The authors are grateful to contributions and comments from, and discussion with, Nicholas Barr, Oliver Bevan, Chelsea Clinton, Rudolf Klein and Ted Marmor. The usual disclaimer applies.

References


Secretary of State for Health (2010), Equity and Excellence: Liberating the NHS. Cm7881, London: Department of Health.


