Hospital Accounting and the History of Health-Care Rationing

Florian Gebreiter
Aston Business School
Aston University
Birmingham, B4 7ET
United Kingdom
f.gebreiter1@aston.ac.uk

Abstract: Focussing on the period from 1948 to 1997, this paper examines the history of rationing in the British National Health Service (NHS), with special reference to the role of hospital accounting in this context. The paper suggests that concerns regarding rationing first emerged in the 1960s and 1970s in response to the application of economic theories to the health services, and that rationing only became an issue of wider concern when the NHS increasingly came to resemble economic models of health services in the early 1990s. The paper moreover argues that, unlike in the US, hospital accounting did not play a significant role in allocating or withholding health resources in Britain. Rudimentary information systems as well as resistance from medical professionals are identified as significant factors in this context.

Keywords: accounting history, hospital accounting, health-care, rationing, NHS, NPM

Introduction

The cost of health care in many developed countries is high and rising. In Britain, for example, the proportion of GDP spent on the NHS has increased from 3.5 % in 1950 to 8.2 % in 2010 (Harker, 2012)\(^1\). Governments across the globe have adopted a range of health-service reforms during the last 30 years in an effort to contain the costs of health care. Accounting has played an important role in this context. In Britain, a succession of governments has instigated a sequence of hospital accounting reforms including ‘management budgets’, ‘resource management’, ‘reference costing’ and ‘payment by results’ since the 1980s (DHSS, 1983; DoH, 1989, 1997, 2002).
Despite such sustained attempts to make health services more efficient and economic, it is frequently suggested that the central ambition of the NHS, namely the provision of all beneficial health care free at the point of use, is not sustainable. From the 1990s onwards, a wide range of commentators have voiced concerns regarding the rationing of health care with many suggesting that it was ‘inevitable’ (e.g. Coulter and Ham, 2000; Harrison and Hunter, 1994; New, 1997).

Against the background of such concerns, the rationing of health care emerged as a significant subject of social science research over the last two decades (e.g. Coulter and Ham, 2000; Harrison and Hunter, 1994; Hunter, 1997; Klein, Day and Redmayne, 1996; New and Le Grand, 1996). Whilst these studies provided significant insights into the rationing of health care, two shortcomings of this literature can be identified. Firstly, the extant literature has a strong focus on the present and future of health-care rationing. Its history, alternatively, has attracted little attention from researchers. Secondly, despite suggestions that hospital accounting is centrally implicated in the rationing of health care in the USA (Fleck, 1987; Preston, Chua and Neu, 1997), the role of hospital accounting in health-care rationing in the context of more socialised health systems like the NHS remains largely unexplored.

The present study seeks to address both of these perceived shortcomings by examining the history of health-care rationing in the NHS with special reference to the role of hospital accounting in this context. The study focusses on the period between the creation of the NHS in 1948 and the election of a Labour government in 1997, which put a (temporary) halt to the market-based health-service reforms the departing Conservative administration had introduced from the 1980s onwards. The next section discusses extant literatures on health-care rationing and hospital accounting, as well as the methodology employed by this paper. The subsequent section examines the history of rationing in the NHS between 1948 and 1989. It argues that concerns regarding rationing first emerged in response to the application of economic theory to the health services but remained subdued as neither the behaviour of patients and doctors complied with the predictions of economic models. Hospital accounting only played a marginal role in rationing debates as well as in NHS resource allocation processes more generally. The next section suggests that wider concerns regarding the rationing of health care in Britain emerged against the
background of the Working for Patients reforms (DoH, 1989), which sought to recast the NHS in the image of economic theories of health care. It also suggests that the government’s ambition for accounting to play a central role in allocating health-care resources failed to materialise due to deficiencies in NHS information systems and unwillingness among medical professionals to engage with financial information. The final section summarises and concludes the paper.

**Literature Review and Methods**

Health-care rationing and hospital accounting have attracted considerable attention from academics over the last two decades. This section offers brief introductions to these literatures with an emphasis on historical studies, followed by an overview of the methods employed by the present study.

**Rationing Health Care**

The rationing of health care, which Schmidt (2004) defines as policies that cause patients to ‘forego medically beneficial treatment within a collectively financed (insurance or tax-based) system of health care provision’ (p. 970), emerged as an important theme in British health-policy discourses in the 1990s. In close succession to each other, a large number of books and articles were published on this subject (e.g. Coulter and Ham, 2000; Doyal, 1997; Harrison and Hunter, 1994; Hunter, 1997; Klein, Day and Redmayne, 1996; New and Le Grand, 1996; Weale, 1998). Written by economists, doctors and social policy experts, virtually all of these contributions suggested that the rationing of health care was ‘inevitable’. It was argued that due to factors like ageing populations and advances in medical technology, demand for health services would necessarily outstrip supply within a system in which care was provided free at the point of use (e.g. Coulter and Ham, 2000; Hunter, 1997; New and Le Grand, 1996).

Academic work in this area has been firmly focussed on the present and future of health-care rationing. Its past, alternatively, has attracted relatively little interest from researchers. A few of the contributions cited above have, however, offered brief accounts of the historical development of health-care rationing as part of wider
examinations of the topic (Hunter, 1997; Klein, Day and Redmayne, 1996; New and Le Grand, 1996). These sources suggested that the rationing of health services was not a recent phenomenon but one of the defining characteristics of the NHS throughout its history. They argued that cash-limited budgets and tight expenditure controls meant that the NHS was a ‘monument to institutionalised scarcity’ virtually from day one of its operation (Klein, Day and Redmayne, 1996, 37; New and Le Grand, 1997). Yet, despite such severe resource constraints, these studies noted that ‘there was little – if any – public discussion of rationing’ in the early decades on the NHS (Klein, Day and Redmayne, 1996, 40; Hunter, 1997). All three historical accounts examined this apparent contradiction between significant under-resourcing and the absence of concerns regarding rationing. They explained this absence by suggesting that medical decisions served as an implicit rationing mechanism (Hunter, 1997; Klein, Day and Redmayne, 1996; New and Le Grand, 1996). More specifically, it was argued that their expertise allowed doctors to create ‘the illusion that decisions about whether or not to treat a condition and how were the result of clinical considerations rather than resource constraints’ (Hunter, 1997, 38). The slow adoption of the expensive but potentially life-saving technology of renal dialysis, for example, was explained in terms of the benefits of concentrating medical expertise in a small number of specialist treatment centres (Klein, Day and Redmayne, 1996).

The detailed consideration of the absence of concerns regarding rationing offered by these studies was not matched by a similarly thorough examination of early discourses on rationing in Britain. Klein, Day and Redmayne (1996) briefly noted that rationing was first explicitly discussed by Powell (1966) and subsequently the subject of a book by Cooper (1975), before suggesting that neither contribution resonated with the British public. Hunter (1997) and New and Le Grand (1996) did not discuss the initial emergence of British discourses on rationing at all. As a result, whilst providing insights into why rationing was not perceived as a significant problem before the 1990s, these three historical accounts tell us little about how concerns regarding rationing first emerged in the UK.

*Hospital Accounting*
Since the 1980s, the hospital has been a significant site for accounting research. Much of this research has adopted case study approaches to examine the effects of ‘New Public Management’ (NPM) reforms on hospitals, doctors and their patients (e.g. Chua, 1995; Jones and Dewing, 1997; Kurunmaki, 2004; Llewellyn and Northcott, 2005; Lowe, 2000; Lowe and Doolin, 1999; Preston, Cooper and Coombs, 1992). Consistent with calls to extend historical studies of accounting beyond their traditional focus on business organisations (e.g. Miller, Hopper and Laughlin, 1991; Walker, 2008a, 2008b), hospital accounting also emerged as a subject of historical inquiry from the 1990s onwards (e.g. Bracci, Maran and Vagnoni, 2010; Gebreiter, 2015; Holden, Funnell and Oldroyd, 2009; Jackson, 2012; Jackson et al., 2013; Jones and Mellett, 2007; Preston, 1992; Robbins and Lapsley, 2008; Robson, 2003, 2007; Samuel, Dirsmith and McElroy, 2005). The resulting literature developed a strong focus on examining the relationship between hospital accounting and a wide range of factors in its social and institutional environment. For example, a series of articles by Robson (2003, 2006, 2007) investigated how ‘dominant individuals’ such as Henry Burdett and Joseph Edmund Stone, institutions like the King’s Fund, and wider ‘political and economic forces’ have shaped British hospital accounting between the late 19th century and the mid-1970s. Other studies have examined how notions of philanthropy and morality (Holden, Funnel and Oldroyd, 2009; Jackson, 2012), developments in economics and engineering (Samuel, Dirsmith and McElroy, 2005), varying levels of government interference (Scott, McKinnon and Harrison, 2003) and changes in social organising principles (Jones and Mellett, 2007) have affected the historical development of hospital accounting practices.

The resource constraints experienced by health services around the world also form an important element of the social and institutional context of hospital accounting practices. In particular, the rising cost of health services, which is often attributed to technological progress and ageing populations, has frequently been discussed by the extant literature in order to contextualise, explain, justify or promote hospital accounting reforms (e.g. Cardinaels and Soderstrom, 2013; Chapman, Kern and Laguecir, 2014; Hopwood, 1992; Jones and Mellett, 2007; Kurunmaki, Lapsley and Melia, 2006; Lapsley and Schofield, 2009). Despite the high profile that it has enjoyed in public-policy debates over the last few decades, rationing and its relationship with hospital accounting have received less attention from accounting
researchers, historical or contemporary. Only two studies, both of which focussed on the role of DRGs in rationing health care in the USA, have addressed this issue. Fleck (1987), a philosopher, suggested that the introduction of DRGs in 1983 provided the US government with an ‘invisible’ mechanism for rationing health care. He argued that DRGs allowed the government to localise and privatise politically sensitive decisions regarding the allocation of life-prolonging resources. Preston, Chua and Neu (1997) similarly suggested that the introduction of DRGs effaced the political nature of rationing decisions. They argued that the DRG system acted as a ‘black box’ in the sense that it provided an ‘inscrutable’ rationing mechanism that was ‘little understood by most actors and generally believed to be correct’ (p. 159). The role of accounting in rationing health care beyond the United States however remains largely unexplored.

This paper pursues two objectives. Firstly, it aims to re-examine the history of rationing in the NHS. Unlike previous histories on this subject, which have focussed on explaining why rationing was not perceived as a problem in the early NHS, this paper examines how concerns regarding rationing first emerged in Britain during the 1960s and 1970s, and how they proliferated in the 1990s. Secondly, the paper seeks to examine the potential role of accounting in rationing health care in the British NHS and thereby extend our understanding of this issue beyond the specific context of the USA, where previous studies on this topic were located.

In pursuing these two objectives, the paper draws on a range of documentary materials including books, reports and government publications as well as professional journals such as the British Medical Journal, The Hospital, The Accountant and The Lancet. These professional journals were particularly suited for the purposes of this paper as they provided a continuous record of debates regarding health-care funding, rationing and accounting throughout the period investigated. Whilst this record was necessarily partial, it reflected a wide variety of perspectives including those of accountants, health economists, hospital administrators and medical professionals. Contributions which concerned themselves with health-care rationing, resource allocation, hospital accounting and related topics were selected from the sources listed above and analysed.
Resources, Resource Allocation and Rationing in the NHS 1948 - 1989

This section examines the history of resource allocation and rationing in the NHS between its creation in 1948 and the publication of the Working for Patients White Paper (DoH, 1989). The next sub-section focusses on the history of resource allocation generally, whilst the subsequent sub-section focusses specifically on the small number of explicit references to health-care rationing during this period.

Resources and Resource Allocation

On 5 July 1948, the vast majority of British hospitals entered into public ownership to form the NHS. From this day onwards, all British citizens would have access to health care on the basis of need rather than their ability to pay. Despite making medical care available free at the point of use to the entire population, the then government and other commentators were not concerned by the cost implications of nationalising the health service. Following the reasoning outlined in the Beveridge Report, it was widely believed that providing free access to health care would cure existing disease and prevent future disease (Beveridge, 1942). The nationalised service would make people not only healthier but also more productive. The cost of the service to the Treasury, initially estimated at approximately £150m for the financial year 1948/49, was thought to remain constant or even decrease in the future.

It soon became apparent that the government had underestimated the costs of the newly created health service. By 1949-1950, the cost of the NHS had increased to £359m. The higher than anticipated cost of the service caused great concern amongst politicians, doctors and accountants. Controlling the cost of health care became the most pressing health policy issue of the early 1950s. In 1950-1951 the government introduced a number of measures which it hoped would have an immediate effect on health-care costs. These measures included the imposition of a cash limit on the NHS budget and the introduction of charges for dentures, spectacles and prescriptions. In addition to these rather crude measures, a wide range of commentators called for the adoption of cost accounting in the hospital service (BMJ, 1952; The Accountant, 1952; The Lancet, 1952). According to these commentators, a departmental hospital costing system modelled on those used in industry would provide for efficiency,
economy and cost control in the nationalised hospital service. In response to these suggestions, the Ministry of Health (MoH) commissioned four reports on hospital costing in the early 1950s (King’s Fund, 1952; MoH, 1955; Nuffield Trust, 1952; RHBT, 1952), which paved the way for the nationwide introduction of a departmental costing system in April 1957.

Concerns regarding the cost of health care remained significant until 1956, when the Committee of Enquiry into the Cost of the NHS, which the government had commissioned in 1953, published its report (Guillebaud, 1956). This report suggested that the steep increases in health expenditure during the early years of the NHS did not reflect any fundamental flaws in the economic viability of the service but had been caused by general inflation and a back-log of demand for dentures and spectacles. The report concluded that the NHS had made efficient use of the resources dedicated to it and recommended that capital expenditure in particular ought to be increased rather than decreased. The findings of the Guillebaud Report put an (at least temporary) end to concerns regarding the cost of health care and both main political parties promised to increase health expenditure at subsequent general elections (Klein, 2006). The departmental costing system was introduced into the NHS in 1957, a year after the publication of the Guillebaud Report. Partly because concerns regarding the cost of health care had subsided by then, and partly due to the perceived limitations of the system, departmental costing did not go on to play a central role in the management of the NHS in general, or in resource allocation decisions in particular (c.f. Bourn and Ezzamel, 1986; Robson, 2003). Departmental costing was criticised for being over-elaborate and unresponsive to the needs of administrators (Hunt, 1961; Langley, 1961). It was moreover criticised for its strong focus on administrative efficiency and its inability to account for the cost implications of clinical decisions, which were increasingly highlighted as the ultimate drivers of hospital costs from the mid-1960s onwards (e.g. Feldstein, 1967; MoH, 1965). The control of expenditure in the NHS was exercised by means of fixed macro budgets, whilst local resources were allocated on the basis of clinical decisions, which were beyond the scope of departmental costing and whose autonomy ‘remained sacrosanct’ (Robson, 2003, 117; Gebreiter, 2015).
From the 1960s to the 1980s, NHS expenditure experienced several periods of high and low growth depending on the general economic situation and government spending priorities. The general trend in health expenditure was, however, upward as the cost of the NHS to the Treasury increased from 3.1% of GDP in 1960 to 5% in 1980 (Harker, 2012). Concerns regarding the allocation of funds within the service emerged at various points during this period, most notably in the 1970s. New approaches towards resource allocation were considered at both the macro- and the micro-levels. Hospital accounting would come to play virtually no role at the macro-level and a small but increasing role at the micro-level.

With regard to resource allocation at the macro-level, it must be noted that the share of the NHS budget allocated to England’s 14 Regional Health Authorities had initially been based on the regions’ hospital expenditure prior to nationalisation. As a result, affluent areas in the south of the country were allocated a relatively higher share of NHS resources than were poorer northern areas. Inflation-linked budget increases perpetuated this discrepancy during the 1950s and 1960s. In 1976, the government accepted proposals by the Resource Allocation Working Party which outlined a new allocation mechanism based on standardised mortality ratios (DHSS, 1976). This mechanism was thought to ensure that NHS funds would be allocated according to regional health-care needs rather than historical coincidence. The Working Party made no reference to a potential role for hospital accounting in this context. Others noted that emerging accounting technologies like ‘specialty costing’ could have been used to adjust regional health budgets for patients who were referred for treatment at hospitals pertaining to other health authorities (e.g. Magee and Osmolski, 1978). This approach was, however, not widely adopted.

Hospital accounting was to play a more prominent role in the context of resource allocation at the micro-level. Starting with the publication of the Cogwheel Report (MoH, 1967), there was an increasing emphasis on the resource implications of individual clinical decisions in the NHS. Suggestions emerged that doctors ought to become more conscious of the costs associated with their choices. Consistent with these suggestions, a number of experiments were conducted with costing systems whose aim was to provide clinically relevant cost information to doctors during the 1970s (e.g. Babson, 1973; Coles, Davison and Wickings, 1976; Russell, 1974). Coles,
Davison and Wicking’s (1976) contribution attracted particular attention, as it suggested that the introduction of clinical budgets at Westminster Hospital had resulted in significant clinician involvement and cost savings. Expectations were raised that accounting would come to play a central role in the micro-allocation of health resources. However, Coles, Davison and Wicking’s (1976) findings could not be replicated at other hospitals and the production of clinically-relevant cost information remained restricted to a small number of experimental sites during the 1970s (DHSS, 1978; Perrin, 1978).

As low economic growth, high inflation and growing social unrest culminated in the ‘winter of discontent’ of 1978-1979, Britain elected a Conservative government which, inspired by neo-liberal ideas, envisioned that accounting would play a central role in micro-level allocation of NHS resources. The government’s vision of a reformed health service was articulated by the NHS Management Inquiry (DHSS, 1983), which was conducted by Roy Griffiths, the deputy chairman of a supermarket chain. A central recommendation of the Griffiths Report was the creation of management budgets, a form of clinical budgeting. Its aim was to highlight to clinicians the cost implications of their decisions and to engage them with resource management. Management budgets were trialled at four test sites across England but received only a ‘lukewarm’ reception by clinicians, who criticised the cost information at the heart of management budgeting as ‘crude’ and ‘unacceptable’ (Stewart, 1984, 731-732). In 1986, the government had to abandon the four test sites and the label ‘management budgets’ as initial scepticism amongst local clinicians had turned into outright hostility (DHSS, 1986). With small modifications, and re-branded as the ‘resource management initiative’, management budgeting experiments resumed at other locations (e.g. Perrin, 1988; Preston, Cooper and Coombs, 1992). Alongside medical audit, the empowerment of patients and the separation of purchasers and providers of health care, the government envisioned the resource management initiative to form one of the central pillars of the post-Working for Patients NHS (DoH, 1989).

**Rationing**
Having briefly discussed the history of resource allocation in the NHS between 1948 and 1989 in general terms in the previous sub-section, the present sub-section focuses specifically on sources which discussed health-care rationing in explicit terms during this period. As noted by previous histories (Hunter, 1997; Klein, Day and Redmayne, 1996), these were small in number. The very first author to address this topic was Enoch Powell, a Conservative politician and former Minister of Health, whose convictions placed him on the right wing of his political party. At a time when economists first started to apply their theories to the health services (Feldstein, 1967; Klarman, 1965), and against the background of initiatives to strengthen the role of quantitative expertise in government (e.g. Fulton, 1968), Powell (1966) published his reflections on the present and future states of the health services. Powell’s (1966) contribution included a chapter entitled ‘Supply and Demand’, in which he mobilised these two concepts borrowed from economics to argue that health care was being rationed in the NHS. With regard to the latter of the two concepts, Powell (1966) described as ‘absurd’ the ‘vulgar assumption […] that there is a definable amount of medical care “needed”, and that if that “need” was met, no more would be demanded’ (p. 26). Instead, he argued that there is ‘virtually no limit on the amount of medical care an individual is capable of absorbing’ (Powell, 1966, 26-27). This ‘infinity of demand’, Powell (1966) argued, was not mitigated by pricing mechanisms because ‘medical care under the National Health Service is rendered free to the consumer at the point of consumption’ (p. 26). As a result, Powell (1966) concluded, ‘supply and demand are not kept in balance’ and ‘supply has to be rationed by means other than price’ (p. 26).

Cooper (1975), the second writer to explicitly discuss this topic, similarly borrowed the notions of supply and demand to suggest that health care was being rationed in the NHS. In a chapter entitled ‘The need to ration’ (p. 46), Cooper (1975), a trained economist, elaborated Powell’s (1966) suggestions that a combination of almost infinite potential demand for health care and the provision of health services free at the point of use would inevitably result in rationing. Cooper’s (1975) specific argument follows:

*Rational economic behaviour dictates that an individual faced with a zero price (free) commodity or service will consume it until further consumption yields him or her no*
further utility – that is until he has no further use for it. In the case of health this is likely to approach infinity or at any rate so high a level of consumption that is may as well be infinity. (pp. 46-50)

The supply of health services, alternatively, was ‘constrained by the size of the current national tax receipts and the health service’s share of them’ (Cooper, 1975, 50). Since, in a service free of charge at the point of use, this mismatch between supply and demand could not be resolved by means of price mechanisms, Cooper (1975) echoed Powell (1966) by concluding that ‘in the NHS rationing must take place in some other manner’ (p. 50).

As demonstrated by the above paragraphs, the first suggestions that health services were rationed in the NHS were derived from economic theory. According to the teachings of economics, there was virtually no limit to the amount of health services that ‘health-care consumers’ would demand in the absence of price constraints. No health service, however well resourced or organised, would be able fulfil such an infinite demand for health care. As a result, the rationing of health care was not only deemed commonplace in the NHS but also inevitable. The real NHS, of course, was populated by patients rather than ‘rational consumers [who] go on demanding health care until its “marginal utility” approaches zero’ (Cooper, 1975, 25). Even Cooper (1975) recognised that British patients were extremely reluctant to adopt the role of health-care customers:

"Few people think of themselves as potential customers of the [health] system. When they are customers they are too ill, and when they have recovered they are too relieved to and thankful to complain. […] People are unaware of what is technically possible. They are unaware of the gross inequalities in provision and, for the most part, have gratefully accepted whatever they have found. (p. 93)"

More generally, British patients ‘demanded’ very little of the NHS. They displayed ‘passivity in the face of long waiting lists’ (Frankel, 1992, 13) and a general attitude towards the health service which Klein (1984) characterised as ‘deferential’. A number of commentators suggested that such behaviours reflected the origins of British hospitals as charitable or Poor Law institutions (Frankel, 1992; Powell, 1966). Aaron and Schwartz’s (1984) observation that the NHS, which covered approximately
50 million people, only received a total of 2809 malpractice claims during the first 25 years of its existence provides a further indication that the early NHS served passive and deferential patients rather than rational health-care consumers.¹

In the real NHS, demand for health services was not determined by consumer preferences but by clinical decisions. During the first few decades of the NHS, such decisions were perceived as an art, which depended above all on the skill, intuition and experience of the individual clinician (e.g. Armstrong, 1977; Gebreiter, 2015). They were subjective, local and inseparable from the clinician taking them. Patients generally accepted clinical decisions unquestioningly, partly because of the great prestige and authority doctors commanded, and partly because the doctrine of ‘clinical freedom’ was seen to guarantee that doctors put the interests and well-being of their patients above all other considerations (e.g. Fox, 1951). The ultimate decision on who to treat, and who not to treat, was therefore the exclusive preserve of individual medical professionals.

It is this ability to provide or withhold treatment on the basis of subjective clinical judgements which later sources argued was at the heart of a system of implicit rationing during the early decades of the NHS (e.g. Hunter, 1997; New and Le Grand, 1996). Sources from the 1970s and 1980s suggest that the extent to which medical practice was affected by resource constraints and the degree to which doctors were complicit in such implicit rationing mechanisms is less clear-cut. Cooper (1975), for example, noted that doctors showed little awareness of their supposed role in rationing health care. A study by two American researchers (Aaron and Schwartz, 1984), which suggested that health care was being rationed in Britain because utilisation rates of various treatments and diagnostic tools were significantly lower than those in the USA, also found conflicting evidence regarding the role of British doctors in rationing. Based on interviews with a wide range of doctors, they suggested that whilst some felt that resource considerations impinged on their practice, the majority indicated that their decisions to withhold treatment were made exclusively on medical grounds. Aaron and Schwartz (1984) did not take the latter suggestions at face value. They argued that ‘British doctors seem to seek medical justifications for decisions forced on them by resource limits’ and that, almost subconsciously, these
doctors ‘gradually redefine standards of care so that they can escape the constant recognition that financial limits compel them to do less than their best’ (p. 101).

This assessment was only partially shared by Klein (1984), a British health policy researcher, who reviewed Aaron and Schwartz’s book for the *British Medical Journal*. Klein (1984) criticised that the book appeared to operate under the assumption that the USA provided the ‘optimum level’ of health care, hinting that differences in treatment rates between the UK and USA could be a function of overtreatment in the USA rather than rationing in the UK.\(^5\) He moreover suggested that decisions to withhold treatment in Britain were as much reflective of clinical preferences as they were of resource constraints:

[D]ifferences in the medical cultures of Britain and America are at least as important as differences in the availability of resources. The two are, to an extent, linked. A humane, clinical conservatism in Britain both sustains and is, in turn, reinforced by constraints in resources. A heroic, aggressive style of medicine in the United States helps to explain – and, in turn, to compound – the high rate of spending. (Klein, 1984, 143-144)

The evidence presented in the above paragraphs provides some support for suggestions that doctors were complicit in an implicit system of rationing. It however also raises the possibility that the relatively low levels of health expenditure in the UK (as compared to, for example, the USA) reflected different incentives and a more conservative clinical culture. Where the balance lay between those two arguments cannot be determined with any degree of certitude on the basis of the available evidence. What can be said with more confidence, however, is that there was very little public debate on rationing in the NHS during the first four decades of its existence (cf. Klein, Day and Redmayne, 1996; Hunter, 1997). Neither doctors, nor patients, nor the wider British public perceived rationing to be a significant problem. At a time when the health services were characterised by powerful, confident and autonomous doctors as well as trusting, passive and deferential patients, concerns regarding health-care rationing were largely restricted to the minds of a small number of economists and American commentators.
No reference was made to hospital accounting in early contributions on the subject of rationing in Britain. More generally, accounting played virtually no role in allocating funds within the NHS at the macro-level (DHSS, 1976), and only a tentative one at the micro-level (e.g. clinical budgeting). In a centralised, medically-dominated health service, clinical decisions and statistical tools like standardised mortality ratios were perceived as the appropriate basis for allocating resources at the micro- and macro-levels respectively.

**The Emergence of Wider Concerns Regarding Health-Care Rationing 1989-1997**

This section discusses three significant developments which occurred in parallel in the wake of the ‘Working for Patients’ White Paper (DoH, 1989); namely the creation of an ‘internal market’, the erosion of professional power, and the emancipation of patients. It argues that as a result of these developments, the NHS increasingly came to resemble economic models of the health service and concerns regarding the rationing of health care proliferated. The section moreover discusses the role of accounting in this context, arguing that the government’s vision of accounting’s playing a central role in allocating (and withholding) health resources did not materialise due to professional resistance and an under-developed information infrastructure.

**The Internal Market**

The central component of the health-service reforms outlined by Working for Patients (DoH, 1989) was the creation of a quasi-market for hospital services by separating the ‘purchasers’ of health care (i.e. Health Authorities) from its ‘providers’ (i.e. hospitals). Under this system, which came to be known as the ‘internal market’, local Health Authorities were to ‘contract’ out health services to a range of competing hospitals, including independent hospital trusts and private sector providers. These contracts would ‘spell out clearly what was required of each hospital in terms of price, quality and nature of the service to be provided’ (DoH, 1989, 24).
The government envisioned that hospital costing would play a two-fold role under the internal market. Firstly, it continued to emphasise clinical budgets and the associated aim of engaging doctors in resource management as key tools in increasing the efficiency of hospitals (DoH, 1989; Health Services Management; 1990; Coe-Legg, 1990). In consequence, the White Paper (DoH, 1989) pledged to expand the resource management initiative and by November 1991 all large acute hospitals in England had started to implement it (Scott, 1991). However, the costing systems which supported resource management were basic and, crucially, did not account for the quality of clinical care which had traditionally been the principal concern of clinicians (e.g. Dearsden, 1990; Green and Harrison, 1989). In consequence, many doctors felt that the NHS costing systems could not adequately support their decision making and, with few exceptions (e.g. Moseley and Fairbanks, 1992), they remained reluctant to engage with cost information in the 1990s (e.g. Buckland, 1994; Jones and Dewing, 1997).

Secondly, the government envisioned that costing information would underpin the newly created ‘market’ for hospital services (DoH, 1989). Based on accurate cost information, hospitals would bid for contracts with Health Authorities, which would then select the most competitive provider. However, such ambitions were not compatible with the rudimentary information and costing systems operated by the NHS in the late 1980s (e.g. Perrin, 1988). In order to improve the costing information supporting the internal market, the government set up a National Casemix Office, encouraged the creation of Healthcare Resource Groups (Sanderson, 1992) and published a succession of ‘Costing for contracting manuals’ (e.g. NHS Executive, 1993a, 1994a, 1994b). Despite these efforts, the cost information available to most British hospitals remained relatively basic and unable to account for the quality of medical care (NHS Executive, 1996) by the time the incoming Labour government abolished the internal market in 1997. A study published in the same year showed that not one of more than 100 NHS purchaser organisations surveyed was using the cost of Healthcare Resource Groups as a basis for contracting (NCO, 1997). Rather than relying on carefully-priced diagnostic groups, the purchasing activities of Health Authorities continued to rely on block contracts which covered a large and often fairly ill-defined set of services.
Although the contracting process largely failed to account for the precise price, nature and quality of the services provided, it did place some pressure on local Health Authorities to be more transparent in the allocation of resources. As a result, during the early 1990s, a number of authorities started to publish lists of health services that they were unwilling to fund (BMA, 1995; Klein and Redmayne, 1992). Whilst these lists consisted of marginal services like tattoo removal and the treatment of bat ears, they were a factor in bringing the issue of rationing to the attention of the wider public (cf. Klein, Day and Redmayne, 1996; Hunter, 1997).

**Doctors**

The decades that followed the creation of the NHS coincided with the ‘golden age’ of the medical profession (Abbott, 1988; Starr, 1982). Buoyed by medical breakthroughs ranging from the discovery of antibiotics to the performance of heart transplants, doctors were almost universally respected, trusted and admired. However, from the mid-1970s onwards, clinicians started to come under attack from a variety of angles. Marxists denounced them as agents of capitalist élites; nurses and other health service employees grew more confident and vocal; epidemiologists proposed that the increases in life expectancy during the twentieth century arose largely from public health and welfare measures rather than clinical interventions; and philosophers argued that the ‘medicalisation’ of society had done more harm than good (Illich, 1975, Klein, 2006; McKeown, 1976; Preston, Cooper and Coombs, 1992).

In addition to these critiques, changes within the medical profession also started to challenge the power of doctors. In particular, the emergence of new forms of medical knowledge was of relevance in this context. As suggested earlier, clinical knowledge had historically been perceived to be subjective and experience-based, and therefore inseparable from individual doctors (Armstrong, 1977; Gebreiter, 2015). From the 1970s onwards, there was increasing interest in statistical approaches to clinical practice, including Bayesian models for diagnostic decision-making processes and randomised controlled trials to assess the effectiveness of a wide range of clinical interventions (e.g. Bunker, Barnes and Mosteller, 1977; Cochrane, 1972). The emergence of such explicit and universal notions of clinical decision making implied that medical knowledge could become accessible to non-clinical professionals.
The diminishing status of the medical profession was both reflected in and reinforced by government health policy in the 1980s and 1990s. Two policies in particular represented a challenge to professional power and autonomy. Firstly, *Working for Patients* called for the introduction of the medical audit, defined as ‘a systematic critical analysis of the quality of medical care, including the procedures used for diagnosis and treatment, the use of resources, and the resulting outcomes for the patient’ (DoH, 1989, 39). In the reformed NHS, every hospital doctor, including fully-qualified consultants, would have their practice audited (Bowden and Walshe, 1991; DoH, 1989; Smith, 1990). Secondly, the government encouraged the development of clinical guidelines, documents which set out in considerable detail how specific diseases ought to be treated (e.g. NHS Executive, 1993b). Such guidelines, the government hoped, would inform clinical practice and form the basis of resource-allocation decisions (NHS Executive, 1996). Implicit in these two policies was, of course, the suggestion that much of the clinical work performed across Britain was sub-standard and in need of monitoring and/or improvement.

The decrease in trust, prestige and autonomy of the medical profession outlined in the above paragraphs is argued to be a significant factor in the emergence of concerns regarding health-care rationing during the 1990s. In this environment, it became increasingly difficult for doctors to determine the demand for health care by means of implicit decisions under the mantle of clinical autonomy. The wishes and choices of patients would play a much greater role in this context, as suggested by the next subsection.

*Patients*

The notion of the patient underwent significant changes in the 1980s and 1990s. Against the background of an increasingly consumerist society, the passive and deferential patient of the post-war decades would be replaced by a new, informed, vocal and assertive type of health-care consumer. The Conservative government was sympathetic towards this development and the White Paper (DoH, 1989) included a number of policy initiatives aimed at supporting patient choice.
As suggested by Greener’s (2009) history of consumerism in the British health services, patients had historically very little choice in the NHS. In theory, patients had enjoyed the right to move to another general practitioner if they were dissatisfied, but such action was neither encouraged nor common. This tendency was to change in the post-White Paper NHS. Under the internal market, ‘money will flow to where the patients are going’ and hospitals would be incentivised to ‘satisfy the needs and preferences of patients’ (DoH, 1989, 33). Patients would be encouraged to consider a wide range of options, ranging from ‘a real choice between GPs’ (p. 55) to purchasing optional extras such as single rooms or personal telephones when in hospital.

In addition to patient choice, the government highlighted that patients needed to be in a good position to make their choices. Consistent with this, the White Paper (DoH, 1989) outlined the government’s ambitions to make a wide range of information about the health services on offer available to patients, including ‘indicators of hospital performance which cover the quality as well as the efficiency of the services provided’ (p. 36). Once in hospital, patients would also get ‘clear and sensitive information of what is happening’ and advice on ‘clinical matters, such as the nature of an illness and its proposed treatment’ (p. 7).

The government’s efforts to empower the users of health services resulted in the publication of the ‘Patient’s Charter’ (DoH, 1991), which set out a range of rights and ‘encouraged individual patients to be more assertive in their use of services’ (Ham, 1993, 2). Of course, many of the ambitions regarding patient choice and information outlined in Working for Patients and the Patient’s Charter remained just that. To this day, the NHS is struggling to develop practical performance indicators and many patients still expect doctors to make choices on their behalf (e.g. DoH, 2012). However, a growing number of patients started to assert their choices and wishes in the 1990s. A refusal to comply with those by doctors and health authorities was increasingly denounced as rationing. The tragic case of ‘Child B’ came to symbolise both the increased confidence and assertiveness of patients and the rising concerns regarding the rationing of health care in the 1990s. Child B was a 10 year old girl suffering from acute myeloid leukaemia. After an unsuccessful bone marrow transplant, the girl was refused further ‘aggressive’ treatment by NHS doctors on account of a low likelihood of success and the considerable pain and discomfort she
would suffer in the process. The girl’s father refused to accept the doctors’ judgement and took the NHS to court, which, against the background of considerable media coverage, ruled in support of the medical opinion. As a result of the publicity created by the case, a wealthy donor agreed to fund private treatment of Child B, which however proved unsuccessful (Hunter, 1997; New, 1997).

As professional power declined, patients emancipated themselves and a quasi-market was created in which funds would follow the choices of health-care consumers, the NHS started to resemble the models of health services created by early health economists (e.g. Cooper, 1975). Doctors, under pressure from a range of constituencies including politicians, health economists and nurses, no longer had the undisputed power and authority to regulate the demand for health services. In the post-White Paper NHS, demand was increasingly determined by a new type of patient who was encouraged by various government policies to express her wishes and preferences (DoH, 1989, 1991). These new patients not only wanted high quality health care, but also wanted it promptly, on their terms and at their convenience. They would be the customers of the NHS, much like they were customers of any other business, with the crucial difference that in a health service free at the point of use, no price constraints would apply. As the resources available to the NHS were unlikely to increase significantly under the Conservative government, many observers believed that the health service would not be able to accommodate the new demands placed on it by health-care consumers. In consequence, many came to see the rationing of health care as inevitable (e.g. Coulter and Ham, 2000; Harrison and Hunter, 1994; New, 1997).

Similarly to the USA, where DRGs operated at the heart of health-care rationing (Fleck, 1987; Preston, Chua and Neu, 1997), the UK government had envisioned accounting playing a central role in allocating, as well as withholding, resources in a post-Working for Patients NHS (DoH, 1989). Such ambitions were frustrated by the information infrastructure and the reluctance of many doctors to engage with accounting information in the NHS, both of which are argued to reflect the historically nationalised and non-commercial nature of the service. Unlike in the USA, whose fee-for-service tradition meant that hospitals had elaborate information and billing systems in place (Preston, 1992), the NHS possessed only modest abilities
to record and price hospital activities in the late 1980s. Despite considerable investment in this area during the 1990s (e.g. NHS Executive, 1993a), the costing systems of most NHS hospitals remained unable to support the contracting process, as originally envisioned, by the time the internal market was abolished (NCO, 1997). The technical shortcomings of NHS costing systems can partly explain why doctors by and large failed to engage with cost accounting. The cost information produced was often not accurate or relevant enough in order to support medical decision making. In particular, the inability of hospital costing systems to account for the quality of medical care was frequently emphasised as a key deficiency in this context (e.g. Dearsden, 1990; NHS Executive, 1996). However, the historical nature of the NHS is also argued to constitute an important reason why British doctors developed little enthusiasm for accounting. In the USA, doctors had always been entrepreneurs as well as clinicians, whose success depended as much on business skills as on medical acumen (Starr, 1982). In the NHS, hospital doctors were salaried employees. Emphasis was placed, above all, on clinicians’ medical performance. As a result, NHS doctors, with only a small number of exceptions (e.g. Mosley and Fairbanks, 1992), neither had the skills nor the interest to engage with financial information (Buckland, 1994; Jones and Dewing, 1997).

Conclusions

This paper has examined the history of health-care rationing in the British NHS with special reference to the role of hospital accounting in this context. With regard to the history of rationing, it has traced the initial emergence of concerns regarding health-care rationing to the application of economic theories to the health services (Cooper, 1975; Powell, 1966). These sources suggested that in a service which was free at the point of use, patients would demand infinite amounts of health care. This study has argued that, during this period, the NHS did not conform to such theories as powerful doctors rather than patients determined the demand for health care. Health-care rationing only emerged as an issue of wider social concern in the wake of the Working for Patients White Paper (DoH, 1989), which aimed to weaken the power of the medical profession, emancipate health-care consumers and introduce market forces into the British health-care system. In an NHS which increasingly came to resemble economists’ models of health services, empowered patients rather than
medical professionals would determine the demand for health care. It was against this background that health-care rationing came to be seen as inevitable in the UK.

The findings of this study suggest that the emergence of concerns regarding health-care rationing was not, as often claimed, an inevitable consequence of ageing populations and technological progress. Instead, it would appear that such concerns depend on very specific notions of patients, doctors and their interactions which only gained wider traction in Britain in response to the neo-liberal health service reforms set out in the Working for Patients White Paper (DoH, 1989).

Unlike in the USA, where DRGs were at the heart of resourcing and rationing health care (Fleck, 1987; Preston, Chua and New, 1997), in the NHS, the role of accounting in this respect remained limited between 1948 and 1997. British politicians, administrators and health policy experts have periodically looked towards accounting as a mechanism for more effective resource allocation in the NHS, but neither the departmental costing system (MoH, 1955), nor various clinical budgeting experiments (e.g. DHSS, 1983) or costing for contracting initiatives (e.g. NHS Executive, 1993a) have fulfilled these aspirations. This research has identified an under-developed information infrastructure as well as apathy, if not outright hostility, among doctors as significant factors in this context and linked both of them to the historically nationalised and non-commercial nature of the NHS.
References:


Beveridge, W. 1942. *Social insurance and allied services.* London: HMSO.


Langley, C. 1961. Finance officers’ weekend school. The Hospital, November.


RHBT (Regional Hospital Board Treasurers). 1952. *Hospital Cost Accounting*. London: Committee of Regional Hospital Board Treasurers.


1 Whilst, generally, health expenditure as a proportion of GDP has tended to increase during this period, the increase has been far from linear. By 1960, health expenditure had fallen to 3.1% of GDP, from where it increased to 5% by 1980. Under the Thatcher government, health expenditure declined to 4.7% of GDP in 1990. The 1990s saw a small increase to 5.5%, whilst Labour’s investment in the NHS during the 2000s took health expenditure to just over 8% of GDP by 2010 (Harker, 2012).
2 Diagnosis Related Groups are a system which classifies hospital cases into one of 467 groups. These groups, which are homogeneous in medical and resource consumption terms, formed the basis of the prospective reimbursement system adopted by the US government to fund the Medicare programme from 1983 onwards.
3 *The Hospital* changed its name to *Hospital and Health Services Review* in 1972 and *Health Services Management* in 1988.
4 By comparison, in 2012-2013 alone more than 16000 medical negligence claims were lodged against the NHS (Donnelly, 2013).
5 Some of the British doctors interviewed by Aaron and Schwartz made similar points. One suggested that it ‘seems very seldom that the U.S. physician ever states that there is no surgery that would help, no drug that is advantageous, and no further investigation that is required. There seems to be an irresistible urge always to do something, even though in many cases the doctor concerned must realise that there is no possibility of benefit.’ (cited in Aaron and Schwartz, 1984: 66)
6 Healthcare Resource Groups are the UK version of DRGs.