

## Guest Editorial

# Rationing in the fiscal ice age

RUDOLF KLEIN\*

*Emeritus Professor of Social Policy, University of Bath, Visiting Professor London School of Economics and London School of Hygiene, UK*

One of the few predictions about the National Health Service (NHS) in the approaching fiscal ice age that can be made with reasonable certainty is that the long-simmering debate about rationing will boil up more fiercely than ever before. Internationally the NHS's ability to manage scarce resources has frequently been held up either as a shining exemplar or as a dire warning, not least in the political struggles over healthcare reform in the United States. But, of course, it is neither. The British experience can best be seen as a case study of the dilemmas and difficulties involved in managing the allocation of scarce resources. And it is precisely because the dilemmas and difficulties remain unresolved – and perhaps can never yield a final resolution – that they have prompted public and academic arguments over the decades: arguments that will flare up ever more fiercely as resource constraints bite in coming years. What follows, therefore, is a selective discussion of issues that are likely to feature in the debate.

First, though, some semantic ground clearing. Even the term 'rationing' is not without ambiguity, often used synonymously with 'priority setting'. The two are linked but best kept conceptually distinct. Priority setting describes decisions about the allocation of resources between the competing claims of different services, different patient groups or different elements of care. Rationing, in turn, describes the effect of those decisions on individual patients, that is, the extent to which patients receive less than the best possible treatment as a result. Refusing to prescribe snake oil is not rationing.

Next, it is important to distinguish between different forms of rationing (Klein *et al.*, 1996). The most visible, and politically sensitive, forms are denial and delay: a refusal to prescribe a particular drug and making patients queue on the waiting list. But the most pervasive form of rationing is the least explicit and least visible: rationing by dilution. Decisions at the coal face of healthcare delivery – not to order an expensive diagnostic test, say, not to make a specialist referral or to reduce ward staffing levels in order to balance the budget – normally attract

\*Correspondence to: Rudolf Klein, 12A Laurier Road, London NW1 1SG, UK.  
Email: rudolfklein30@aol.com

little attention unless they explode into a scandal: the Mid-Staffordshire case, where ‘too few doctors and nurses’ (Healthcare Commission, 2009: 7) resulted in patient deaths, is a case in point. But such decisions are as much a form of rationing as the refusal to prescribe a drug. In both instances the result is the same: patients receive less than the best possible treatment (see, e.g. Schubert *et al.*, 2008).

Hence, decisions about priorities and any consequent rationing take place at all levels of the NHS. In contrast to earlier decades (Department of Health and Social Security, 1976), government no longer makes explicit decisions about the allocation of resources to different sectors of the NHS. Setting output targets has replaced setting input priorities. However, this clearly affects the distribution of resources. So while rationing by delay has been dramatically reduced – and disappeared from the headlines – as a result of government target-setting and spending, the investment required to shrink waiting lists may well have aggravated other, less visible forms of rationing. Next in the hierarchy of decision making come Primary Care Trusts (PCTs), responsible for 80% or more of NHS spending, varying greatly in their approach to setting priorities or deciding what treatments should be available to whom. Lastly, some hundreds of providers and thousands of general practitioners – whether budget holders or not – are daily taking decisions which affect what patients get. And while we know a great deal about what is happening at the national level, and have some fragmentary insights into what is happening at the PCT level, we know very little about what is happening at the coal face: the known unknowns multiply as we descend the decision-making ladder.

### **A technocratic chastity belt?**

Centre stage at the national level is the National Institute for Health and Clinical Excellence (NICE). Its responsibility for providing guidance to the NHS has many dimensions but, crucially for this analysis, it represents an attempt to depoliticize decisions about who should get what: in effect, NICE was intended to be a technocratic chastity belt to protect the NHS against ministerial meddling. Science and evidence, not political whims or media panics, would shape NICE’s decisions and NHS practice: clinical and cost effectiveness would be the criteria when appraising technology and issuing guidance. Such at any rate was the theory when NICE was set up. In the event, however, there has been continuing controversy around NICE’s role and methodology.

In general terms, there are the doubts about the appropriateness of using cost utility analysis in determining what should or should not be on offer in the NHS and the consequent reliance of NICE on quality-adjusted life years (QALYs) in its technology assessments (for a critical discussion of the issues involved, see Oliver and Sorenson, 2005). Crucially, should the exclusive aim of policy be to maximize the production of ‘health’ – as measured in terms of the gain in

QALYs offered by different interventions – as distinct from, say, widening access or reducing inequalities? Next, granting the centrality of QALYs in NICE's methodology, how should the outer limits of acceptable spending be set? NICE's upper threshold is normally set at £30,000 per QALY. But, as NICE recognizes, this is an arbitrary figure. In the words of its chairman, 'there is no empirical basis for deciding at what value a threshold should be set'. Consequently, there are 'social value judgements' which are incorporated into NICE's decision-making process (Rawlins and Culyer, 2004). Decisions are not driven by arithmetic alone.

Given the arbitrariness of the QALY threshold – however modified by social value judgements – NICE is vulnerable on two counts. The first is that NICE's decisions about new technologies – mainly expensive drugs – ignore the opportunity costs to the NHS. NICE's now mandatory recommendations may displace more cost-effective forms of spending (House of Commons Health Committee, 2007). For example, the implicit cost per QALY of spending by PCTs on cancer services appears to be about £22,000 (Martin *et al.*, 2008), that is, well below NICE's threshold. It is an issue that is likely to become ever more important in an era of frozen budgets. All the more so, since NICE attracts the wider criticism that the overall impact of its activities has been to drive up NHS spending as a result of concentrating on new technologies – read drugs – while paying insufficient attention to assessing and weeding out existing activities (Maynard *et al.*, 2004). NICE has responded by publishing estimates of savings produced by some of its recommendations (NICE, 2009) and its web site now offers 'cost saving support'.

The second source of NICE's vulnerability stems from the fact that, as noted, social value judgements are inescapably – and rightly – involved in making its decisions. This immediately raises the question of what legitimacy NICE has for making such judgements (Syrett, 2003). Here the outcome of the 2008 debate as to whether patients should be allowed to make top-up payments for expensive end-of-life cancer drugs not approved or yet appraised by NICE without losing their entitlement to continuing NHS treatment is instructive. The debate was largely driven by patient groups – funded, at least in part, by pharmaceutical companies – and featured prominently in the media. The Secretary of State for Health set up an inquiry (Richards, 2008) which recommended that top-up payments should be permitted, provided that privately funded activity was clearly separated from NHS activity. Further, the Richards report argued that there should be 'greater flexibility from all those involved, including NICE, in determining how the benefits that some drugs deliver are valued', citing specifically drugs for small groups of patients in the last months or years of life. NICE duly showed greater flexibility, introducing new criteria for drugs for small patient groups with short life expectancies (Raftery, 2009).

The episode can be interpreted in two very different ways. On the one hand, it can be seen as a defeat for the notion that rationing decisions in the NHS can be

insulated from politics in the widest sense. Even if NICE's decision to revise its criteria was not taken under direct government pressure, it was almost certainly taken in order to head off such pressure. Further, putting a premium on the last weeks or months of life is clearly inefficient, threatening to displace more cost-effective activities (Maynard and Bloor, 2009). From this perspective, the episode represents a regrettable surrender to lobby power and media pressure.

Contrariwise, the episode can be seen as justification of NICE's rejection of an exclusive focus on efficiency, while at the same time, raising questions as to how social values should be defined and by whom. It is an issue with which NICE itself has struggled. It engages in a variety of consultations. It has developed guidelines on social value judgements for its various committees. Most interestingly, perhaps, it has developed a form of deliberative democracy by setting up a Citizens' Council of 30, which periodically meets to consider specific issues. Its members are provided with relevant evidence and time to consider competing arguments; the results of their deliberations are both thoughtful and critical (see, e.g. NICE, 2008). Whether this is sufficient to give NICE decisions legitimacy is an open question, further considered in the conclusions. For some of the same issues apply to the priority and rationing decisions taken by Primary Care Trusts, to which we now turn.

### **Framgedted decision making**

The Department of Health's handbook setting out the requirements that Primary Care Trusts have to satisfy if they are to achieve 'world class commissioning' status includes instructions about how PCTs should determine strategic priorities (Department of Health, 2009, Competency 6). They have to demonstrate, among other things, that they use a mature programme budgeting technology, that the PCT board works with clinicians, local general practitioners (GPs) and the public to develop investment or disinvestment proposals and that projects and initiatives are evaluated against 'defined prioritization criteria'. In short, it is all about process. Nothing is said about how 'prioritization criteria' should be defined, weighted, applied and used. Much the same is true of guidance issued by the National Prescribing Centre (2009) on behalf of the Department of Health. Absent a national template or a national survey of how PCTs set about the task – NICE only commissioned such a study at the beginning of 2010 – we therefore have to rely on a series of snapshots.

As always in the NHS, variation is the norm. But it is variation around a theme: how best to take decisions about priorities given multiple, not necessarily congruent and possibly conflicting, criteria. 'It is now recognized that prioritization based on a single criterion such as cost effectiveness, quality adjusted life years, equity analysis or clinical effectiveness often leads to poor resource allocation decisions', in the words of one PCT policy document (Okoro, 2008). Accordingly this PCT, like many others, uses multi-criteria decision analysis

(MCDA). This involves scoring candidates for investment on 11 criteria – ranging from the magnitude of the benefit to impact on health inequality – using a five-point scale (see also Airoidi and Bevan, 2010). Other techniques are on offer and in use, notably Programme Budget and Marginal Analysis (PBMA), designed to create an economic framework within which account can be taken of equity and other criteria (Donaldson *et al.*, 2008; Wilson *et al.*, 2008). Different methodologies and different local contexts inevitably lead to different decisions being taken. Further, national, NICE-approved priorities may be ranked lower than local priorities (Iqbal *et al.*, 2006; Wilson *et al.*, 2007). Thus, postcode prioritization – and the consequent rationing – seems to be built into the system.

The same is true of rationing decisions in the strictest headline sense. These are the decisions that PCTs make in response to exceptional funding requests, that is, the requests of clinicians who want to prescribe a drug not normally funded by the NHS (and which may have been turned down by NICE or not yet approved) for a specific patient. Here a survey conducted for the Richards report (2008) provides some national data. Eighty PCTs (roughly half) responded to the survey, and reported that some 14,000 applications had been made in the previous year. The number of applications per Trust ranged from 0 to 1017%; the approval rate again ranged from 0 to 100% (see also Russell, 2009).

Variations, like deviations from national policy, are no surprise in the NHS. Even so, the range in the number of applications stands out. For this brings us to the bottom rung of the rationing ladder: decisions by clinicians. The range no doubt reflects a number of factors, but high among them is the variation among clinicians in their willingness to accept, or their readiness to challenge, top-down rationing decisions. One of the first, most innovative if also flawed studies of rationing in the NHS (Aaron and Schwartz, 1984) argued that clinicians internalize resource constraints in their practice. But what we do not know is the extent to which this results in the denial of potentially beneficial treatment – rationing in the strict sense – as distinct from a sensible reluctance to engage in heroic medicine regardless of the cost. In any case, treatments may be refused on grounds of humanity rather than cost (Ham and Pickard, 1998). Rationing at the coal face of the NHS – by clinicians and managers – is the black hole of NHS studies, no doubt in part because of the difficulty of knowing when economy and efficiency measures turn into, or should be defined as, rationing by dilution. This continues to be both a conceptual and a methodological puzzle.

In summary, the practice of rationing in the NHS remains an example of ‘muddling through elegantly’ (Hunter, 1993), though some would question the appropriateness of the adverb. Not only is there inconsistency in the criteria and methods used to make judgements. But much rationing is diffuse and implicit, with a low public profile. Over the years, there have been changes: notably, of course, the introduction of NICE. But the hope that a combination of evidence-based medicine and economic analysis could provide a ‘rational’ basis for allocating scarce resources, insulated from political processes, has not materialized.

Evidence-based medicine does not necessarily tell clinicians how to treat individual patients, so making the case for allowing medical discretion in the use of resources (Mechanic, 1997). Moreover, the evidence often changes over time (Groopman, 2010). Economic analysis has to compete with non-utilitarian criteria for deciding how to allocate resources, as is evident from the experience of NICE and the use of techniques like PBMA and MCDA which seek to accommodate other perspectives.

For the future, efforts to devise more sophisticated methods of analysis, incorporating multiple criteria, will no doubt continue. However, their sophistication – indeed their ‘rationality’ – is likely to be judged in terms of their ability to recognize the limits of any purely technical approach. For there is another – complementary, if not necessarily competing – paradigm for defining rationality. It is one which puts the emphasis on the *processes* of decision making and, in doing so, acknowledges the role of politics in the widest sense, that is, responding to public sentiments and preferences. This argues that the legitimacy of rationing decisions depends on accountability for reasonableness (Daniels, 2000) and the institutions of public involvement (Landwehr, 2009). It is a debate that has a long history, with political scientists and economists tending to take opposing view in line with their disciplinary biases (Klein and Williams, 2000).

Devising appropriate processes is, however, as challenging as designing robust analytic techniques. In the case of the NHS the government appears to have taken the view that promoting public consultation is the answer. Thus, a statutory duty to involve the public has been placed on PCTs. However, there is little evidence of any public appetite for involvement (Lomas, 1997). Indeed a recent UK study points in the opposite direction (Chisholm *et al.*, 2009). Most members of the public think that decisions about priorities and rationing should be taken by doctors; inconsistent public views reflect the consultation methods used. Conceivably, NICE’s Citizens’ Council may provide a better model. However, in the hard times ahead no generally accepted decision-making model is likely to emerge, far less any model that insulates central government from the consequences of financial stringency. The best that Ministers can hope for is that most rationing will continue to take the form of dilution rather than excision and that decisions will be taken in the name of clinical discretion and thus be politically invisible.

## References

- Aaron, H. J. and W. B. Schwartz (1984), *The Painful Prescription*, Washington, DC: The Brookings Institution.
- Airoldi, M. and G. Bevan (2010), *Improvement in Practice: Commissioning with the Community*, London: The Health Foundation.
- Chisholm, A., K. Briggs and J. Askham (2009), *NOT NICE: Can PCTs Engage Patients and the Public in Commissioning New Health Technologies?*, Oxford: Picker Institute.

- Daniels, N. (2000), 'Accountability for reasonableness', *British Medical Journal*, **321**: 1300–1301.
- Department of Health (2009), *World Class Assurance Handbook – Year 2*, London: Department of Health, <http://www.dh.gov.uk>
- Department of Health and Social Security (1976), *Priorities for Health and Personal Social Services in England*, London: HMSO.
- Donaldson, C., A. Bate, P. Brambleby and H. Waldner (2008), 'Moving forward on rationing: an economic view', *British Medical Journal*, **337**: 905–906.
- Groopman, J. (2010), 'Health care: who knows best', *The New York Review of Books*, **LVII(2)**: 12–15.
- Ham, C. and S. Pickard (1998), *Tragic Choices in Health Care: The case of Child B*, London: King's Fund.
- Healthcare Commission (2009), *Investigation into Mid Staffordshire NHS Foundation Trust*, London: Healthcare Commission.
- House of Commons Health Committee (2007), *National Institute for Health and Clinical Excellence First Report of Session 2007–2008*, London: The Stationery Office HC27.
- Hunter, D. (1993), *Rationing Dilemmas in Health Care*, Birmingham: National Association of Health Authorities and Trusts.
- Iqbal, Z., A. Pryce and M. Aza (2006), 'Rationalizing rationing in health care: experience of two primary care trusts', *Journal of Public Health*, **28**: 125–132.
- Klein, R., P. Day and S. Redmayne (1996), *Managing Scarcity: Priority Setting and Rationing in the National Health Service*, Buckingham: Open University Press.
- Klein, R. and A. Williams (2000), 'Setting priorities: what is holding us back – inadequate information or inadequate institutions?', chapter in A. Coulter and C. Ham (eds), *The Global Challenge of Health Care Rationing*, Buckingham: Open University Press.
- Landwehr, C. (2009), 'Deciding how to decide: the case of health care rationing', *Public Administration*, **87(3)**: 586–603.
- Lomas, J. (1997), 'Reluctant rationers: public input to health care priorities', *Journal of Health Services Research and Policy*, **2(2)**: 103–111.
- Martin, S., N. Rice and P. C. Smith (2008), *The Link Between Health Care Spending and Health Outcomes for the New English Primary Care Trusts*, York: Centre for Health Economics CHE Research Paper 42.
- Maynard, A. and K. Bloor (2009), 'NICE wobbles', *Journal of the Royal Society of Medicine*, **102**: 212–213.
- Maynard, A., K. Bloor and N. Freemantle (2004), 'Challenges for the National Institute for Clinical Excellence', *British Medical Journal*, **329**: 227–229.
- Mechanic, D. (1997), 'Muddling through elegantly: finding the proper balance in rationing', *Health Affairs*, **16(5)**: 83–92.
- National Institute for Health and Clinical Excellence (2008), *Report on NICE Citizens Council Meeting: Quality Adjusted Life Years and the Severity of Illness*, London: National Institute for Health and Clinical Excellence, 31 January–2 February 2008. <http://www.nice.org.uk> [20 January 2010].
- National Institute for Health and Clinical Excellence (2009), 'Cost saving guidance', London, <http://www.nice.org.uk> [20 January 2010].
- National Prescribing Centre (2009), *Defining Guiding Principles for Processes Supporting Local Decision Making about Medicines: Final Report*, Liverpool: National Prescribing Centre, <http://www.doh.gov.uk>
- Okoro, C. (2008), 'Priority Setting Policy', <http://www.ealingpct.nhs.uk> [2 December 2009].

- Oliver, A. and C. Sorenson (2005), 'The limits and challenges to the economic evaluation of health technologies', in J. Costa-Font, C. Courbage and A. McGuire (eds), *Economics of Innovation and New Health Care Technologies*, Oxford: OUP.
- Raftery, J. (2009), 'NICE and the challenge of cancer drugs', *British Medical Journal*, **338**: 271–272.
- Rawlins, M. D. and A. J. Culyer (2004), 'National Institute for Clinical Excellence and its value judgments', *British Medical Journal*, **329**: 224–227.
- Richards, M. (2008), *Improving Access to Medicines for NHS Patients: A Report to the Secretary of State for Health by Professor Mike Richards CBE*, London: Department of Health.
- Russell, J. (2009), *Exploring Individual Funding Requests for Treatments in the NHS: A Background Paper (draft)*, London: Division of Medical Education, University College London, Mimeo.
- Schubert, M., T. R. Glass, S. P. Clarke, L. H. Aiken, B. Schaffert-Witlier, D. M. Sloane and S. De Geest (2008), 'Rationing of nursing care and its relationship to patient outcomes', *International Journal for Quality in Health Care*, **20**: 227–337.
- Syrett, K. (2003), 'A technocratic fix to the "legitimacy problem"? The Blair government and health care rationing in the United Kingdom', *Journal of Health Politics, Policy and Law*, **28**(4): 715–746.
- Wilson, E., S. Peacock and D. Ruta (2008), 'Priority setting in practice: what is the best way to compare costs and benefits?', *Health Economics*, **18**(4): 467–478.
- Wilson, E., J. Sussex, C. Macleod and R. Fordham (2007), 'Prioritizing health technologies in a primary care trust', *Journal of Health Services Research & Policy*, **12**(2): 80–85.