Reducing avoidable inequalities in health: a new criterion for setting health care capitation payments

Katharina Hauck, Rebecca Shaw and Peter C. Smith*
Centre for Health Economics, University of York, UK

Summary

Traditionally, most health care systems which pretend to any sort of rationality and cost control have sought to allocate their limited funds in order to secure equal opportunity of access for equal need. The UK government is implementing a fundamental change of resource allocation philosophy towards 'contributing to the reduction of avoidable health inequalities'. The purpose of this essay is to explore some of the economic issues that arise when seeking to allocate health care resources according to the new criterion. It indicates that health inequalities might arise because of variations in the quality of health services, variations in access to those services, or variations in the way people produce health, and that the resource allocation consequences differ depending on which source is being addressed. The paper shows that an objective of reducing health inequalities is not necessarily compatible with an objective of equity of access, nor with the objective of maximising health gain. The results have profound consequences for approaches towards economic evaluation, the role of clinical guidelines and performance management, as well as for resource allocation methods. Copyright

Capitation and health policy

In many developed nations, the system of health care finance is increasingly being used as an important instrument in seeking to secure a fair distribution of health care resources. To this end, many nations have increasingly turned to the use of 'capitations' as a basis for guiding the distribution of health care expenditure [1]. A capitation payment can be defined as the amount of health service funds associated with a citizen for a particular time period. Effectively, a capitation system puts a 'price' on the head of every member. Clearly, the expected health care expenditure needs of citizens vary considerably, depending on personal characteristics such as age, morbidity and social circumstances. Considerable effort has therefore, been expended on the process known as risk adjustment, which seeks an unbiased estimate of the expected costs of a citizen relative to all other citizens.

One of the earliest developments in the use of capitation methods in the finance of health care was the work in England of the Resource Allocation Working Party in the 1970s [2]. This sought to allocate a fixed National Health Service (NHS) budget to geographical regions in accordance with an equity criterion of seeking to secure 'equal opportunity of access for those at equal risk'. The methods adopted by the Resource Allocation Working Party have been superceded by more empirically based approaches [3,4]. However, the underlying equity objective has not changed.

*Correspondence to: Centre for Health Economics, University of York, YO10 5DD, UK

Received 9 November 2000
Accepted 9 October 2001

As with most systems of capitation, the premise underlying NHS capitation methods is that current patterns of expected utilisation should form the basis for making the equity criterion operational. However, for two reasons, the NHS has been reluctant merely to use unadjusted predictions of utilisation as the basis for capitation payments. First, current utilisation might to some extent reflect systematic variations in supply, implying that existing inequities might be perpetuated if no adjustment were made for such variations. And second, uncritical use of current utilisation as the basis for setting capitation payments might introduce a perverse incentive for local agents to increase current utilisation in order to attract higher capitation payments for their population in the future. These considerations have led to the development of a sophisticated econometric capitation methodology, principally, on the basis of small area socio-economic data [5]. The English capitation methods have been the subject of intense scrutiny, and have been implemented, or are under consideration for implementation, in a number of jurisdictions [6–10]. The methods seek to identify the national average response, in terms of health care expenditure, to a set of local socio-economic 'needs' indicators, after adjusting for supply factors.

It is important to note that, in common with most capitation methods, the English approach is conservative, in the sense that it assumes that (on average) the NHS is currently meeting the desirable concept of need, whatever that concept might be (for example, capacity to benefit, level of sickness, life expectancy, and so on). The methods fail to reflect legitimate health care needs that are not currently met by the NHS. We do not intend to enter here into the debate as to what is meant by need, although this clearly should be a germane focus of enquiry [11]. For the purposes of this paper, however, by 'unmet need' we merely seek to indicate that certain groups of the population systematically fail to receive the NHS health care which other groups with similar health status receive, implying a departure from the principle of equal access for equal need underlying the NHS. The use of empirical utilisation data as the basis of capitation payments will therefore perpetuate the inequity implied by the existence of unmet need, however need is defined (see [4] for a discussion of these issues).

The Labour government elected in May 1997 brought with it a policy of wishing to address the persistent and growing inequalities in health which had become evident in the United Kingdom. The Government set up an independent enquiry, chaired by Sir Donald Acheson, which recommended numerous policy options [12]. It then produced a policy document *Saving lives: our healthier nation* [13] which put in place a public health agenda, with the objective of 'improving the health of everyone, especially the worst off' – that is, of improving health and reducing health inequalities.

The commitment to reducing health inequalities has in turn resulted in a reappraisal of the capitation criterion in use in England. The Advisory Committee on Resource Allocation, the body charged with developing capitation methodologies in England, has been instructed by ministers to undertake a fundamental review of methods, incorporating a new criterion for determining capitation payments: 'to contribute to the reduction in avoidable health inequalities'. This criterion represents a radical departure from that of seeking to offer equal opportunity of access and in effect seeks to secure a redistribution of health. It steers health policy quite determinedly away from the narrow concept of *health care* equity towards the broader concept of *health* equity, with its implications for diverse policy areas such as income redistribution, housing, education, environment, transport, and so on.

The purpose of this paper is to put forward a simple economic model of health production, and to examine the implications of the new criterion for NHS capitation methods. In the following section we develop a model of the current capitation criterion. In the next section on causes of inequalities of health we further investigate various sources of inequalities in health, and we discuss which of these can be addressed by a change in capitation methodology. The paper introduces the new capitation criterion, and discusses some of its political implications, and the last section offers some concluding comments.

A model of the current capitation criterion

In this section we explore from a theoretical perspective why inequalities in health might arise,

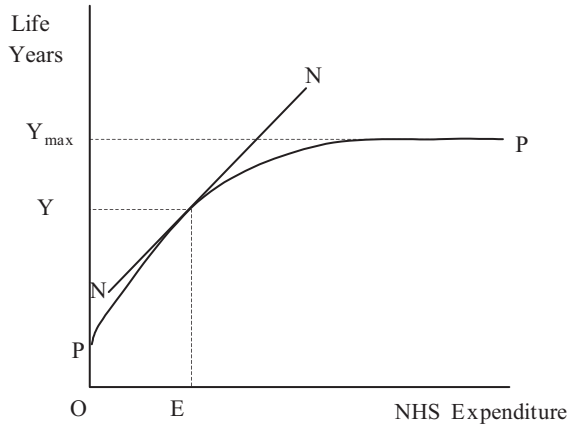


Figure 1. The health production function

and the implications for health care expenditure of seeking to reduce observed inequalities. The core of our exposition relies on an individual's health production function. This traces the efficient relationship between lifetime health care expenditure and health outcome, and is illustrated as the curve PP in Figure 1. For a given lifetime expenditure E on health care, and given current best clinical practice, the production function shows the maximum attainable health outcome (say life expectancy) Y of the individual. The maximum attainable life expectancy is Y_{mak}. The health production function is of course highly stylised, and requires careful examination before being used for analytic purposes.

First, we assume a single health care provider, which we can call the NHS. In practice, there may be available other providers of health care (such as private sector providers). For the purposes of this paper we think of these as being potential exogenous influences on the NHS production function shown in Figure 1. We also wish to side step the issue of which concept of 'health outcome' should be employed. The reader may wish to think of this as quality-adjusted life years. However, for expository purposes, we shall restrict discussion to a measure based on unadjusted life years. We do not believe that the choice of outcome measure materially affects the theoretical argument.

We define health care expenditure to be lifetime expenditure by the NHS, discounted to birth. The capitation criterion under investigation is directed at health inequalities *avoidable by the NHS*, and we therefore concentrate on health outcomes that can be affected by that agency's actions. In specifying such a function we are of course presuming that extra health care activity can contribute to increased health, a claim that could be open to challenge. It is moreover important to acknowledge that there are many other exogenous factors that may influence the nature of the health production function and consequent inequalities. These include the individual's genetic characteristics, occupation, use of non-NHS health care, lifestyle and other external influences such as the environment, the economy and the actions of governmental and other agencies. Changes in these factors might change the form of the NHS production function. For example, if an individual takes up a healthier lifestyle, this might give rise to an upward shift. We do not pursue these external influences further here, but it is worth noting that their inclusion in the model as a vector of circumstances is not in principle problematic.

Also, for ease of exposition, we assume a constant health care technology over the patient's lifetime. Of course, the rapid change in technologies that occurs in practice considerable complicates the practical problem for the health care system (if it is to secure productive efficiency). This interesting issue is however not germane to this theoretical discussion. More generally, we restrict the analysis to the deterministic case, and do not introduce uncertainty arising from technologies, individual characteristics, external circumstances or NHS budget constraints. In practice, the effectiveness of health care is likely – to a greater or lesser extent – to fall some way short of the ideal indicated by the production function. *Random* inefficiencies of this sort do not materially affect the argument. *Systematically* larger inefficiencies suffered by particular groups relative to others are however discussed in some detail below.

The question now arises: given the shape of an individual's health production function, how much expenditure should the health care system devote to that individual? In systems which are not budget constrained, we might in principle expect to observe expenditure up to the point where marginal benefit is zero. However, within a budget-constrained system of health care we must assume that some other criterion applies.

Many commentators argue that in these circumstances any decision rule for deciding how much to spend should be based on maximising the health output of the system, given its budget constraint. This principle gives rise to a simple decision rule: apply a uniform cut-off cost per life year saved,

above which no treatment is offered. The cut-off can be represented by the slope of the line NN in Figure 1, which yields the optimal expenditure for the individual under scrutiny, given the global budget constraint. The same sloped line is applied to all individuals, whatever the shape of their health production functions. This model underlies almost all the literature on economic evaluation in health care and the use of health benefit measures such as quality-adjusted life years. There is probably a widespread consensus amongst health economists that it is – or at least ought to be – the principal efficiency criterion for allocating resources in health care [14]. We term it the health maximisation model. It is important to note that – *if* we define need in terms of marginal capacity to benefit from health care – the health maximisation model is consistent with the founding principle of the UK NHS – that those in equal need should have equal access to services [15].

Causes of inequalities in health

Implicit in the new NHS capitation criterion, with its emphasis on avoidable health inequalities, is the suggestion that currently NHS resources are not being allocated in a socially desirable fashion. In particular, it suggests that, relative to their more healthy counterparts, the less healthy are receiving less health than is socially desirable. Three classes of circumstances within the NHS might give rise to this state of affairs:

- systematic variations in health care quality (variations in technical efficiency);
- systematic variations in utilisation of health care services (variations in allocative inefficiency),
- systematic variations in health production functions (variations in people's efficiency in producing health).

We now consider these sources of inequalities in turn, and we discuss which of them can be addressed by a change in capitation methodology.

Variations in health care quality

Suppose all individuals have the same production function and that the same cut-off criterion is applied to all individuals. That is, given the budget constraint, optimal expenditure E is being directed at all individuals. However, services for some classes of individuals are technically inefficient in the sense that they offer poorer quality than those for healthier individuals do – that is, outcomes lie below the production function frontier. This implies that treatments for two equally needy individuals differ due to variations in technical efficiency. This situation is represented in Figure 2 by the point L for the disadvantaged individual, giving rise to health outcome Y_L, as opposed to Y_H for the individual receiving better quality care.

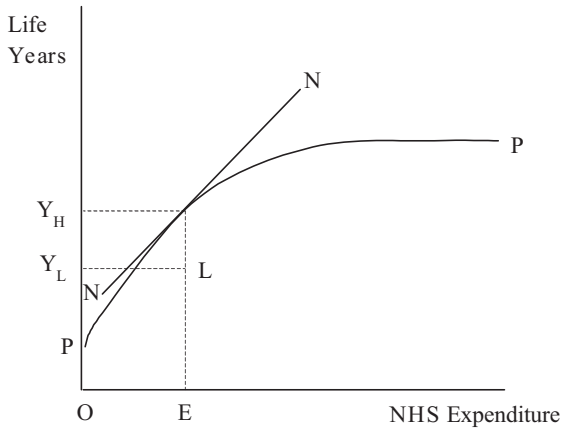


Figure 2. Inequalities in health arising from variations in technical efficiency for two individuals

Services to less healthy populations may be less technically efficient than other services for a number of reasons – expenditure may not be allocated optimally across an individual's lifetime, health care staff may be less motivated to secure good outcomes or may communicate poorly with less healthy individuals, recruitment of staff may be more difficult or capital configurations less appropriate in areas where the less healthy live, and so on. In this case, it is important to identify the true production possibilities, and to distinguish between improvements in outcome that can be secured by improved use of existing health care resources, and those that require additional resources. Addressing inequalities arising from technical inefficiency requires no change to capitation methods, because existing allocation of expenditure is optimal – it is the use of resources, which is inefficient.

It is important to note, however, that this builds on the assumption that technical inefficiency is

exogenous to the capitation system. It might be the case that the chosen capitation method provides an incentive to behave inefficiently. For example, capitation payments positively weighted for the current sickness of the population could provide an incentive not to use resources efficiently for fear of improving the population's health status and thereby losing budget. For a discussion of behavioural responses to fixed budgets see References [16–18].

For the purpose of this paper, however, we assume exogeneity of inefficient behaviour to capitation systems. Policy attention should, therefore, focus not on changing capitation methods, but on other instruments to secure better use of resources in services for disadvantaged populations. Various quality initiatives now put in place within the NHS, such as the publication of performance data and systems of audit and inspection, may help secure progress towards this objective [19].

Variations in utilisation of health care services

Suppose that all individuals have the same production function and all are being treated technically efficiently (that is, on rather than below the production function). However, a stricter cut-off criterion is applied to some classes of individuals than to others, implying the existence of allocative inefficiency. This may, for example, be due to market or informational failures on the demand or supply side of health care. Inequalities in utilisation have the consequence that, although needs are identical, expenditure on health care is less for some groups than others. Figure 3 illustrates the principle for two individuals, with the stricter treatment criterion applied to the disadvantaged individual L resulting in lower expenditure E_L and poorer outcome Y_L than for the other individual H. Under these circumstances, use of capitation payments E_L and E_H based on empirical data will perpetuate the implied inequity.

If a stricter cut-off criterion is currently being applied to some individuals than to others, a fundamental principle underlying many health care systems (and certainly the NHS) is being breached – that of equal access to health care for those in equal need. There is indeed evidence of considerable unmet need and of substantial inequalities in utilisation in UK health care [20]. Minority ethnic groups, disadvantaged socio-economic groups, the elderly and persons living in remote areas experience inequalities, most notably in primary care, in prevention and health promotion, and in the treatment of coronary heart disease. These inequalities in utilisation are present even after controlling for 'need'.

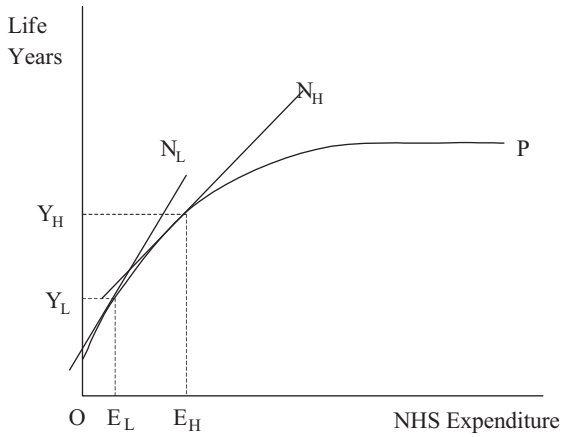


Figure 3. Inequalities in health arising from variations in access for two individuals

Inequalities in utilisation unrelated to need imply that health maximisation is not being secured, because the underserved have a greater capacity to benefit from expenditure than the relatively 'overserved'. A redirection of resources towards 'underserved' individuals is required, with an implication that capitation payments for disadvantaged populations should rise relative to the remainder of the population. This does not require definition of a new criterion for setting health care capitation payments. Rather, it requires the formulation of strategies aimed at eliminating allocative inefficiency in the provision of health care. The policy implication is therefore to design interventions that reduce utilisation inequalities. The nature of these will of course be highly dependent on the reason for inequalities in access to services. In practice very few studies have sought to address such policy issues [20–22].

For the purposes of capitation, attention should therefore focus on the magnitude of the associated unmet need, and on the expenditure consequences of rectifying the problem. In terms of Figure 3, the requirement is to quantify the shifts in expenditure $E_L E_H$ required to ensure that all citizens receive the same level of care. By definition, uncritical analysis of *existing* expenditure patterns will not yield useful information for this purpose. In

principle, we should therefore seek out variations in the slope of the cut-off criterion applied to different social groups. In practice this is likely to be difficult. However, it may be that there exist areas of the country where the unmet need has been eliminated, and that analysis of existing expenditure patterns within those areas may yield an acceptable basis for setting national capitation payments. Sutton and Lock show how this could be done in a Scottish context [23], although the rather arbitrary method of selecting 'exemplar' areas adopted in that study indicates the type of practical problems likely to be encountered. Of course, even if capitation payments can be corrected to account for unmet need, there remains a performance management problem of ensuring that the increased funds associated with unmet need are indeed directed towards the currently underserved population.

Variations in health production functions

Suppose that all individuals are being treated with technical and allocative efficiency, in accordance with the health maximisation model. However, individuals have different health production functions, so that their health outcomes vary. This situation can be illustrated in Figure 4, which compares two individuals with different health profiles, in the sense that – at the same level of health expenditure – individual L is unambiguously less healthy than individual H. This is due to determinants of health which are beyond the immediate influence of the health services, such as the social and economic environment, genetical endowments, or lifestyle choices of the individuals. The cut-off criterion is indicated by the slope of the straight lines, and gives rise to health outcomes Y_H and Y_L. The implied capitation payments are E_H and E_L. Application of an equal cut-off criterion implies smaller health inequalities in comparison to an equal allocation of expenditure to H and L.

If all patients are being treated in accordance with the health maximisation principle, but the outcome is nevertheless unacceptable, then a reallocation of resources according to some equity criterion is required, under which resources are redirected towards less healthy individuals. Avoidable inequalities of this sort arise, even though quality of and access to health care are equal for identical citizens, because of differences between individuals which are outside the control of the

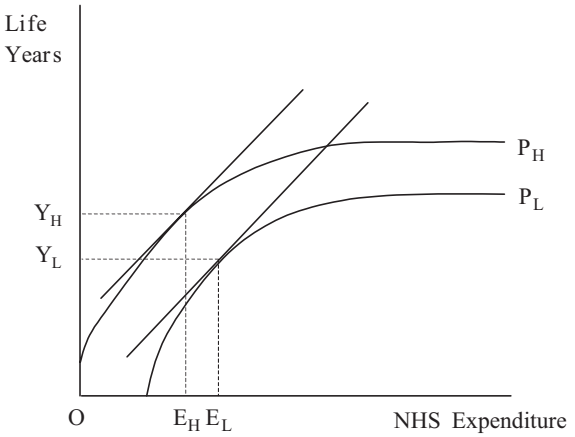


Figure 4. Inequalities in health arising from different production functions for two individuals

health services. Policy attention to such inequalities reflects a concern with principles of *vertical* equity between individuals, rather than the traditional concern with horizontal equity embedded in most capitation methodology. In principle, society should address vertical equity issues by considering an optimal reallocation of all resources, both private and public. However, our focus is purely on the health care sector, and in this context the unacceptable health inequalities imply that a fundamental revision of capitation methods may therefore be required.

A model of the new capitation criterion

Policy to correct for variations in people's efficiency in producing health implies an interest in increasing the level of health care for the less healthy relative to that received by the healthy in order to compensate for such disadvantage. Like Case 2, this implies a shift of health care resources in the form of capitation payments towards the less healthy. In contrast to Case 2, however, the objective here is to rectify a perceived injustice in individual endowments, and not inefficiencies within the health care system. In the extreme case, of wishing to *eliminate* avoidable health inequalities, a situation as in Figure 5 might obtain. Expenditure on individual L is increased in order to secure the same life span as currently enjoyed by individual H. This results in increased capitation payment E_L^*. Note that the marginal cost per life

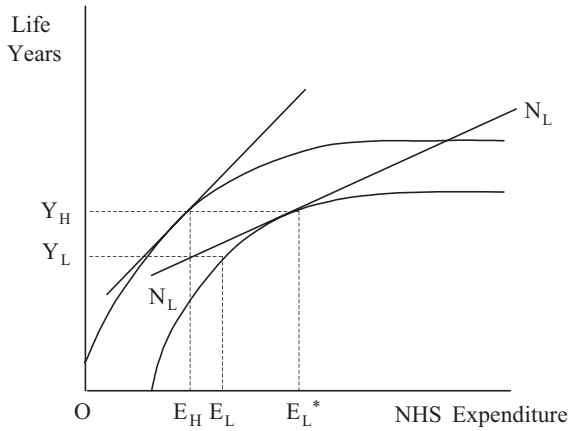


Figure 5. Expenditure change required to equalise life expectancy

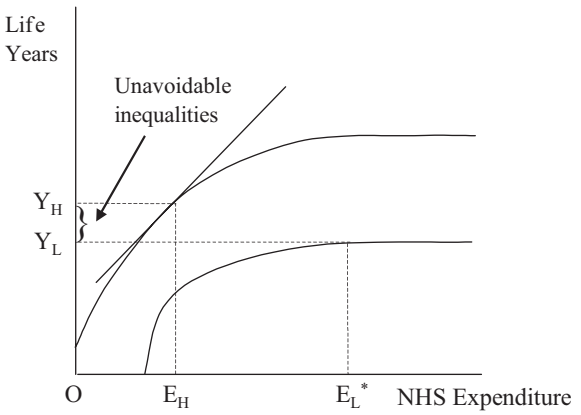


Figure 6. Inequalities unavoidable by the health services

year saved becomes higher for individual L (the associated line $N_L N_L$ becomes shallower than the original NN). This might imply that the unhealthy individual receives treatments which the healthy individual does not receive, or that the unhealthy individual receives more expensive treatments, or treatments of a higher quality.

The situation set out in Figure 5 would result in an unambiguous rise in the health care budget requirement. If this were considered unrealistic, the solution would be simultaneously to reduce expenditure on individual 1 whilst increasing expenditure on individual 2. That is, health inequalities would be reduced partially by *worsening* the outcome for healthier individuals. In Figure 5, a revenue neutral solution would then result in a common life expectancy somewhere between Y_H and Y_L. (Whether this is politically feasible is another matter!).

The strategy of *eliminating* avoidable mortality is of course extreme. In practice, both limited technological capacity and strength of public preferences might give rise to a policy reluctance to seek to eliminate variations entirely. A more realistic criterion is therefore to *reduce* avoidable inequality. Figure 6 shows a situation where some unhealthy individuals are unable to achieve the same life span as individual H, in which case the health services would – under the criterion of 'eliminating avoidable inequality' – spend up to the point where the marginal benefit of health care expenditure and the slope of the cut-off is zero. The remaining inequalities – symbolised by the distance between Y_H and Y_L could only be eliminated by reducing the health status of individual H. If this is politically undesirable, the remaining inequalities are unavoidable by the NHS.

It is likely that a broader view of social policy would indicate that interventions in other public policy areas – such as housing, public transport or income redistribution – are effective in eliminating health inequalities. Successful policies in other areas would result in an upward shift of the health production function in Figure 6. This argument can be extended to the case where it is possible but inefficient to reduce inequalities with health care interventions. In a situation where inequalities in health could be further reduced only with very high health care expenditure, public expenditure in other policy areas might lead to exogenous improvements in health production. This combination of strategies requires less public expenditure than a solely health care policy strategy, if the expenditure on alternative policies are smaller than the saved health care expenditure. It is possible, that health care policy is entirely inefficient in tackling inequalities in health care, and all inequalities are reduced more economically with policies in other areas. In order to make this assessment, the marginal effectiveness in reducing inequalities in health of alternative public policies (and possibly even portfolios of policies) should be compared. In principle, a socially optimal health inequalities policy would allocate resources across policy areas so that the marginal benefit of public expenditure (in terms of reducing health inequalities) would be equal in each policy area.

There is another reason why society might not want to adopt the complete elimination of inequalities as an objective. The new criterion

requires a sacrifice in overall population health. In Figure 5, total health gain for a fixed budget is less than in the health maximisation model outlined in Section 2 – the common life expectancy under the new capitation criterion is less than the average of Y_H and Y_L. The policy maker's problem now becomes one of balancing total health gain (an efficiency objective) against reductions in inequalities (an equity objective) [24,25]. A possible algebraic formulation of this problem is given in Appendix A. The problem can be illustrated as in Figure 7, which traces the health production possibilities arising from the health production functions for two individuals with different levels of health (Appendix B shows how Figure 7 can be derived from the individual production functions). It indicates – for a fixed budget constraint – the possible mixes of maximum health outcomes Y_L and Y_H that the NHS could in principle secure for the two individuals. The point H^* indicates the maximum aggregate health attainable for the two individuals subject to the given budget constraint. The point Q^* is the point where the two would secure equal health, and the distance Q^*Q_0 indicates the aggregate loss in health brought about by pursuit of such pure equality. In practice, it seems likely that there exists a social welfare function which results in a policy intermediate between the points Q^* and H^*, reflecting the politically preferred balance between efficiency and equity objectives.

The new criterion for setting capitation payments requires a clear normative definition of the concept of equity in health which policy makers have in mind. Any deviation from the equal cut-off criterion may imply that individuals with the same capacity to benefit from health care receive different amounts of health care resources. Unequal treatments require political justification, and this is the role of the equity concept. There is a substantial if not always enlightening theoretical economics literature on equity concepts in health and health care [26], and there has been little empirical examination of what meanings or precise specifications stakeholders attach to the concept [27]. Therefore, it will be difficult to find agreement on a particular equity concept. Moreover, once identified, the theoretical equity concept needs to be translated into an unambiguous resource allocation pattern. This is without doubt an acute political problem [28].

Finally, it is worth noting that the new capitation method implies better medical treatment of

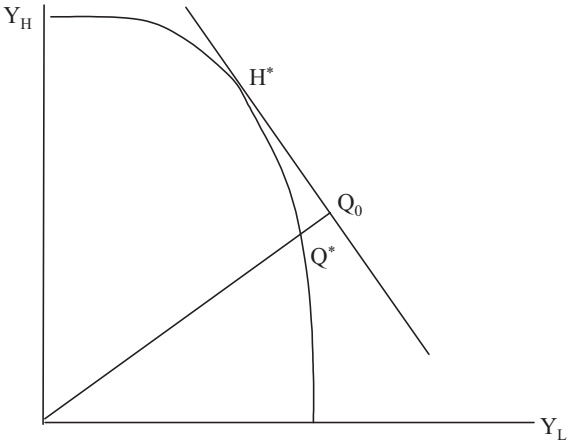


Figure 7. The health production possibility frontier

unhealthy groups of the population. This gives rise to major practical difficulties in defining criteria for membership of the targeted group, and ensuring that health care is delivered in accordance with policy intentions. Furthermore, it might result in incentives for individuals to acquire membership of those groups that are given privileged access. The variety of practical difficulties that emerge when seeking to make operational principles of vertical equity – as distinct from horizontal equity – are considered elsewhere [29].

Conclusions

This paper has sought to link the economic literature on health inequalities with the policy issue of capitation payments in the light of a new capitation criterion adopted in England. It has demonstrated that there are three broad categories of causes of health inequality relevant to the health sector: variations in efficiency, variations in access to care, and variations in personal health production. This last poses the most fundamental challenge to capitation policy, as addressing it implies a desire to move away from a policy of equality of access (horizontal equity) towards one of targeting health care at particular classes of individual (vertical equity). There is clearly a major challenge in seeking out the evidence on which the change to the capitation methods would be based. Two broad classes of information required relate to the effectiveness of interventions in reducing health inequalities, and public

preferences regarding the importance of reducing health inequalities. Both sorts of evidence are in short supply [30,31].

Furthermore, there is no guarantee that mere alteration of capitation payments will ensure that additional resources reach deprived populations. By definition, the new vertical equity criterion requires that the health sector alters the way in which it delivers health care to those with poor health expectancy. Yet, in general, directing extra 'health inequality' resources at needy areas will not necessarily lead to reduction in health inequalities. Rather, it may merely lead to perpetuation of existing patterns of utilisation in an area, albeit at a higher level than before. Important performance management and auditing issues are, therefore, raised if the policy reflected in the revised capitation payments is to be translated into desired action by health care professionals.

The discussion has emphasised the role of health services in addressing health inequalities, and has made only general reference to the broader influences of social policy on inequalities. This emphasis reflects the current administrative reality – that health ministries perceive their principal role to be one of delivering health care. Yet, there is no reason in principle why health ministries should not be responsible for addressing the health inequality implications of *all* areas of public policy. Under this arrangement they would be responsible for auditing the impact on health inequalities of major public sector initiatives, and for levying 'taxes' (or providing subsidies) to encourage policies that contribute to health inequality policy. Nurturing this role would be one approach towards the optimal distribution of all public resources.

It is also important to note that an emphasis on health inequalities offers a profound challenge to the evaluation of health care technology. In principle, it implies that technologies should be evaluated differently according to the health status of the individual – that is, the need to target certain unhealthy groups may mean that certain treatments are recommended for those groups which are not considered cost-effective for healthier groups. This consideration complicates the task of designing and evaluating trials enormously, and implies a move towards Williams' notion of equity-adjusted QALYs as the basis for economic evaluation. The principle also offers considerable challenges in framing intelligible clinical guidelines for practitioners. Yet the logic of incorporating a health inequality criterion into resource allocation leads inevitably to its incorporation into economic evaluation of technologies, with all the attendant complications.

The new capitation criterion, therefore, raises many more general issues relating to the distribution of resources in health care and the broader public sector. Many of these issues are immensely challenging. However, we believe that the economic models presented here offer a systematic and coherent framework for addressing these challenges.

Acknowledgements

We are grateful for comments from our colleagues Nigel Rice, Aki Tsuchiya and Alan Williams, from some of the participants at the inaugural meeting of the UK Health Equity Network, at which this paper was presented, and from two anonymous referees. Hauck is supported by a CHE research studentship. Shaw is supported by the ESRC Health Variations Programme project L128251050. Smith is in part supported by the Department of Health research programme at CHE and the ESRC.

Appendix A – An algebraic exposition

Assume that NHS expenditure is distributed efficiently over the lifetime of the patient, and that the objective is to maximise health outputs (summed across all individuals), with a fixed budget constraint. If the individual's health production function is $f_i(.)$, then a health maximisation objective is allocate expenditure x_i so as to maximise $\sum_i f_i(x_i)$, subject to $\sum_i x_i = B$, where B is the global budget constraint. This yields the familiar result that the optimal level of expenditure x_i^* on individual i is given by $\partial f_i/\partial x_i = \phi$, indicating an identical benefit/cost cut-off ratio N for all individuals i.

Suppose now the policy maker's problem becomes one of balancing total health gain (an efficiency objective) against reductions in inequalities (an equity objective). One possible algebraic formulation of the problem is as follows.

$$\text{Maximize} \quad \alpha \sum_i f_i(x_i) - 1/\varepsilon \sum_i \{L - f_i(x_i)\}^{\varepsilon}$$

$$\text{subject to} \quad \sum_i x_i = B$$

where L is some benchmark 'maximum' achievable health outcome (lifespan), α reflects the importance of the total health outcomes objective relative to the inequalities objective, and $\varepsilon > 1$ indicates the degree of aversion to health inequalities. A value of $\varepsilon = 1$ would indicate no interest in health inequalities, while increases imply an increase in the relative importance attached to large deviations from L.

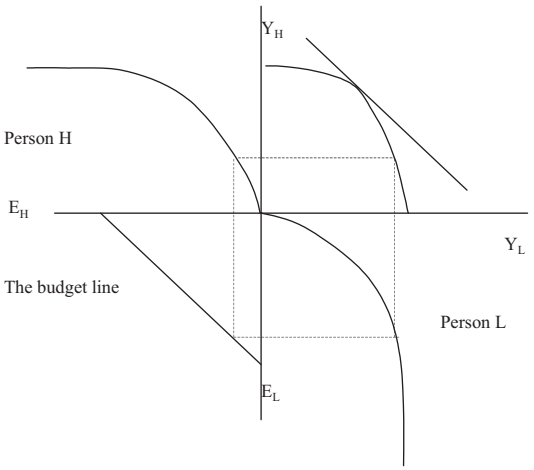
The first-order condition for individual i is given by

$$\frac{\partial f_i}{\partial x_i} = \frac{\lambda}{\alpha + \{L - f_i(x_i)\}^{\varepsilon - 1}}$$

where λ is the Lagrange multiplier. Thus, in general, the solutions x_i^* will be such that individuals with high life expectancy are given higher treatment cut-off ratios N. Note that, in the special case when $\varepsilon = 2$, α and L are interchangeable, in the sense that an increase in L is directly equivalent to an increase in α. The problem therefore reduces to one of minimising squared health variations, defined as $\sum_i \{L - f_i(x_i)\}^2$. The choice of 'lifespan' parameter L reflects the importance of maximizing health relative to reducing inequalities and no additional choice of α is required – choice of a high value of L then leads to relatively low variations in personal values of N. In these circumstances, $\partial f_i / \partial x_i \propto \{L - f_i(x_i)\}^{-1}$ – that is, at the optimum, cut-off rates will be inversely proportional to life years lost, as defined by $\{L - f_i(x_i^*)\}$.

Appendix B – Derivation of health production possibility frontier

This appendix indicates how the (two person) production possibility frontier can be derived from the individual health production functions. The production frontiers for person H and person L are replicated (in a transposed state) in respectively the top left and bottom right corners of the diagram. The fixed expenditure budget constraint $E_H + E_L$ is represented by the straight line in the bottom left quadrant. All expenditure choices must conform to this constraint. They are then reflected, via the production functions, into the top right quadrant, which therefore, yields the production possibility frontier, which is reproduced as Figure 7 in the main text.



References

1. Rice N, Smith, P. Capitation and risk adjustment in health care financing: an international progress report. *The Milbank Quarterly*, 2001; **79**: 81–113.
2. Resource Allocation Working Party. *Sharing Resources for Health in England*. HMSO: London, 1976.
3. Royston GHD, Hurst JW, Lister EG, Stewart PA. Modelling the use of health services by populations of small areas to inform the allocation of central resources to larger regions. *Socio-Econ Plann Sci* 1992; **26:** 169–180.
4. Smith P, Rice N, Carr-Hill R. Capitation funding in the public sector. *J Roy Statist Soc Ser A*, 2001; **164**: 217–257.
5. Carr-Hill RA, Sheldon TA, Smith P, Martin S, Peacock S, Hardman G. Allocating resources to health authorities: development of methods for small area analysis of use of inpatient services. *British Med J* 1994; **309**: 1046–1049.
6. Welsh Office. *NHS Resource Allocation Review: Emerging Findings Report to the Health and Social Services Committee*. Welsh Office: Cardiff; 2001.
7. Department of Health and Social Services. *Allocating Resources to H&SS Boards: Proposed Changes to the Weighted Capitat Formulae – A Third Report from the Capitation Formula Review Group*. Department of Health and Social Services: Belfast; 2001.
8. Scottish Executive Health Department. *Fair Shares for All: Final Report*. Scottish Executive Health Department: Edinburgh; 2000.
9. Häkkinen U, Mikkola HM, Nordberg M, Salonen M. *Tutkimus Kuntien Terveyspalveluiden Valtionosuuksien Perusteista [A Study on the Foundations of*

the State Subsidies for Municipal Health Services]. Sisäasiainministeriö: Helsinki; 1996.
10. Pampalon, R. L'approche québécoise en matière d'indicateurs de besoins pour l'allocation régionale des ressources, in CREDES. In *Allocations des Ressources, Géographie des Soins*. CREDES: Paris; 1998.
11. Culyer AJ. Need – the idea won't do – but we still need it. *Social Sci Med* 1995; **40**: 727–730.
12. Acheson D. *Independent Inquiry into Inequalities in Health*. HMSO: London, 1998.
13. Department of Health. *Saving Lives: Our Healthier Nation*. The Stationery Office: London, 1999.
14. Culyer AJ. Health, health expenditures and equity. In *Equity in the Finance and Delivery of Health Care: An International Perspective,* Doorslaer EV, Wagstaff A, Rutten F. (eds). Oxford University Press: Oxford, 1993.
15. Department of Health and Social Security. *Sharing Resources for Health in England: Report of the Resource Allocation Working Party*. HMSO: London, 1976.
16. Whynes D, Heron T, Avery A. Prescribing cost savings by GP fundholders: long term or short term? *Health Econom*, 1997; **8**: 335–343.
17. Shmueli A, Glazer J. Addressing the inequity of capitation by variable soft contracts. *Health Econ* 1999; **8**: 335–343.
18. Croxson B, Propper C, Perkins A. Do doctors respond to financial incentives? UK family doctors and the GP fundholding scheme. *J Public Econ* 2001; **79**: 375–398.
19. Department of Health. *The NHS Plan: A Plan for Investment, A Plan for Reform*. The Stationery Office: London, 2000.
20. Goddard M, Smith P. *Equity of Access to Health Care*. Centre for Health Economics, University of York: York, 1998.
21. Gordon D, Shaw M, Dorling D, Davey-Smith G. *Inequalities in Health: the Evidence Presented to the Independent Inquiry into Inequalities in Health, chaired by Sir Donald Acheson*. The Policy Press: Bristol, 1999.
22. NHS Centre for Reviews and Dissemination. *Review of Research on the Effectiveness of Health Service Interventions to Reduce Variations in Health*. Centre for Reviews and Dissemination: York, 1995.
23. Sutton M, Lock P. Regional differences in health care delivery: implications for a national resource allocation formula. *Health Econ* 2000; **9**: 547–559.
24. Williams A. Beyond effectiveness and efficiency lies equality!. In *Non-random Reflections on Health Services Research,* Maynard A, Chalmers I. (eds). BMJ Publishing Group: London, 1997.
25. Wagstaff A. QALYs and the equity-efficiency trade-off. *J Health Econ* 1991; **10**: 21–41.
26. Williams A, Cookson R. Equity in health. In *Handbook of Health Econom,* Culyer AJ, Newhouse J (eds). Elsevier: Amsterdam, 2000.
27. Pereira J. *What Does Equity in Health Mean?* Discussion Paper 61. Centre for Health Economics: University of York, 1989.
28. Culyer AJ, Wagstaff A. Equity and equality in health and health care. *J Health Econ* 1993; **12**: 431–437.
29. Mooney G. And now for vertical equity? Some concerns arising from aboriginal health in Australia. *Health Econ* 1996; **5**: 99–103.
30. Lindholm L, Rosen M, Emmelin M. How many lives is equity worth? A proposal for equity adjusted years of life saved. *J Epidemiol Community Health* 1998; **52**: 808–811.
31. Andersson F, Lyttkens CH. Preferences for equity in health behind a veil of ignorance. *Health Econ* 1999; **8**: 369–378.