Review of the Weighted Capitation Formula

Gwyn Bevan
Professor of Management Science
Department of Management and LSE Health and Social Care
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
England
Tel: 020 7955 6269
FAX: 020 7955 6885
Email: R.G.Bevan@lse.ac.uk

June 2008

A report submitted to the Secretary of State for Health
Index to Tables

Table 1: Changes to the weighted capitation formula since 1999 33
Table 2: Ranges and ratios of indices for different components 53
Table 3: Work commissioned in developing the weighted capitation formula since 1999 65

Index to Figures

Figure 1: AREA index for additional need and under-75 SMR .........................19
Figure 2: Actual and Estimated Spend on Health Care for High Risk Populations ...30
Figure 3: Range in % Distance from Target from 1998/99 to 2007/08 ...............32
Figure 4: Weighting of different components in the formula .........................52
Figure 5: Ratio of maximum to minimum of indices for HCHS, Pharma, PMS .......54
Figure 6: Ratio of maximum to minimum of indices for HCHS, Pharma, PMS, HIV
weighted by spend on each component .......................................................55
Figure 7: Indices for additional need and age for HCHS ................................56
Figure 8: Indices for MFF and age for HCHS ..............................................56
Figure 9: Indices for additional need for HCHS, Pharma and PMS ....................57
Figure 10: Indices for age for HCHS, Pharma and PMS ................................58
Figure 11: Indices for MFF for HCHS and PMS ..........................................59
Figure 12: Indices for EACA and HIV/AIDS .............................................60
Figure 13: Unified vs HCHS weighted populations (000s) ..............................60
Review of the Weighted Capitation Formula

Summary

i. The Advisory Committee on Resource Allocation (ACRA) commissioned this Report, on a review of the weighted capitation formula for revenue allocations to Primary Care Trusts (PCTs), for the Secretary of State for Health. The terms of reference were to: carry out a study, which will chart the work done, and the decisions made, by ACRA in relation to the review of the formula, prior to making revenue allocations for 2003/04 to 2005/06 and 2006/07 and 2007/08; examine current academic opinion and international best practice in this area; and produce a Report to ACRA in April 2007.

ii. Since the Report was submitted to the Secretary of State, significant progress has been made in developing capitation formulas and new lines of research have been commissioned. This Report refers to, but does not review, these important developments, and has been revised to read as a retrospective review as of June 2008. One of the points made in this Report is the gap in publishing ACRA’s work: reports were published regularly until 2000, but then there was a gap until new material was published in 2008. One of the objectives of this Report is to fill that gap hence it has been written for a general readership.

iii. Although capitation formulas aim to promote equity, there are, however, questions over what definition of equity is being sought by these formulas, and limitations to what can be achieved by resource allocation alone. The seminal 1976 Report of the Resource Allocation Working Party (The RAWP Report) interpreted the underlying objective of its terms of reference to be ‘to secure through resource allocation equal opportunity of access for people at equal risk’ and this was seen as the bedrock principle against which all methods of developing capitation formulas ought to be judged. But official Reports that have stated this to be their objective, have recommended methods that were actually designed to equalise spend per capita across populations, rather than within them in terms of age, socio-economic or ethnic groups. From 1999 ACRA was asked to develop formulas to include a health inequalities adjustment (HIA), with the objective of reducing inequalities in health, which applied for two years only (2001/02 and 2002/03).

iv. Funding based on a capitation formula is directed at achieving equity and not funding existing demand or supply. There have recently been claims that the formula used for resource allocation was a cause of financial deficits in the NHS. These claims have been examined by the chief economist of the Department of Health, who showed any such link to be spurious. As deficits are a failure to match expenditure to income, they can never be good grounds for increasing incomes. Seeking to fund demand or supply regardless of relative need is in effect to abandon the principles of equity that have governed resource allocation to the NHS since 1977. Purchasers need to manage demand and design contracts with providers so that payments to providers can be afforded: through ‘Payment by Results’ (PbR) for secondary care, and the Quality and Outcomes Framework (QoF) for primary care. The cash limit of a
PCT, determined by the estimated relative needs of its population, will not automatically equal payments under PbR and QoF. Nor will changes in a PCT’s budget be increased in line with increases in volumes and quality of hospital and primary care services delivered.

v. There are four fundamental elements to weighted capitation formulas: estimates of:
   a. the number and characteristics of populations in each area;
   b. weights for demography;
   c. weights for additional need; and
   d. adjustments for unavoidable variations in costs.

vi. Even the basic building block of population statistics at the level of general practices is insecure: we still lack a good explanation for ‘list inflation’: the estimates of population derived from lists of patients registered with GP practices tend to exceed those from censuses. These differences are material for some PCTs and unknown at the level of general practices. There is no single body responsible for the accuracy of population data on registered lists. It is not possible to derive sound estimates using formulas at the level of practices without resolving adequately the problem of ‘list inflation’.

vii. There are, in principle, two ways of deriving weights for demography and additional need; and adjustments for unavoidable variations in costs. First, ideally from the bottom up, on a normative basis: by estimating what care ought to be provided and cost in different parts of the country. Unfortunately this bottom-up approach proved difficult. We lack detailed data on: the prevalence and incidence of disease in England and its severity; and evidence of what treatment ought to be provided and cost. So, to make good these fundamental gaps in data and evidence, a top-down approach offers an attractive alternative in estimating relative need and unavoidable variations in costs, but this requires making assumptions in analyses of empirical data.

viii. The stumbling block encountered by all empirical analyses is seeking to estimate what ought to be from what is, so that we avoid perpetuating underlying inequities and inefficiencies. And empirical methods are further handicapped by fundamental problems with data: in estimating additional need, we lacked data at the individual level on measures of risk and utilisation; in estimating unavoidable labour costs, known as the staff Market Forces Factor (MFF), the analyses were handicapped by lack of data on travel to work areas.

ix. Capitation formulas are subject to: ambiguities over definitions of equity; conceptual and methodological difficulties in estimating need and unavoidable variations in costs; and problems and limitations with data. Hence any capitation formulas will be imperfect and open to criticism and challenge. Even if data and evidence were to become available to allow the development of a bottom-up approach in estimating needs, this would create new problems, as the estimate of what we ought to spend would differ from the resources that were available.
x. The RAWP report was able, within a year, to recommend a simple formula, to
guide allocations to 14 Regional Health Authorities, driven by three elements
only: resident populations; weighting for demography by estimated national
average costs and additional needs by Standardised Mortality Ratios (SMRs).
The only element that has (essentially) stayed the same is RAWP’s approach
to weighting for demography. The RAWP report recognised the problem of
unavoidable costs, but was unable to recommend a method to take account of
this. The RAWP approach was characterised by elegant economy supported by
heroic assumptions on the relationship between the SMR and relative need for
resources.

xi. Later work has sought to develop an evidence base in accounting for
additional needs and unavoidable costs from analyses of empirical data:
estimating additional need from analyses of small-area variations in utilisation;
and unavoidable costs in the staff MFF from the general market approach
(based on analyses of labour markets outside the NHS). There have been
criticisms of the methods used to account for need. But, people in the NHS
perceived the weakest element of the formula, to be the estimates of the staff
MFF. The scope of the MFF was extended, following a review in 1993: from
the South East to the whole of England; and in coverage of staff (from 35% to
nearly 60% of total revenue expenditure). Since then the staff MFF has been
seen to lack face validity and be a constant source of dissatisfaction and
criticism. There are questions over the savings that can be made in areas with
low MFF, problems in translating indices of labour markets into what ought to
be differences in labour costs in the NHS, and questions over whether medical
and dental staff ought to be included.

xii. The reorganisation of the NHS, with general practices as the building block,
has created three new sets of problems for ACRA. First, relating estimates by
the Office for National Statistics (ONS) to practices: with ‘list inflation’; and
the ‘attribution problem’ of weighting variables that use data from censuses
for practice populations where the populations in wards are registered in
different practices. Second, extending the scope of formulas, to cover primary
care and prescribing. Third, considering the applicability of formulas at the
small scale of general practices (see below).

xiii. ACRA has sought, since 1999, to develop innovative approaches to the
development of formulas in resource allocation. This was essential in
developing the Health Inequality Adjustment (HIA: no other country has
attempted this). ACRA, in 2007, commissioned two exploratory studies of the
bottom-up approach to resource allocation and a person-based approach to
analyses of data on utilisation. ACRA explored the specific cost approach to
estimating unavoidable costs, which was informative but did not offer a sound
alternative to the general labour market approach. The research ACRA has
been able to commission has been focused on producing results that can be
used to change formulas with short time scales using available data and there
has been limited exploration of approaches using different kinds of data.

xiv. DH publishes clear accounts of the bases for the formula used for resource
allocation. But there has been nothing published on developments in resource
allocation for the past five years and no report from ACRA since 1999. It would be beneficial to have publications on: the rationale of the current formula and accounts of alternatives that ACRA has explored and why these have been rejected; what is happening to inequalities in the NHS and the impacts from various changes made to formulas.

xv. The process that has developed since 1997 with standing policy and technical group advisory groups on resource allocation has worked well. These groups have: considered the scope of developments in the light of Ministerial priorities; set priorities; tried to develop new ways of accounting for a HIA, additional need and avoidable costs; been focused on making changes for each year with best use of data and methods. Hence ACRA has identified what had to be done to provide a sound basis for resource allocation each year and also sought to map out options for the longer term. Work on improving the population base, which has not been a subject for research, has been done in-house. Research for ACRA has been conducted in an exemplary way: the contracts were let through competitive tendering, the studies were done by experts using state-of-the-art techniques, their work was subject to challenge from experts, and from 1997, governed by a policy and technical group. Since the mid-1990s, however, a small group of researchers have consistently won contracts for the work commissioned by ACRA, which makes this vulnerable to criticisms of being a cliquish activity. This is a well-recognised characteristic in contracting for such work. The best way of opening up competition would be by ACRA funding longer-term research programme into resource allocation as suggested below.

xvi. ACRA agreed two sets of criteria that would be applied in evaluating formulas for resource allocation:
   a. Essential: technical robustness; transparency; objectivity; plausibility; freedom from perverse incentives; reliability of calculation.
   b. Desirable: comprehensibility to non-specialists; durability; practicality; clarity of contribution of indicators; flexibility; stability; materiality.

xvii. All ACRA’s recommendations have to satisfy the criteria of practicality, durability, stability, reliability of calculation, and consistent applicability. Throughout there are tensions between: transparency, comprehensibility and technical robustness; and flexibility and materiality. The criterion of technical robustness requires sophisticated statistical methods to estimate weightings for additional need, which are no longer transparent and comprehensible to non-specialists. Estimates of the MFF have been made using the general labour market approach, which appears to be the only approach that is technically robust, but the MFF has consistently been seen to lack face validity.

xviii. ACRA has to make judgements on the materiality of refinements to the formula: there are particular problems where an element affects a few PCTs only, or where there may be a set of related issues none of which is material in its own right, but may have a material cumulative impact. Part of the problem here is developing a national formula that has to be applied consistently across PCTs with very different characteristics in terms of the needs for, and supply of, health care. (One possible approach would be to allocate to each Strategic
Health Authority (SHA) a discretionary sum for them to allocate to PCTs to reflect special local circumstances.) There looks to be a good case for ACRA reviewing the materiality of the allowances for: the English Language Difficulties Adjustment (ELDA); the Emergency Ambulance Cost Adjustment (EACA); and HIV/AIDS. There are also questions over the materiality of developing different indices for the different components of expenditure. It would be helpful to enforce stronger criteria on materiality so that ACRA can focus on getting the big things right: better data on population, more satisfactory ways of accounting for unavoidable variations in costs, and economical ways of accounting for demographic and additional need that are sensitive to differences across different types of health care.

xix. ACRA and the Technical Advisory Group (TAG) operate on the tradition of voluntary input from members. As the NHS is the key group that ought to have confidence in the formulas, it is appropriate that each body is chaired by NHS chief executives who are respected by their peers. It would be inappropriate to transform ACRA into an independent executive body responsible for determining allocations to PCTs subject to terms of reference, but otherwise independent of Government, like the Monetary Policy Committee (MPC). There is, however, scope for clarifying ACRA’s contributions to policy. It would be helpful for DH to publish annual reports on ACRA’s deliberations, recommendations, and Ministers’ decisions. The model used in New Zealand in their reviews of their funding formula is a process of five-yearly reviews by an advisory group (with members from District Health Boards and technical experts), which begins with a formal consultation with District Health Boards. This model is worthy of consideration for England.

xx. There is much to be proud of in the development of capitation formulas for the NHS in England. That does not, however, mean that what we have is perfect. Although the terms of reference of my review did not ask for recommendations, it is appropriate to end with a few suggestions for improving the processes through which capitation formulas could be developed in future:
   a. To clarify the objective of formulas: are they to promote equity of access for equal risk or reduce health inequalities, or both?
   b. To clarify ACRA’s remit: ought this be limited to developments of the formula, or include decisions on the pace of change, and relationships with payment systems in the NHS (PbR and QoF)?
   c. To expand the membership of ACRA and TAG to include representatives of other services, such as a Chief Executive of a mental health care trust.
   d. To agree criteria for materiality of refinements to the formula.
   e. To enable ACRA and TAG to benefit from access to expert advice on specific issues such as labour economics.
   f. To consider formalizing a process of five-yearly reviews that begins with formal consultation with the NHS.
   g. For DH to publish:
      i. the rationale of the current formula and accounts of alternatives that ACRA has explored and why these have been rejected;
ii. annual reports on how ACRA’s recommendations have and have not been implemented; and
iii. reports on the impacts of resource allocation on inequalities in use of services and in health outcomes.

h. To widen ACRA’s remit so that ACRA could commission longer-term research in resource allocation and, in this way, increase the expertise available bidding for short-term contracts.
Review of the Weighted Capitation Formula

1. Introduction

1. The Advisory Committee on Resource Allocation (ACRA) commissioned this Report, on a review of the weighted capitation formula for revenue allocations to Primary Care Trusts (PCTs), for the Secretary of State for Health. The terms of reference were to: carry out a study, which will chart the work done, and the decisions made, by ACRA in relation to the review of the formula, prior to making revenue allocations for 2003/04 to 2005/06 and 2006/07 and 2007/08; examine current academic opinion and international best practice in this area; and produce a Report to ACRA in April 2007.

2. Since the Report was submitted to the Secretary of State, significant progress has been made in developing capitation formulas and new lines of research have been commissioned. This Report refers to, but does not review, these important developments, and has been revised to read as a retrospective review as of June 2008. One of the points made in this Report is the gap in publishing ACRA’s work; reports were published regularly until 2000, but then there was a gap until new material was published in 2008. One of the objectives of this Report is to fill that gap hence it has been written for a general readership.

3. The second section of this Report outlines: the objectives, implications and elements of weighted capitation formulas. The third section summarises official reports that describe developments of, and the rationale for changes in, methods of resource allocation between 1976 and 1999. (Since 1999, however, there has been no report from ACRA, although papers on elements of formulas were published until 2002). The fourth section of the Report describes ACRA’s work from 1999, and the rationale for changes to revenue allocations from 2003/04. The final two sections comment on the processes of ACRA and TAG; and offer conclusions and recommendations.

4. There are eleven Annexes. The first six give more details on the different elements in the development of formulas between 1976 and 1999: the population base; weighting for demography; weighting for additional need; allowances for unavoidable variations in costs; primary care and integrated formula; other issues (joint finance and the private sector). Annex 7 describes the formula as used from 1999/2000 to 2001/02. Annex 8 describes the innovation derived from ACRA’s work on the Health Inequality Adjustment (HIA), which was applied in 2001/02 and 2002/03. Annex 9, gives the evaluation criteria for resource allocation formulas used by ACRA. Annex 10 gives a list of abbreviations. Annex 11 gives the list of resource allocation papers on the Department of Health website: these include Resource Allocation Research Papers (the RARP series) and Resource Allocation Working Papers (the RAWP series).
5. My expertise in NHS resource allocation has been gained through various projects and membership of advisory groups over 30 years. I have been a member of research teams that reported on: the Management of Financial Resources in the National Health Service for the Royal Commission on the NHS (Perrin et al, 1978); studies for the Department of Health and Social Security (DHSS) of the medical service increment for teaching (SIFT) and a review of the literature on resource allocation for DHSS (Bevan, 1987; Mays and Bevan, 1987); problems of subregional resource allocation (Bevan and Spencer, 1984; Bevan and Brazier, 1987; Bevan 1988; Beech et al, 1990); and of setting budgets organised around general practices, including developing approaches for the Welsh Office in setting budgets for GP fundholders in Wales, and reviewing setting budgets for and managing risk by Total Purchasing Pilots (Baxter et al, 2000; Bachman and Bevan, 1996; Bevan, 1984, 1989, 1997, 1998; Bevan and McLeod, 2001). I have been a member of the Resource Allocation Working Group of the Welsh Office (1987 to 1991); the Medical Practices Committee of England and Wales, which used to approve applications for increases in the number of GPs in an area, or the replacement of GPs who had retired (1993 to 1994); and the Technical Advisory Group on Resource Allocation of the Department of Health (since 1997). I have reported on methods of resource allocation for the World Health Organisation (Bevan, 1991), the Department of Health and Social Services in Northern Ireland (Bevan and Sheldon, 1994), and for the Public Inquiry into children’s heart surgery at the Bristol Royal Infirmary (Bevan, 2001).

2. Objectives, elements and implications of capitation formulas

2.1. Objectives

6. Aneurin Bevan (1991), in a speech in the House of Commons, opening the debate on the second reading of the NHS Bill in 1946, identified what were then three problems of inequities of access to health care. First, access was influenced by ability to pay, which was tackled by access to the NHS being free at the point of delivery and financed by general taxation. Second, there was an inequitable distribution of General Practitioners, and Medical Practice Committees (MPCs) were created to tackle this problem, albeit with limited powers of regulation of where GPs could work through ‘negative direction’: MPCs were empowered to prevent GPs moving into ‘over-doctored’ areas, but could not direct GPs to work in ‘under-doctored’ areas (Webster, 1988). Third, there was an inequitable distribution of hospital services, about which little was done until 1976. From 1948 to 1974, hospitals were financed by a process of incremental budgeting; and exceptional arrangements were made for the elite teaching hospitals, so that they remained outside the state hierarchy of regional hospital boards and hospital management committees. These exceptional arrangements undermined attempts to promote equity through the programme of capital development and funding the revenue consequences of capital schemes (Bevan et al, 1980). The observation of

1 A fuller account of the history of GP distribution is given in RAWP 5
Julian Tudor Hart (1971) that the NHS operated an ‘inverse care law’, in which ‘the availability of good medical care tends to vary inversely with the need for it in the population served’, was justified by variations in ratios of supply to populations, reported in analyses by Cooper and Culyer (1971).

7. The crucial developments that tackled the problem of the inequitable distribution of supply of hospital services in the NHS were the 1974 reorganisation; and the introduction of a system of funding using formulas as recommended by the landmark Report of the Resource Allocation Working Party (The RAWP Report) (Department of Health and Social Security, 1976). As Smith (2007, p.12) observes their recommendations ‘took formula funding in the UK to a new level of intellectual coherence and sophistication, and have been highly influential internationally’.

8. Although there is consensus that the underlying purpose of publicly-financed health care is to improve ‘equity’, there is less clarity of what kinds of equity ought to be sought (Mooney, 1994). The RAWP Report interpreted the underlying objective of its terms of reference to be ‘to secure through resource allocation equal opportunity of access for people at equal risk’ and this was, until 1999, seen as the bedrock principle against which all methods of developing capitation formulas ought to be judged. The methods recommended by the RAWP Report, and used for resource allocation from 1977/78 to 2001/02, sought to tackle the inequities in the distribution of hospital services identified by Aneurin Bevan in 1946. These methods sought to equalise resource use per capita, taking account of relative need of populations and unavoidable differences in costs between providers. However, such methods do not, of themselves, correct problems of inequities in access caused by variations across groups, by age, social class, or ethnic group in seeking care when ill, or discrimination by providers on grounds other than clinical need. Hence there are two different kinds of inequities in access, which are, and are not, tackled by changes in the distribution of supply.

9. RAWP’s interpretation of their terms of reference defined the objective of formulas for over 20 years, but RAWP’s actual terms of reference required them to recommend methods to promote equity and efficiency. Members of the working party recognised that achieving equity in terms of expenditure per capita that took account of variations in risk only, would not achieve equity in access because of variations in costs per unit of service. These variations could arise from two reasons: unavoidable variations in costs, a problem which RAWP recognised ought to be taken into account in formulas; and

---

2 The crucial organisational change, that highlighted inequities in the distribution of supply in relation to populations and provided the basis for the development of a capitation formula to tackle this problem, was the 1974 reorganisation. This brought (undergraduate) teaching hospitals into the regional structure; and created new health authorities defined in terms of geographical areas, which were responsible both for planning health care for the population of, and running services within, those areas (Bevan et al, 1980). For a fuller account see Mays and Bevan (1987) and RAWP 4 <http://www.dh.gov.uk/en/Managingyourorganisation/Financeandplanning/Allocations/DH_4108515>.

3 Its terms of reference were: ‘To review the arrangements for distributing NHS capital and revenue to RHAs, AHAs and Districts respectively with a view to establishing as soon as is practicable, a method of distribution that is responsive objectively equitably and efficiently to relative need and to make recommendations’ (p. 5).
variations in efficiency, which RAWP saw as being tackled by other policy instruments (see below). Smith (2007, pp. 60 to 73) lays out a formal theoretical framework for analysis of issues underlying the use of formula funding.

10. In 1999 ACRA was asked by the Minister of State for Health (John Denham) to undertake a review to develop a new formula ‘to contribute to the reduction in avoidable health inequalities’, and hence a shift from seeking equity in access to reducing inequity in outcomes. But the outcome of ACRA’s work on this objective in terms of the formula was limited in the scale of funding and applied for two years only (see below).

2.2. Fundamental elements of formulas

11. There are four fundamental elements to weighted capitation formulas, which can be thought of as a four stage process:
   a. allocate resources on a equal per capita basis;
   b. take account of differences in demography (age and sex);
   c. take account of additional differences in need; and
   d. take account of unavoidable variations in costs.

12. These stages hence require estimates of: the number and characteristics of populations; weights for demography; weights for additional need; and adjustments for unavoidable variations in costs. Each stage is problematic because of limitations in data and formidable methodological problems. This is illustrated by examining what would be required to develop a normative basis of a formula for need only: i.e. ignoring difficulties over data on populations and accounting for unavoidable variations in costs.

13. A normative basis for a formula for need only would begin with information on the needs of populations. We lack basic epidemiological data on local rates of prevalence of chronic conditions (such as diabetes and osteoarthritis) and the incidence of acute conditions (such as strokes and acute myocardial infarctions). Even if such data were collected routinely this would not, of itself, give information on the care that was needed. To make a start on this we would also require detailed information on the degree of severity; but even if we had this detailed information we lack comprehensive evidence on cost-effectiveness to estimate what costs ought to be for all types of illness and treatments. It is hence, hardly surprising, that ACRA found (see below) developing a normative basis for resource allocation from the bottom up looked at best to be a long-term prospect. Furthermore, even if it were possible to develop a formula on a normative basis, this would create new sets of problems: as it is likely that the costs of what was ‘needed’ would exceed the available funding, and so explicit rationing decisions would be required that denied care to those who ‘needed’ it according to the formula.

---

4 This limitation in basic epidemiological information is powerfully illustrated by attempts to estimate the ‘avoidable’ burden of disease: see Hollinghurst et al (2000), Bevan et al (2007).
14. In the absence of a normative approach, those who have developed formulas have necessarily had to rely on empirical approaches to estimating weights for need; and also for adjustments for unavoidable costs. The stumbling block encountered by all empirical methods is seeking to estimate what ought to be from what is (Sheldon et al, 1993; Mays, 1995), so that we avoid perpetuating underlying inequities and inefficiencies. For example, successive formulas since the RAWP Report have taken account of demography by (essentially) using estimates of national average costs by demographic group, but patients may be denied the treatment they need because they are too old \(^5\) (Bowling, 1999). Whatever empirical approach is used to estimate variations in need relies on proxy (and not direct) estimates and is hence ‘essentially contested’: i.e. intrinsically open to dispute (Gallie (1955-6) quoted by Lukes (1974)). This is because the only way of resolving the dispute would be to produce direct evidence of relative need, but if we had that direct evidence we would not have to rely on proxies.

15. This section now considers the different elements in formulas and problems in their estimation.

16. *Estimates of populations* are required by age and sex, and other characteristics that capture additional need. There are two sources of data on the size of populations: estimates of geographically-defined areas by the Office of National Statistics (ONS) based on decennial censuses; and lists of patients registered with GP practices. There are good reasons why data from these different sources will differ (see paragraph 57). But, even after accounting for these known differences, estimates of populations on registered lists typically exceed those derived from censuses. Hence this is known as ‘list inflation’, this shorthand term is, however, misleading as it implies that only the registered lists are in error. The substantial differences between these two estimates when reconciled for accounting differences, have persisted in inner city areas where both estimates are likely to be in error, and have defied explanation.

17. From the start of the NHS in 1948, ‘list inflation’ was a problem in two crucial processes: paying general practices by capitation; and for Medical Practices Committees in deciding on whether or not to approve applications for numbers of GPs in an area. ‘List inflation’ was also a problem, from 1991, with the introduction of GP fundholding, which required setting budgets allocations for Hospital and Community Health Services (HCHS) at practice level; and from 1996, with the emphasis on a primary-care driven NHS (in allocations to PCTs and making estimates of allocations for practice-based commissioning). When ACRA was established in 1997 as the successor body to the Resource Allocation Group (RAG), it recognised the advantages of basing formulas on registered populations, but little was done to tackle the problem of ‘list inflation’. It is extraordinary that, ever since the start of the NHS, decisions have been made on the distribution and payment of GPs, and later on allocations of other resources, using problematic data on GPs’ registered lists;

\(^5\) It was well known around the time of the RAWP Report that there was rationing by age for access to renal failure (Challah et al, 1984).
and that there is still no single body that has responsibility for the accuracy of these data.

18. **Weighting for demography** matters as need obviously varies by age and sex. Weighting for age is vital as the demographic composition of areas varies enormously: with, e.g., high proportions of young families in new towns and of the elderly along the south coast. Weighting for sex is of marginal importance in formulas: although needs of the sexes differ, their composition varies little across areas. The RAWP Report recommended weighting for demography (age and sex) using national average rates of utilisation, which has been relatively uncontroversial and been followed ever since.

19. **Weighting for additional need**: Demographic data alone, however, do not capture variations in need. The approach of the RAWP report to accounting for additional need is characterised by elegant economy based on a heroic assumption: using (mainly) the ratio of actual to expected mortality in the form of Standardised Mortality Ratios (SMRs), and assuming that a health authority with e.g., 10 per cent higher (lower) mortality than expected ought to receive 10 per cent more (less) resources. This method was applied from 1977/78 to 1990/91.

20. There are good reasons for the RAWP Report using mortality data as the best proxy measure of additional need in populations. Mortality data are a good measure of the cumulative social and health experience of people living in an area and a sensitive indicator of general health care needs; reliable, accurate and routinely available and provide a more timely indicator of relative need than indices derived from census data; and independent of supply of, and access to, services (Mays, 1987; Mays and Bevan, 1987; Sheldon et al, 1993). But there are a host of questions in how to use mortality data in formulas for allocating resources:

   a. For which services is mortality a good measure of morbidity: is it only relevant for acute diseases?
   b. For what proportion of morbidity is there effective treatment and therefore ought resources be made available?
   c. Mortality is focused on morbidity: to what extent do mortality data capture extra needs for health care of deprived populations?

---

6 This account draws on work with Trevor Sheldon for the Northern Ireland Department of Health and Social Services.
7 The purpose of using SMRs is to measure variations in relative need. Bennett and Holland (1977) argued, for example, that there is little geographical variation in the incidence of chronic disease. What matters is whether the residual variation is or is not correlated with the SMR. We do not have data on the incidence of chronic disease. It is likely to be indicated by the census question on long-standing limiting illness. Carr-Hill et al, 1994a, for example, found that, in the under-75s, this was strongly correlated with the under-75 SMR (0.81).
8 Part of the problem here is confusion arising from assertions that formulas ought to take account of ‘deprivation’ (Mays, 1987). Obviously it is not a function of health care to remedy the causes of social deprivation, but the additional needs for health care that arise from these causes. (Which in turn forces the obvious question: Should extra funds be spent, for example, on health care needs from inadequate housing rather than making housing adequate?) As the RAWP Report pointed out, SMRs are correlated with all indicators of social deprivation. This has been a consistent finding of subsequent empirical research, Carstairs and Morris (1989), for example, found the under-65 SMR to be correlated in Scotland with census variables such as no car (0.74) and unemployment (0.73).
d. Where mortality is a good measure of morbidity, is the SMR an appropriate index?

e. If the SMR is an appropriate index, which age range should be used?

And if the SMR is based on a restricted age range (under 75), then should it, nonetheless, be applied to the whole population?

f. If the SMR is applied to a specified age range, what weight should be used for the SMR?

21. In 1977, the Government could implement the recommendations of the RAWP Report as providing a better basis for allocating resource than incremental budgeting. Over time, however, there were concerns over the troubling questions listed in the above paragraph, and one promising way of offering answers was to examine the empirical relationships between actual utilisation of services and indicators of morbidity and material deprivation. And following a study for Review of the RAWP Formula (RoR) (Royston et al, 1992), from 1991/92, the weightings that were used for additional need were derived empirical methods that related small-area analyses of variations in utilisation to indicators of need. Sheldon and Carr-Hill (1992, p. 412) in their critique of that inadequate attempt to estimate need from data on utilisation, observed that the RAWP Report ruled out such approaches because of the problem of disentangling hospital use from supply factors. All such approaches require assumptions, although these are less explicit than that made by the RAWP Report in its use of SMRs. Some assumptions are required because of serious limitations with data and severe methodological problems:

a. The ecological fallacy (see Smith, 2007, pp. 76 to 78) is when an effect detected at aggregate level of analysis does not exist at individual level. This is a serious problem as we can only link routinely-collected data on population characteristics to utilisation at the small-area level. The ‘ecological fallacy’ may mean, e.g., that in an inner-city area with high levels of unemployment and high rates of utilisation, the individuals who made heavy use of services were not affected by unemployment.

b. Data availability on utilisation of services: the best data we have on utilisation of services is for hospital admissions; it is difficult to derive indicators of need for other services, such as community-based care.

c. Interaction between risk and utilisation: as Grossman (1972) pointed out, there is a two-way interaction between sickness and use of health

---

9 The all-age SMR is the most well known index of mortality. There are others: age specific mortality indices, relative mortality index, Yerushalmy’s mortality index. Research has shown that use of different indices would significantly change authorities’ targets. Mays and Bevan (1987) argued that age-specific SMRs would provide more appropriate weighting than all-age SMRs. And Oliviera and Bevan (2003) found that, for Portugal that age-specific SMRs are preferable on conceptual grounds as they measure deviations on mortality rates against national mortality rates by age group and give equal weight for deaths in different age groups (unlike SMRs) and were found to be more strongly correlated with indicators of material deprivation and other mortality indices (years of life lost and the relative mortality index).

10 The all-age SMR is dominated by deaths in the very elderly the under-75 SMR is likely to be more sensitive to variations in underlying risk.

11 ACRA commissioned research in 2007 to analyse person-based utilisation, which offers a way to overcome the problem of the ecological fallacy.
care; sickness will have an impact on use of hospitals, and vice-versa. This requires use of techniques to estimate simultaneously each of these interactions. Sheldon and Carr-Hill (1992, p. 412) argued that a general weakness of any regression analysis is that ‘even the use of simultaneous modelling techniques cannot deal adequately with a system where demand, utilisation and supply are so inextricably intertwined’.

d. **Spatial interaction between supply and utilisation**: distance to hospital affects utilisation, and patients’ willingness to travel to hospital is influenced by the existing distribution of hospital services. Hence the relationship between distance and utilisation varies between inner-cities and sparsely-populated rural areas. Models that assume this relationship is the same across all areas will result in biased estimates of the impact of distance on utilisation.

e. **Medical practice variations and inappropriate treatment**: studies in the UK (and the US) have found correlations between variations in admission rates in the aggregate and generic social indicators. In contrast other studies of specific conditions and procedures have revealed variations in rates of treatment which, it is consistently argued, cannot be explained by differences in need and are attributed to variations in medical practice. (This could be seen as another example of the ecological fallacy.) Studies have also found high levels of inappropriate care. The evidence of variations in medical practice and levels of inappropriateness indicate that current practice can be very different from what ought to be provided.

f. **Relational distance** is a way of describing the fact that patients who are well educated and articulate are relationally close to the GPs and hospital doctors and tend to be more skilled users of ‘free’ services than those who are not. Hence data may show that those who are materially deprived do indeed make heavier use of services, but this may still be an underestimate of their need. The poor, for similar levels of illness, may make less use of health care than the rich do, so their actual utilisation may underestimate their need.

---

12 Techniques of simultaneous estimation had been deployed in US econometric studies, which sought to establish the impact of health care on mortality. See, for example, Auster et al (1969), Hadley (1982).
17 This term was developed by Black (1976).
18 Asthana et al (2004) cite examples from the literature of where deprived populations do and do not have significantly lower rates of health service use according to estimated need.
22. The York Study (Carr-Hill et al 1994a), which followed that for the RoR, also used empirical methods to estimate need from analyses of small-area variations in utilisation and was rightly admired for its technical sophistication, but was later criticised for its omission of variables that measured health (Asthana et al, 2004). The subsequent AREA study (Sutton et al, 2002) did include health variables. In practice different sets of estimates of weightings for additional need are strongly correlated with each other, but these different sets will produce different target allocations by PCTs. The AREA study was criticised by Stone and Galbraith (2006), who raised questions over the degree of variations in utilisation explained by their preferred models and that other models could have been chosen. Hence even experts applying state-of-the art techniques of estimation to analyse data on utilisation cannot produce uncontested weightings for additional need.

23. Stone and Galbraith (2006) argued for the normative approach, which was not a realistic option in the short term19. The only realistic alternative was to go back to RAWP’s approach of elegant economy based on obvious heroic assumptions as argued by Sheldon (1997):

The empirical work over the past years seems to have validated the original idea of the Resource Allocation Working Party to use a measure of the death rate as an indicator of relative need. There are no unique and valid indicators of health care need, and, no matter how sophisticated the analysis, research based on the use of services tends to underestimates the effect of poverty because the middle classes are better at accessing health services. As a result of this, a similar result could be produced by basing a formula simply on population size and age, weighted by the under 75 year standardised mortality ratio.

24. It is now possible to go further, as sound empirical work has not only validated the choice of SMR, but also suggested that RAWP’s heroic assumption of a weight of one for the SMR to have been sound. Although the flawed work for RoR estimated this weight to be 0.44, the York and AREA studies produced estimates of about 0.8 and 0.920. Figure 1 shows the scatter of observations for PCTs with the AREA index to be strongly correlated with the under-75 SMR. Figure 1 also shows two PCTs with similar weighting for additional need based on the AREA index (of 1.5) have very different SMRs: 1.30 (Easington) and 1.70 (North Manchester). It is a moot point as to whether we regard the AREA index or the under-75 SMR as the better indicator of relative need. It is hence plausible to offer a simple index for additional need based on the under-75 SMR with a weight of one as offering rough justice and transparency. The advantages of using mortality data were described above: they are accurate, reliable, available annually and applicable at PCT and practice level; and in comparison with relying on survey data, mortality data lack the complications of revisions from decennial censuses and the attribution problem for practices.

---

19 At the time this Report was drafted, ACRA’s experience with attempts to develop the bottom-up approach was disappointing (see paragraph 74). Work commissioned by ACRA in 2007, led by Sheena Asthana from the University of Plymouth, looks, however, to have promise.

20 Regressing the AREA index against the under-75 SMR gives 95% confidence interval estimates of: the coefficient on the SMR to be from 0.86 to 0.93 (the t statistic is 50.2); and the intercept to be from 0.071 to 0.142 (the t statistic is 5.9).
Unavoidable variations in costs

As the Resource Allocation Group (RAG) observed (NHS Executive 1996, p.15 - 16) a ‘£1 buys less health care in London and the South East than elsewhere’. This is because trusts have to pay the local market rate, which results in higher costs: from higher grading of posts; higher use of agency staff; and lower staff productivity, from high turnover and vacancies due to paying below the local market rate. The justification for including nursing and other specialist staff is not that of competition with other employers, but that ‘there is a proven recruitment and retention problem in and around London, suggesting that staff are mobile geographically within the NHS and the compensation of London weighting is insufficient to overcome the preference to live/work elsewhere’. There is a question, however, over whether medical and dental staff should be included in estimates of the staff MFF.

The main focus of effort in accounting for unavoidable variations in costs has been on developing the staff Market Forces Factor (MFF). There are two ways of estimating the effects of labour market forces on costs in the NHS:

a. The specific cost approach, by analysing actual variations in costs between NHS organisations in different areas; or
b. The general labour market approach, by analysing variations in costs in other labour markets.

Source: Department of Health (2005b)\textsuperscript{21}.

\textsuperscript{21} 2006-07 and 2007-08 Exposition Book Table 4a.5: (column E +1) for SMR, and column X for AREA index, www.dh.gov.uk/allocations
27. The problem with the specific cost approach is that although this captures variations in costs it does not indicate whether these are avoidable, and hence relying on this approach could perpetuate inefficiency and generate perverse incentives. So all estimates of the MFF have been derived from the general labour market approach. This is open to criticisms that general labour market indicators do not adequately proxy the additional (unavoidable) costs faced by NHS providers in different parts of the country and that NHS labour is less mobile geographically than other types of labour in different occupations. Another problem in using the general labour market approach has been the absence of data by travel-to-work areas: e.g., we know that the costs of labour are very high in the City of London, but how ought those high costs be reflected in allowing for market forces in the NHS? More generally, the MFF is estimated by zones with sharp differences in rates between zones (known as 'cliff edges') that lack face validity: the 1996 RAG Report (NHS Executive 1996, pp.16 -17) pointed out that neighbouring HAs with similar external comparator pay rates were assigned to different zones resulting in a net difference of 7 per cent in their allocations.

28. Initially the MFF covered the London and the South East only (see Annex 4). The 1996 RAG Report observed this was inadequate, as variations in labour markets also applied in the rest of England (NHS Executive 1996, p.15). But extending the MFF to cover the whole of England created another problem, because, as the 1997 RAG Report (NHS Executive 1997, p.11) highlighted, the MFF has an asymmetrical impact. It obviously compensates areas for the higher costs incurred where local rates of pay are higher than NHS national rates. In areas (in particular Cornwall) where local rates of pay were (and still are) lower than NHS national rates, the Report observed: ‘Whilst these HAs should enjoy some of the benefits e.g. better quality staff, this is unlikely to be sufficient to offset having allocations that reflect a going rate they cannot pay’.

29. RAG’s concern over the MFF is similar to that faced by the RAWP Report in deciding what weight to give to the SMR: we expect health service employers in areas with high labour costs to have to pay more for similar staff than employers in areas with low labour costs; but ought we simply give a weight of unity to the different indices for labour? We can see that if e.g., secretarial pay is 20 per cent higher in an area than the NHS rate, then this broadly reflects opportunity costs of trying to pay less; but if secretarial pay is 20 per cent lower than the NHS rate, does paying the NHS rate generate 20 per cent more benefits? In addition to this problem of asymmetry, there are complex questions over the impacts of general labour markets on nurses. RAG suggested that a possible approach to this problem would be to determine a cut-off point where NHS rates are deemed to equate to the going rate.

30. With the introduction of ‘Payment by Results’ (PbR) the MFF has become explicit: the MFF is now paid directly to providers to subsidise their costs, whereas before it was just one of a number of elements in the formula. Hence it is even more important that the MFF be seen to be based on good evidence and fair. Unfortunately this was not so. Both RAG (NHS Executive 1996, p.

22 A full account of the developments of the MFF and its criticisms is given in RAWP 1.
15) and ACRA (see below) have been concerned about the criticisms of (see e.g. the Health Select Committee below), and lack of confidence in, the MFF, which is often seen as a fudge that unfairly benefits some parts of the NHS.

31. There have also been concerns over the adequacy of allowing for the additional costs of rurality and sparsity, which raise three different issues:
   a. The measures of additional need from empirical analyses tend to reflect measures of urban rather than rural poverty\(^ {23}\);
   b. There may be extra costs from delivering ambulance and community-based care in sparsely-populated areas;
   c. The costs of the minimum quantum of certain types of care required by a sparsely-populated area may exceed that population’s fair level of funding.

32. Other causes of unavoidable variations in costs causes of excess costs, such as teaching, training and research, are financed outside capitation formulas, have not been considered by ACRA and are outside the remit of this Report\(^ {24}\). This also applies to the funding of specialised services (which have low patient numbers but need a critical mass of patients to make treatment centres cost effective\(^ {25}\)).

33. The purpose of this section has been to make it clear that after 30 years of work, we know for certain that there is no perfect formula for allocating resources. Formulas can only offer rough justice. ACRA has continued to grapple with the best ways of handling ‘list inflation’; providing evidence for estimates of additional need and unavoidable costs; and indeed what equity objective we ought to pursue in capitation formulas.

---

\(^{23}\) For example, car ownership is an indicator of relative affluence in urban areas, but in rural areas, owning a car may be a necessity rather than a luxury.

\(^{24}\) The RAWP Report recommended protection for the higher service costs of teaching medical and dental undergraduate students only through the Service Increment for Teaching (SIFT). The medical SIFT rate per student has been based on the estimated excess costs per medical student of medical teaching hospitals (over non-teaching hospitals); and dental SIFT rate per student on the estimated total costs per dental student of dental teaching hospitals. The estimates of medical SIFT have been obscure and their definition and scope have changed over time. The current rate is based on an estimate using data from the 1980s, which has been used to justify refinements in its distribution for which there is no sound evidence (Bevan, 1999). There have also been developments to finance costs of training and funding of research leading to an exercise to estimate costs of research through an accounting exercise, which is problematic. There are three problems in accounting for costs of teaching, training and research in capitation formulas. First, there is the problem of accounting for the costs of products that are jointly produced: the graphic example being attributing costs in rearing sheep to the production of mutton and wool. For the NHS this is much more problematic as the same people provide patient care, teaching, training and undertake research. Second, whilst main teaching hospitals as compared with other hospitals tend to be more expensive and do much more teaching of medical students, some training and research, which does not, of itself mean that the higher costs are justified by those differences. Particularly as there are large variations in costs of main teaching hospitals that are not explained by these differences or in the mix of patients they treat. Third, the extra costs of teaching hospitals are all pervasive: it is hard to relate extra funding to the activities this is supposed to finance and hence release that extra funding commensurately with changes in those activities (Perrin, 1987).

\(^{25}\) The arrangements for funding these services until the Bristol case were reviewed by Bevan (2001). A recent report from Sir David Carter has resulted in new arrangements.
2.3. Using formulas to make allocations

34. The RAWP Report not only laid out the key elements of capitation formulas, but also how these should be used in making allocations. The Report made the crucial distinction between the target, as an eventual aim, and each year’s allocation, which was made so as to move actual allocations towards their targets. The guidance on resource allocation for 2006/07 to 2007/08 outlines the following four elements used to set PCTs’ actual allocations as follows (Department of Health, 2005, p. 18):
   a. weighted capitation targets;
   b. recurrent baselines (the actual current allocation which PCTs receive);
   c. distance from target (DFT), the difference between (a) and (b) - if (a) is greater (less) than (b), a PCT is said to be under (over) target; and
   d. pace of change policy, which determines the level of increase which all PCTs get to deliver on national and local priorities and the level of extra resources to under target PCTs to move them closer to their weighted capitation targets – this policy is decided by Ministers for each allocations round.

2.4. Allocations and budget setting

35. RAWP methods as developed and applied from 1977/78 to 1990/91 were applied within a financial framework of hierarchical budget setting that did not offer a way of promoting efficiency (Bevan and Robinson, 2005). To provide incentives for efficiency and improvements in quality, it was necessary to develop other policy instruments, and for the NHS these are provided by separation between purchasers and providers with two systems of payment to providers: ‘Payment by Results’ (PbR) for a variety of providers of hospital care, with different contractual obligations; the Quality and Outcomes Framework (QoF) for primary care. The problem with this more complex financial framework is that each PCT’s income (determined by capitation) will not automatically equal payments under PbR and QoF. It is not even clear that, if a PCT’s volumes of cases were at utilisation rates that were consistent with its capitation target, the PCT would be able to afford its payments under PbR. Nor will a PCT’s allocation be increased in line with increases in volumes and quality of hospital and primary care services delivered. The composition of targets does not provide a normative basis of how money should be spent, even though there have been elements within formulas for specific services.

36. There have recently been criticisms that the formula used for resource allocation was a cause of financial deficits in the NHS or long waiting times. This is discussed further below. The fundamental point to be made here is that as deficits are a failure to match expenditure to income, they can never be good grounds for increasing incomes. Seeking to fund demand or supply regardless of relative need is in effect to abandon the principles of equity that have governed resource allocation to the NHS since 1977 (Bevan, 1998).

37. It is up to PCTs to manage all demands for payments from their single source of income. Hence purchasers need to manage demand and design contracts
with providers so that payments can be afforded. An obvious model for PCTs of managing demand for secondary care is that of the Health Maintenance Organisation (HMO), with GPs as gatekeepers managing referrals to secondary care (Robinson and Steiner 1998). A lesson from the Total Purchasing Pilot is that involving GPs in demand management is easier at the level of individual practices but that creates problems in managing risk for small populations (Martin et al, 1997; Baxter et al, 2000).

3. The development of capitation formulas in the NHS


38. The RAWP report was produced within a year and recommended a simple formula to guide allocations to 14 Regional Health Authorities (RHAs) driven by three elements only: resident populations; weighting for demography by estimated national average costs; and weighting for additional needs by SMRs. Following the RAWP Report, there was a limited review of RAWP methods: the Report of the Advisory Group on Resource Allocation (the AGRA Report: Department of Health and Social Security, 1980). This review was curtailed, recommended minimal changes, but provided the rationale for two innovations: funding of multi-regional clinical services and developing the MFF.

39. The first significant review, by the NHS Management Board, the Review of the Resource Allocation Working Party Formula, produced interim and final Reports (RoR: Department of Health and Social Security, 1986 and 1988). This was the first official Review that based the weighting for additional need (using data from the 1981 census) on empirical analysis of data on small-area variations in utilization. This recommended: changing the index of the SMR by reducing the age range from all-age to under-75 and using across all conditions rather than condition-specific; reducing the weight on the SMR to 0.44; and including a measure of ‘deprivation’. The change made from

26 The impact of year-to-year variations in risk has been examined for all inpatient admissions (Martin et al, 1997) and for rare costly referrals (Bachmann and Bevan, 1996). Results of each show that these variations can be unmanageably large for annual budgets for populations of 10 000 or less. But, once the budget is extended to about 30 000 person years, the variation is significantly reduced. This does not, however, mean that a practice ought to have to have a population of at least 30 000: a practice of 10 000 ought to be able to manage year-to-year variations in expenditure, provided that it is allowed a three year period over which to manage these. In turn this means that, for small practices, underspends in any given year ought to be carried forward to provide insurance against future overspends.

27 Its terms of reference were: ‘To consider detailed improvements in the methodology recommended by RAWP as new data and the result of research studies become available; and to advise on any practicable and desirable changes which could improve the process of resource allocation for the hospital and community health services’ (p. 1).

28 Its original terms of reference were: ‘To report to Ministers by the end of 1986 with proposals for improving the way in which the national RAWP formula measures relative need for the hospital and community health services, taking account of consultation with the NHS and related interests such as professional bodies and medical and teaching interests, of work carried out on modifications to the RAWP formula at sub-regional level, and of any relevant new statistical analysis. The Report should include recommendations as to the timing of any proposed changes (paragraph 1).’

29 such as the UPA8 score after investigation of its stability at regional level
1991/92 was to use the new SMR index with a weight of 0.5 (Secretaries of State for Health for Health, Wales, Northern Ireland and Scotland, 1989) with no measure of ‘deprivation’. Unfortunately, the research undertaken that justified the reduced weight on the SMR was methodologically flawed (Mays, 1989; Sheldon and Carr-Hill, 1992) 30. As a result, regions with high additional need (as measured by the SMR) were underfunded from 1990 to 1995.

40. In 1993, with the availability of new data from the 1991 census, the Department of Health and Social Security invited a competitive tender for a new empirical analysis to estimate a new weighting for additional need. Its primary purpose was: ‘to improve the sensitivity of the current formula for allocating Hospital and Community Health Service (HCHS) funds to Regional Health Authorities’31 (Carr-Hill et al, 1994a, p. 1). A team from York University were awarded the contract for this work and produced an impressive report based on state-of-the art techniques to derive estimates from small-area variations in utilization (Carr-Hill et al, 1994a)32: this was well described by as Ken Judge, in evidence to the Health Select Committee (1996a, p. xv) as being ‘widely acknowledged to be the most impressive and sophisticated undertaken so far in this field’. The research team reported to two advisory groups: ‘a technical group, comprising academics and technical experts from the NHS; and a steering group, drawn principally from the NHS, which was concerned with issues of policy and implementation’ (Carr-Hill et al, 1997). Exceptionally, however, there was no report published from these steering groups. The only accounts are from the research team’s description of the models they developed; the Department’s description of how the policy of using results of these models for allocations for 1995/96; and criticisms of that policy by the researchers (Carr-Hill et al, 1997), and the Health Select Committee of the House of Commons (Health Select Committee, 1996a, p. xviii).

41. The York study (Carr-Hill et al. 1994a) developed two different models: for acute and psychiatric inpatient services and produced an index for each33. A

---

30 This estimate was derived by using the method of Ordinary Least Squares. There is a two-way interaction between measures of morbidity and utilisation of health care: a small area with relatively high levels of morbidity is likely to have relatively high levels of utilisation; but we also expect those relatively high levels of utilisation to reduce the level of morbidity from what it would otherwise have been. Hence simply regressing small area variations in utilisation on SMR is likely to produce an under estimates of the coefficient: to derive unbiased estimates it is necessary to model both effects simultaneously.

31 The secondary objectives were to: address the possibility of developing a national formula for general practitioner fundholding procedures; assist the four Thames Regional Health Authorities produce an acceptable pan-Thames formula; investigate the relationship between age and sex and the use of hospital beds; determine some of the substitution and complementary relationships between different types of care (Carr-Hill et al, 1994a, p. 1). The main report (Carr-Hill et al, 1994a) covered the last two. Another report by Carr-Hill et al (1994c) explored the possibilities of deriving a capitation formula for guiding the setting of GP fundholder budgets.

32 This Report was summarised in two papers in the British Medical Journal: Carr-Hill et al (1994b) and Smith et al (1994).

33 This did take account of the interaction between need and utilisation and also by using the technique of two stage least squares regression analysis; and used the technique of multi-level modelling to
problem in resource allocation formulas is the lack of data on community services. The convention is to use weighting for additional needs as applied to hospital services to community services. The Government argued that, as the York study derived weighting for additional needs for hospital services only, this justified not applying those weightings to most community and other health services (which accounted for 24 per cent of HCHS expenditure)\(^\text{34}\). Applying the results of the York study consistently across all services would have resulted in a weighting of additional need of approximately 0.8 on SMRs (about half-way between the weights of 1 used from 1977/78 to 1990/91; and 0.5 used following from 1991/92 to 1994/95)\(^\text{35}\). Applying the weightings to 76 per cent of HCHS only (64% acute and 12% psychiatric) limited the scale of the redistribution to areas with material deprivation and was criticized for favouring Conservative constituencies\(^\text{36}\). The Department’s responses were to justify this practice as an interim measure and to commission research into estimating additional needs for community health services (Health Select Committee, 1996a, p. xx).

42. The Health Select Committee (1996b) took evidence in 1995. The establishment of the Resource Allocation Group (RAG)\(^\text{37}\) and Technical Advisory Group (TAG) in 1995 look to be responses to the lacunae in governance identified by criticisms by academics and the Health Select Committee. The tasks for RAG included: consideration of the 24 per cent of expenditure omitted from weighting for additional needs; the report from the study of community services by a team based at the Universities of Plymouth and Kent (Buckingham et al, 1996); and the MFF, which the Health Select Committee (1996a, p xxxi) stated ‘in its present form is deeply flawed’. RAG produced two reports on its work for 1995 and 1996 (NHS Executive, 1996a); and for 1997 (NHS Executive, 1997).

43. Using results of a study of community services by a team based at the Universities of Plymouth and Kent (Buckingham et al, 1996) RAG (NHS Executive 1997, pp. 7 - 10) recommended weightings for all HCHS\(^\text{38}\). These

control for ‘district’ effects on small areas (i.e. where a district’s policy has an effect on utilisation of services in small areas in that district).

34 The details of this practice were reported by the RAG Report for 1996 (NHS Executive 1996, pp. 10 –13). The extra services covered were: acute and psychiatric outpatients and daypatients; maternity and geriatric inpatients; maternity and psychiatric community health services; ambulances. RAG recommended not weighting the following services that accounted for 24%, for additional need (with expenditure as a % of HCHS in parentheses): community services (other than maternity and psychiatric) (11.41%); people with learning disabilities (5.25%); administration (7.83%).

35 The best fit to the York results in terms of the under 75 SMR was obtained with a weight of 0.8.

36 And analysis for the Health Select Committee (1996a) showed the impact of applying weights fully. There would have been losses of 10% or more in target allocations for: East Surrey; West Surrey; Cambridge and Huntingdon; and West Sussex. There would have been gains of 10% or more for Newcastle and North Tyneside; Gateshead and South Tyneside; St Helens and Knowsley; Sunderland; Lambeth, Southwark and Lewisham; Camden and Islington; East London and the City; Liverpool; and Manchester.

37 RAG was established to provide a systematic way of involving the NHS in the development of policy and practice of resource allocation with its main objectives being: to establish a work programme for developing a coherent and coordinated approach to new HA allocations; and to develop financial frameworks for both standard and new forms of fundholding.

38 The changes made were: for community services, technical changes to the age/cost curve, and using six indices derived by the study by Buckingham et al (1996) for: district nurses, health visitors,
laid the basis for the formula for allocations for HCHS for 1999/2000 for HAs and PCGs. There were different indices for four services (with their percentage share of HCHS in parentheses) (NHS Executive, 1999): acute (70.22%) and psychiatric (14.54%) hospital services, based on results from the York Study; and non-psychiatric (11.87%) and psychiatric community services (3.34%), based on results from the Plymouth / Kent Study.

44. RAG was replaced by the Advisory Committee on Resource Allocation (ACRA)\(^39\), which was established in September 1997, and produced a report on its work for 1998 (NHS Executive, 1998). After that the Department of Health produced annual Reports on Resource allocation each year from 1999 with the last report being published in 2005 (Department of Health, 2005).

45. Other issues and adjustments to formulas within the scope of this Report\(^40\) that have been considered after the RAWP report of 1976 are:
   b. *Inner cities*: The second RAG Report (NHS Executive, 1997) considered three issues posed by the needs of inner-city populations, ethnicity (people with English language difficulties), mentally disordered offenders, the homeless and rough sleepers\(^41\). The Report recommended taking account of homelessness and rough sleepers. The ACRA Report (NHS Executive, 1998) recommended an English Language Difficulties Adjustment (ELDA) for interpreter, advocacy and translation services for inner cities with ethnic populations.
   c. *AIDS and Drugs misuse*: The first RAG Report (NHS Executive, 1996) was the first official report to consider special allocations for these services; and recommended these special allocations be continued (subject to changes in the method of funding).

---

\(^39\) ACRA’s terms of reference (p. 5) were: to advise the Secretary of State for Health on the distribution of resources across primary and secondary care, in support of the goal of equitable access to healthcare for all; and to develop and apply methods which are as objective and needs-based as available data and techniques permit.

\(^40\) As explained above, this Report does not consider allowances for teaching, research and training.

\(^41\) The rough sleepers adjustment was a supplement to the formula introduced for 1996/97 allocations. A monetary adjustment was added to the targets of the 30 HAs that had the highest proportion of rough sleepers in their population (Department of Health, 2003).
d. **Joint finance**: was considered by both RAG Reports (NHS Executive, 1996 and 1997) but not by the ACRA Report (NHS Executive, 1998).

e. **Private sector**: The ACRA Report (NHS Executive, 1998) considered the impact of private medical expenditure on in determining equitable allocations but recommended no change to take account of this in the capitation formula.

### 3.2. Changing formulas for NHS reorganisations

46. Official Reports on Resource Allocation from 1976 to 1988 were designed for making allocations by the Department (of Health and Social Security and then of Health) to 14 RHAs. These reports also offered guidance for RHAs in making allocations to subregional Health Authorities (initially Areas and later Districts), which were responsible for running HCHS in their areas, and their populations needs for these services. Primary care was administered separately: initially by Family Practitioner Committees and later Family Health Service Authorities. Hence the various reports on resource allocation over that period did not consider capitation formulas for funding primary care (other than making some allowance for the administration costs of Family Practitioner Committees). This meant that the Department’s capitation formulas:

a. were designed for making allocations to RHAs;
b. were limited to coverage of hospital and community health services (HCHS); and
c. had to take account of cross-boundary flows of patients who lived in one health authority and were treated in another (and were designed to fund providers’ catchment populations (Bevan and Ingram, 1987)).

47. From 1991, the introduction of the internal market stripped District Health Authorities (DHAs) of their provider functions, so that, from 1991, they were funded on the basis of their resident populations (and contracted with providers) and hence formulas no longer had to take account of cross-boundary flows (Secretaries of State for Health for Health, Wales, Northern Ireland and Scotland, 1989).

48. From 1991, DH was also responsible for giving guidance to Districts on making allocations to the many general practices that had opted for various forms of fundholding (NHS Executive, 1996b). So capitation formulas had to be developed for practice populations and the subset of HCHS included in fundholding. Applying capitation formulas to practice populations raised two

---

42 The aim of joint finance as stated by 1996 RAG Report (p. 19) had become facilitating the continuing shift from hospital to community care. From 1989/90 to 1996/97 these funds had been allocated to RHAs based on last year’s expenditure, and each RHAs made allocations to DHAs using their own methods. For 1996/97 DH allocated these funds to DHAs based on last year’s expenditures. RAG (1997, p.16) proposed a formula for allocating joint finance, introduced in 1997/98, which used four groups weighted for need ; and adjusted for MFF.
problems: the unknown reliability of estimates at practice level in areas with high ‘list inflation’\(^{43}\) and the ‘attribution problem’ (see below).

49. From 1 April 1996, the NHS was reorganised so that DH was responsible for making allocations directly to DHAs, this removed the exercise of local choices in both the development of formulas and decisions over the pace of change in moving allocations towards targets. DHAs were made responsible for primary care. This and later reorganisations have sought to develop the general practice as the building block of the NHS, which had two implications for the formula: increasing the scope of services included and reducing the scale of populations. Much of the work of RAG, TAG and ACRA has been generated by adapting formulas to these changes. Formulas for funding primary care were developed to cover two elements from 1996/97 (see Annex 5): prescribing; and Discretionary General Medical Services (practice staff, premises improvements, and computers)\(^{44}\).

50. Although GP fundholding was abolished, the emphasis on practice-based commissioning raises many of the problems of using a capitation formula for fundholding, and has a direct antecedent in the extension of fundholding through the Total Purchasing Pilot (Bevan, 1997; Bevan and McLeod, 2001). Applying formulas to general practices makes their application more problematic for three reasons:

a. The small scale of practice populations: it is much easier to derive a sound estimate of relative need for a population of a RHA (with a population of over 3 million) than at the level of a general practice. Once we accept that any capitation formula will only offer rough justice, we can regard an RHA’s target as a good estimate with small margins of error in estimation. But, at practice level, a target will be unreliable and have large margins of error: we can expect estimates for a third of practices to have errors of estimation greater than 10 per cent (Martin et al, 1997, p.11). Bevan (1987) argued that: ‘At the level of single practices, there is likely to be limited gains from hoping to explain differences in risk.…. At this level, however, clinical information on hospital admissions can be used to assess whether, for example, high admission rates are likely to reflect increased risk or medical practice styles’.

b. ‘List inflation’: we know the scale of unresolved discrepancies between two estimates of populations at the level of the PCT but not at the practice. A recent Report to ACRA has estimated the difference from using these two estimates to be over £50m in the target for two London PCTs. In a PCT with high ‘list inflation’ this is likely to vary

\(^{43}\) Although this would have not been much of a problem for GP fundholders as these practices tended to be in the affluent suburbs.

\(^{44}\) Discretionary GMS is also known as General Medical Service Cash Limited (GMSCL): when cash limits were introduced on public expenditure from 1976, certain services were deemed to be demand-based, and therefore exempted from fixed cash limits. Most General Medical Services provided by GPs and their prescribing costs were classed as being demand-based. The inclusion in the global formula of non-discretionary GMS (payments to GPs for reimbursement of their costs, and their incomes as determined by the contract) was considered by ACRA in 2001 (see paragraph 89 below).
across practices and may be concentrated in a few but ‘list inflation’ is not known at practice level. It is not possible to derive sound estimates using formulas at the level of practices without resolving adequately the problem of ‘list inflation’.

c. The ‘attribution problem’ arises in weighting variables that use data from censuses for practice populations where the populations in wards are registered in different practices. The convention is to use weights of the proportion of each ward in each practice (this is, e.g., how Jarman scores used to be calculated for deprivation payments to GPs), which assumes that risks in ward populations are attributed at random. If this assumption does not hold then serious errors can arise as it has been consistently found that most expensive five per cent of the population account for 60 per cent of total expenditure on health services. This is illustrated by Figure 2, derived from hospital expenditure in the Netherlands (Lamers and van Vliet, 1995). Figure 2 gives the percentages of the populations that account for the top one, five, ten and 25 per cent of expenditure; and what the (much lower) estimated expenditure would have been based on using average spend by age and sex (which are the only data directly available at practice level). To take an extreme case for illustrative purposes, assuming that these results applied to England, if a ward (of 5 000 people) were served by two practices, and one ended up with all the high risk one per cent (50 people), accounting for age and sex would mean that it would receive two per cent of the ward’s share of resources, but be faced with need for 30 per cent of those resources.

45 The finding of the skewed distribution of expenditure is common. Scheffler (1989) reported US data showing that the ranking the elderly by spend, the top 7.7% of the accounted for over 71% of total expenditure. Matsaganis and Glennerster (1994) reported that the top 5% in a fundholding practice accounted for 68.4% of total expenditure.

46 The normal problems associated with the skewed distribution of expenditures shown in Figure 2 are that of ‘cream-skimming’: i.e. insurers will try to exclude the high-risk population. It has been argued, however, that it is very difficult to identify who these people are, and hence there are limits to what those who aim to practice cream skimming can achieve. The problem, which Scheffler (1989) identified, is rather different: that practices may end up with a high-risk population, which is not identified by routinely-reported data: e.g., using national methods of risk rating by ward. He alleged that this would be the Achilles heel of GP budget holding. Scheffler argued that the problem caused by a high-risk population would be acute with small populations, and implied that using larger populations would solve the problem. See Scheffler (1989), Matsaganis and Glennerster (1994), van de Ven (1994).
3.3. Developments from 1999

51. *The new NHS* (Secretary of State for Health, 1997) committed the Government to ‘put in place new mechanisms to distribute NHS cash more fairly’, so that ‘the healthcare needs of populations, including the impact of deprivation, will be the driving force in determining where cash goes’ with ‘a new formula to set fair shares for the new Primary Care Groups’. In November 1998, the Minister of State for Health (Alan Milburn) announced a wide ranging review of the formula used to make allocations to Health Authorities and Primary Care Groups/Trusts, which was followed by a letter of March 1999 from the next Minister of State for Health (John Denham) to the chair of ACRA. The main objective for this review was ‘to contribute to the reduction in avoidable health inequalities’. The Minister also asked ACRA to cover three principal concerns about the fairness of the current formula:

a. problem of moving from basing population estimates based on censuses to registered lists;
b. accounting for unavoidable costs; and

c. accounting for additional need.

52. The review was carried out under the auspices of ACRA. From 1999/2000, the bases of the unified formula were frozen until 2001/02 except for routine changes to update data (NHS Executive, 1999). Annex 7 describes the basis of the formula for 1999/2000.

53. In June 1999, ACRA considered papers that set the scene for the wide-ranging review of capitation formulas in response to the letter from the Minister of State. These papers summarized comments from ACRA members, and sought to define the scope for this objective. Two main ways for taking the work forward emerged.

a. A ‘radical’ approach which would start from scratch and attempt to produce a single measure as a health inequalities adjustment (HIA) to be applied at HA and PCG level. This approach was recognised to be problematic: much of NHS activity provides a sickness service and is not related to reducing health inequalities; and there was no readily available (evidence based) measure of health (or health inequality) that could be applied in a robust fashion.

b. A ‘supplementary’ approach in which there would be a core allocation for the ‘sickness service’, and various supplements for specific service areas for interventions that were known to be effective in reducing health inequalities.

54. In September 1999, ACRA organized a forum for members of TAG and ACRA to discuss these approaches and agreed on the ‘supplementary approach’. An outcome of this work was the introduction of an interim HIA for 2001/02 and 2002/03 only. For 2002/03, £88m was distributed to the 50 HAs with the highest YLL index in addition to £60m targeted funding for Health Action Zones carried forward from the previous year. The HIA was superseded by a new formula from 2003/04 but the allocations for the HIA and Health Action Zones were incorporated in the baseline allocations of the 50 HAs. Annex 8 describes ACRA’s work in support of the HIA.

55. The terms of reference for this Report include charting the work done, and the decisions made, by ACRA in relation to the review of the formula, prior to making revenue allocations for 2003/04 to 2005/06 and 2006/07 and 2007/08. This Report has aimed to put this review in context in terms of the objectives and limitations of formulas; intrinsic problems and developments since the RAWP Report based on official reports that are publicly available that cover the period from 1976 to 1999. Table 1 outlines changes made to the formulas since 1999. Figure 3 shows how Distances From Targets (DFTs) increased sharply in 2003/04 following the replacement of HAs by PCTs and the introduction of a revised formula based on the AREA Report (Sutton et al, 2002) (which makes it difficult to identify the separate impacts of either change).
56. The following sections examines ACRA’s work programme in relation to revenue allocations for 2003/04 to 2005/06 and 2006/07 and 2007/08 on:
   a. The population base (including rough sleepers)
   b. Accounting for need for HCHS
   c. Accounting for unavoidable costs for HCHS
   d. Primary care and integrated formula
   e. Other matters

Figure 3: Range in % Distance from Target from 1998/99 to 2007/08

Source: Department of Health (2007)
Table 1: Changes to the weighted capitation formula since 1999

<table>
<thead>
<tr>
<th>Year</th>
<th>Unified formula</th>
<th>Population base</th>
<th>Need</th>
<th>Unavoidable costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999/2000 to 2000/01</td>
<td>A unified formula was introduced covering HCHS, prescribing and discretionary GMS.</td>
<td></td>
<td></td>
<td>An English Language Difficulties Adjustment (ELDA) was introduced as a supplement to the formula, for the extra costs of interpretation, advocacy and translation services.</td>
</tr>
<tr>
<td>2001/02</td>
<td></td>
<td>An interim health inequalities adjustment (HIA) was introduced</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002/03</td>
<td>Extended to include non-discretionary GMS and funding for HIV/AIDS treatment and care and HIV prevention(^{47})</td>
<td></td>
<td>The adjustment in the new non-discretionary GMS component replaced the existing additional need adjustment in the discretionary GMS component</td>
<td>Staff MFF was revised</td>
</tr>
</tbody>
</table>

\(^{47}\) The inclusion of GMSNCL was promulgated in *The NHS Plan* (Secretary of State for Health, 2000). Funding for HIV/AIDS treatment and care and HIV prevention each had had its own formula and allocations policy.
Table 1: Changes to the weighted capitation formula since 1999 (continued)

<table>
<thead>
<tr>
<th>Year</th>
<th>Unified formula</th>
<th>Population base</th>
<th>Need</th>
<th>Unavoidable costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003-04 to 2005-06</td>
<td></td>
<td>2001 census</td>
<td>New need adjustments were introduced in the HCHS and prescribing components (replaced both the existing formula and the HIA)</td>
<td>Staff MFF was revised</td>
</tr>
<tr>
<td>2006-07 to 2007-08</td>
<td></td>
<td>(ONS) population projections; rough sleepers adjustment dropped; and a growth area adjustment introduced(^{48})</td>
<td>Primary medical services component was introduced into the formula(^{49})</td>
<td>The HCHS staff MFF was reviewed to support the implementation of ‘Payment by Results’(^{50})</td>
</tr>
</tbody>
</table>

Sources:
Department of Health (2003) for 2003/04 to 2005/06; and describes the HIA and the introduction of non-discretionary GMS and funding for HIV/AIDS.
Department of Health (2005a) for 2006-07 to 2007-08.

\(^{48}\) This was in support of the Department for Communities and Local Government sustainable communities initiative
\(^{49}\) This followed the devolution of this funding to PCTs in 2004-05, and replaced the non-discretionary and discretionary GMS components.
\(^{50}\) The main resulting change was an increase in the number of zones in the staff MFF from 119 to 303 to match the geography of PCTs.
4. ACRA’s Work programme 1999 to 2005

4.1. Population base

57. ACRA (NHSE Executive, 1999) had to recommend the best way of estimating the population base for PCGs, which were to become PCTs. The Report recommended: continuing to use population projections for resident populations and adjusting these for ‘fringe’ patients (who were registered with a GP in a different HA/PCG from where the patients lived). The Report saw using registered populations as preferable in the longer term and recommended cleaning practice lists for this purpose. Registered populations exceeded ONS estimates in all regions, except East Anglia. These excesses were greatest in London and the South East; and concentrated in the elderly and males aged between 25 and 44.

58. The Report highlighted three explanations for ‘list inflation’:
   a. prisoners and armed forces personnel are included in ONS estimates and entitled to use HCHS; but not registered with GPs and do not make use of their services;
   b. inaccuracies in registered populations caused by inadequacies in administration and recording of transfers (with large numbers of patients being double counted); and
   c. time lags between the two systems in taking account of births and deaths.

59. ACRA’s subsequent work on the Population Base consisted of two main streams:
   a. the problems of ‘list inflation’; and
   b. the availability of data from the 2001 census.

60. ‘List inflation’. ACRA established a Population Base Group (PBG) in 1999 to make recommendations to reduce list inflation, ensure that it remained at minimal levels and improve the accuracy of GP and health authority data at the point of registration. (Unregistered populations would be counted for resource allocation purposes.) The Group reported in 2000 and Ministers were consulted Ministers on the report.

61. For revenue allocations post 2007/08, ACRA agreed that formulas ought to be based on the best and most robust population base using all available data sources. This would include: consideration of unregistered populations and specific population groups (e.g. asylum seekers), links with a wide range of organisations responsible for data collections, and links with wider policy issues (e.g. practice based commissioning51) and review work being carried out on the need and cost elements of the formula. A working group was

---

51 There may be a tension between work on finding the most robust population base for PCT allocations and the population base that will be used for practice based commissioning e.g. GP registered lists. This may result in the allocations to PCTs being made on a different basis to how the funding is allocated to practices.
established (chaired by Dave Roberts) to take this approach forward\textsuperscript{52} and reported in 2007.

62. \textit{Availability of data from the 2001 census}. ACRA considered problems arising from the 2001 census in making allocations for: 2003/04 and 2004/05, and post 2007/08. The issues included: the availability of estimates based on these censuses, the need to make adjustments, that three of the eight Jarman components (social class, overcrowding and ethnicity) would not be derivable from the 2001 census consistently with current definitions. The outcomes were:

\begin{itemize}
  \item a. to use the best available ONS estimates for PCTs for the year of allocations\textsuperscript{53} and allocating additional funds to the four growth areas identified by the Office of the Deputy Prime Minister (ODPM)\textsuperscript{54}, and
  \item b. to question the adjustment for ‘rough sleepers’\textsuperscript{55}.
\end{itemize}

63. As ONS do not produce population projections by PCTs the Department has developed a method of deriving these projections using ONS projections for the 354 lower-tier Local Authorities in England\textsuperscript{56}. For allocations for 2003/04 to 2005/06, the Department used 2001 population estimates and retained the adjustment for ‘rough sleepers’\textsuperscript{57} (Department of Health, 2003). For allocations for 2006/07 and 2007/08 the Department used ONS 2003-based long-term sub-national population projections and dropped the adjustment for ‘rough sleepers’ (Department of Health, 2005).

\section*{4.2. Weighting for need for HCHS}

64. In June 1999, ACRA recognized that the utilisation-based approach appeared the most appropriate way of producing target allocations for the core element. In September 1999, ACRA identified the following phased research programme to update the formula for that ‘core element’:

\begin{itemize}
  \item a. Within two years, to develop an updated ‘small-area’ utilisation based model for general services as a means of ‘benchmarking’ an individual-level based model (see (c) below)\textsuperscript{58}.
\end{itemize}

---

\textsuperscript{52} Its tasks included: defining the responsible population; specifying requirements of a population base for Practice-Based Commissioning (PBC); description of current population bases and data sources; description of potential data sources; impact assessment of new policy initiatives (e.g. PCT mergers); assessment of process to support data provision (mainly about ‘future proofing’ data sources); specifications for sample data sets; Commissioning of analysis and data modelling.

\textsuperscript{53} The 2003 mid-year population estimates available in August 2004, took account of adjustments resulting from the Manchester and Westminster matching exercises.

\textsuperscript{54} For 2004/05 and 2005/06, an additional \textsterling 20 million of revenue funding was allocated and for 2005/06 an extra \textsterling 20 million of capital funding was made available.

\textsuperscript{55} This was based on an estimate of 8600 persons, however the latest estimate from the Rough Sleepers Unit stood at 532 persons. ACRA were informed that there was no obvious way to assign estimates to PCTs, and discussed the proposal that the adjustment should be dropped from the formula given concerns over its materiality.

\textsuperscript{56} These were chosen because of their geographical coterminosity with PCTs and are: London boroughs, unitary authorities, metropolitan and non-metropolitan district councils.

\textsuperscript{57} for PCTs within the 30 former HAs which received the adjustment in 2002/03

\textsuperscript{58} It was recommended that this work ought to: be based on most recent cost-weighted HES data available; exploit non-census based ‘explanatory’ variables (e.g. social security benefits, unemployment rates, etc); pay regard to how to account for differential access (e.g., ‘unmet need’ and
b. Within a year, to undertake further analysis of how ‘age’ relates to need for ‘core’ health care services; 

65. In December 2000, ACRA considered an interim report (which had been considered by TAG) on the ‘bottom up’ general service module scoping study commissioned from York University. The report outlined alternative approaches to bottom up data collection: many existing surveys were not fully representative, and even where they were they raised the problems of how to translate results into cost data, the alternative of designing a special survey was feasible but costly. So, given the timescale of the ‘bottom up’ research programme, a fallback ‘top down’ approach was adopted.

66. In April 2005, at a joint meeting of ACRA and TAG, a paper giving a stocktake of resource allocation was considered, and a wide range of issues were identified: links to Departmental policies; new money allocated to PCTs outside of the formula to PCTs; reviewing/updating the additional needs element of the formula; GMS global sum; a bottom-up approach to resource allocation; practice based commissioning; rurality; asylum seekers; prison health care; community pharmacy services; general dental services.

67. ACRA’s principal streams of work on additional need were:
   a. further empirical analysis of the weights to be used in the capitation formula, which continues the ‘top-down’ approach; and
   b. exploration of an alternative ‘bottom-up’ approach.

68. The following two sections describe each of these streams of work. The view taken in April 2005 on other work areas was as follows:
   a. General dental services were identified as the priority amongst the other work areas.
   b. Resources for community pharmacy services could be allocated through the current formula.
   c. The formula was insufficient in terms of identifying needs of those in prison and asylum seekers. Each required a supplement to the formula.
   d. Rurality was seen as an issue of unavoidable costs rather than additional need.

---

59 key research issues were likely to be: accounting for births, deaths and premature mortality, availability of data, and costing.

60 This would be based on a survey of NHS utilisation, and/or relevant practice based data. For resource allocation purposes such utilisation data would need to be costed (or weighted) and then mapped back to routinely held data at HA, PCG, and practice level.

61 It had the biggest budget and offered the greatest challenge because dental need is not correlated to the current needs formula. A scoping paper linked to the policy objectives is required to enable ACRA and TAG members to understand fully the issues.
69. Empirical analysis of the weights in the capitation formula. Since 1995/96 the weighted capitation formula had been based on various empirical analyses of hospital and community health services (see Annex 3). A new approach, using additional variables on health, had been developed in revising the formula used in Scotland, and offered a way of capturing need that was unmet for reasons other than variations in access to supply: unmet need was captured for ethnicity, for example, by allowing this to have a negative effect on utilisation and a positive effect on morbidity.

70. In 2001, ACRA tendered for a new small-area study to estimate weights in the capitation formula from empirical analysis including variables on states of health (using data from the Health Survey for England\(^63\)) as used in work on revising the formula used in Scotland. The objectives included (Sutton et al, 2002, p.9):

a. updating and replacing the indices of additional needs for HCHS and prescribing;
b. investigating the possibility of constructing a single index of need at PCT level;
c. investigating alternative methodologies to replace the age-related components of the existing formulas;
d. proposing a methodology to adjust for specific unmet need; ensuring that formulas were updateable by using non-census data and data from the 1991 census which was consistent with the 2001 census\(^64\); and
e. ensuring that the proposed need indices were robust at PCT level.

71. The tender was awarded to the AREA team from University of Glasgow/ISD Scotland. They developed three models for: acute and maternity services; mental health services; and prescribing\(^65\). The final Report (the AREA report: Sutton et al, 2002) recommended a new set of weights (see Appendix 1) which:

a. used a new age cost curve based on health resource groups (HRGs)\(^66\) (and attributed costs of births to the age of the mother); and

---

\(^{62}\) This issue had been visited a number of times and the conclusion was always the same. It was therefore agreed that there was no need to revisit this issue (with the proviso that the formula now extended into primary medical services, and primary medical services are substituted for secondary care in some rural areas).

\(^{63}\) These are household data, and the model requires making the assumption that institutional populations mirror the wider population in the community.

\(^{64}\) As detailed results would not be available from the 2001 census until 2003.

\(^{65}\) There were five important differences from the York study, in the proposed development of the AREA system of weighting: using a single index for socio-economic characteristics (the Index of Multiple Deprivation); using non-census variables; incorporating age and additional need in an integrated way (a one-stage approach); aggregating across programmes to produce one single set of weights; measuring access using the use of the ‘distance to 5 NHS providers’ variable. Of these proposed differences the only differences in outcomes of the two studies were the AREA team’s use of non-census variables and the introduction of additional variables in seeking to control for differences in access.

\(^{66}\) The researchers were unable to develop a ‘one-stage’ approach for assessing the affect of age and additional need on utilization; and hence retained the traditional two-stage approach. The age cost curve within the York formula was based on average costs for 12 broad specialty groups.
b. captured aspects of unmet need not previously identified\textsuperscript{67} using data that could be more easily updated.

72. ACRA considered issues over: the adequacy of weightings for community care and mental health; and presentation of the new element for additional need. ACRA’s recommendations and the outcomes for revenue allocations to PCTs for 2003/04 to 2005/06 (Department of Health, 2003) and 2006/07 to 2007/08 (Department of Health, 2005) were as follows:

\textit{a.} ACRA recommended using the new method for constructing the age cost curve, and this was accepted by Ministers.

\textit{b.} ACRA considered that the AREA work presented good evidence of unmet need for certain groups, and represented a step forward. The outcomes were that three new need indices were accepted: for prescribing, acute and maternity HCHS, and mental health HCHS\textsuperscript{68}.

\textit{c.} ACRA questioned whether HIA should be part of future allocations given that the AREA approach offered a new way of capturing unmet need. ACRA recognised that the HIA attempted to address the social gradient in health, so there was a good case for keeping HIA conceptually distinct\textsuperscript{69}. ACRA recommended further testing to look at the impact of predicted prevalence on the model and to assess the need for a separate HIA. The HIA was not maintained as a separate allocation, but, to protect those PCTs who had benefited from that, the existing 2002/03 allocation of £148m was added to those PCTs’ baselines.

73. In 2004, ACRA agreed on work to update fully the three sets of estimates of the AREA report using data from the 2001 census: which had not been available in time for the AREA study. This work sought to be consistent with the original AREA formula, as only the needs variables would be updated. Unfortunately, however, some variables included in the AREA weights could not be updated as the necessary data were no longer available. Hence it was agreed to substitute different but similar variables and minimise the potential problem of inconsistency by staying as close as possible to the three AREA models\textsuperscript{70}. Due to the problems associated with data availability and variable consistency it was, however, decided not to use the updated estimates for the 2006/07 and 2007/08 allocations.

---

\textsuperscript{67} E.g., where ethnic groups make use of lower levels of health care services than other groups but have higher levels of morbidity. It was noted that the new formula for Scotland might include a similar morbidity element.

\textsuperscript{68} The AREA team could not produce a well-specified model for Community Health Services (CHS). The report recommended that the acute and maternity model be used to determine weighted capitation for all the CHS programmes other than for community mental health, where the mental health model was recommended. This was accepted.

\textsuperscript{69} The HIA used an index based on Years of Life Lost (YLLs) see Annex 8 for a description of the HIA.

\textsuperscript{70} For the Acute and Maternity services model, replacing the Education domain score variable with the Skills sub domain variable and excluding the insignificant musculoskeletal morbidity index variable resulted in some changes to the coefficients and a model with slightly better explanatory power. For the Mental Health services model, there had been a small change to the coefficients and a slight increase in explanatory power. There had been no change to the prescribing model.
74. Following the joint meeting of ACRA and TAG, in April 2005, members agreed on a twin-track approach. One track was the AREA study (Sutton et al, 2002), which revisited the issues of age and proximity to death and unmet need within the current method. The second track was to explore a bottom-up approach, which is the subject of the next section.

75. **Exploration of a ‘bottom-up’ approach.** At the joint meeting of ACRA and TAG, in April 2005, members identified limitations of the top-down approach to resource allocation based on small area variations in utilization: it replicates the status quo and is thus inconsistent with policies that seek to change this (such as reducing health inequalities, National Service Frameworks, and chronic disease management in the community). It was suggested that developments in policy might make the alternative bottom-up approach necessary.

76. At the joint meeting of ACRA and TAG, in November 2005, Members considered the results of exploratory analysis of the bottom-up approach by the Department. This encountered a number of serious problems including:
   a. the quality and limitations of QoF data on morbidity;
   b. translating data on morbidity into estimates of use of services; and
   c. translating estimates of use of services into costs.

77. ACRA members identified problems in developing the bottom-up approach for different diseases:
   a. ACRA members suggested that rather than carrying out further research on disease areas that are already well covered, e.g. Coronary Heart Disease (CHD), diabetes, and asthma, that the work should be done on weaker areas. But these weaker areas (such as Schizophrenia) turned out not to be appropriate because of current data limitations.
   b. For CHD, affluent (materially deprived) patients might be more (less) inclined to see their GP, be prescribed statins and receive elective care. This means that, for CHD, materially-deprived patients are likely to have higher mortality, lower prevalence, require more emergency treatment and a more expensive model of healthcare.

77. Whilst the top-down approach based on current patterns of expenditure creates tension with policies that aim to change service delivery, ACRA recognized that, even if the bottom-up approach were feasible, applying expenditure weights based on best practice is no guarantee that the resources would be spent that way, because allocations are not hypothecated. A fair summary of exploratory work until 2007 on the bottom-up approach is that ACRA had a much better understanding of its difficulties than a clear direction for its

---

71 Such as, the proposed shift from hospital to primary care, NSFs and chronic disease management, and practice based commissioning) and feasible (such as the Quality Management and Analysis System (QMAS), which supports the Quality and Outcomes Framework (QOF) and the Electronic Patient Record.

72 lack of reporting by age, sex and severity; biases in reporting between different areas

73 Having cost weights attached to some conditions would take resources away from other areas, but applying best practice to all conditions may result in totals that exceed the NHS budget.
development. Hence ACRA’s emphasis on a twin-track approach, which has continued to develop. ACRA recommended, in March 2008, a new formula based on the top-down approach based on small area analysis and a new HIA. ACRA commissioned in 2007, two studies to explore innovative approaches; from a team led by Sheena Asthana at the University of Plymouth to develop a formula from the bottom up; and from a team led by Jennifer Dixon at the Nuffield Trust to develop a formula based on analyses of utilization by individuals.

4.3. Unavoidable costs for HCHS

78. ACRA discussed the two main methods for adjusting for unavoidable cost variations, the specific cost approach and the general labour market approach, and accepted that the review of the staff MFF should use the latter. ACRA discussed different staff groups included within the MFF (which excluded medical and dental staff\(^\text{74}\)). The review would need to reconsider the inclusion of such groups with the impact of the new consultants’ contract.

79. ACRA commissioned two studies of the MFF from the Institute of Employment Research at Warwick University (Institute for Employment Research, 2001; Davies and Owen, 2004) based on the general labour market (GLM) approach. Institute for Employment Research (2001) reviewed the MFF and was presented to ACRA for information only.

80. The outcomes for revenue allocations for 2003/04 to 2005/06 were to (Department of Health, 2003):
   a. make no change to methods used for the funding the Emergency Ambulance Cost Adjustment (EACA) and the English Language Difficulties Adjustment (ELDA);
   b. for the staff MFF, to base this on the results of the Warwick study (Institute for Employment Research, 2001); include medical and dental staff; remove the cut off in the MFF, which meant reductions in target allocations for areas with MFF below national average costs. (This was not discussed by ACRA; and RAG had argued for the cut off when the MFF was extended to the whole of England\(^\text{75}\)).

81. In 2004 the decisions to remove the cut-off to the MFF and the inclusion of medical and dental staff were considered by TAG and ACRA and agreed to be correct\(^\text{76}\).

82. TAG and ACRA:

---

\(^{74}\) This was based on two factors that work to offset the impact of market forces on these staff: the effect of training places for junior doctors and the opportunities for private practice for consultants.

\(^{75}\) A paper to ACRA reporting on this change justified it as follows: ‘Undamping the MFF would recognise that even if there are national pay scales then in low cost areas the NHS is benefiting from a surplus of labour at the wage rate being offered. This leads to lower turnover and some likely productivity gain relative to higher cost areas. Undamping the MFF would give health economies greater financial scope to make use of the extra flexibilities in the new pay system to award recruitment and retention premia’.

\(^{76}\) Removing the cut-off, and including medical and dental staff had a material effect on a small number of trusts, however, the trusts that lost out, still had costs below tariff.
a. made recommendations for the design of MFF for integrating primary care into a single unified allocation (see below);
b. considered the differences between two methods of covering higher costs as defined by the MFF, 40 per cent was through the weighted capitation formula, which applied to targets only and 60 per cent was through ‘Payment by Results’ (PbR), in which changes to costs would be passed on in full.
c. made recommendations for weighting the land index for Trusts with more than one site on the basis of activity (as recommended by TAG).

83. ACRA considered the second study of the staff MFF by the Institute of Employment Research (Davies and Owen, 2004), in 2004 and agreed in principle: to increase the number of MFF zones from 119 to 303 (treating each PCT as a different zone); but not to change the existing methodology to smoothing ‘cliff edges’ (which was based on distance and population)\(^\text{77}\). This study was used in deriving the MFF for allocations for 2006/07 and 2007/08 (Department of Health, 2005).

84. In 2005, TAG and ACRA reviewed the work programme on unavoidable costs, and prioritized another review of the MFF, and agreed that the work on both the Specific Cost Approach and the GLM approach should be carried out through competitive tenders. Two studies were commissioned from the Health Economics Research Unit (HERU) at the University of Aberdeen and a consortium comprising Crystal Blue Consulting, the Centre for Health Economics York and the City Health Economics.

85. HERU were commissioned to: provide an update on the existing GLM approach to the MFF (research by Institute for Employment Research (IER) at Warwick, 2004); specifically consider seasonality and job responsibility, and look into the possibility of using different comparator groups for different NHS staff groups. ACRA considered recommendations from the HERU study. ACRA agreed to technical recommendations that improved the accuracy of estimates\(^\text{78}\). Their impact was to narrow slightly the range of unsmoothed Standardised Spatial Wage Differentials (SSWDs). ACRA agreed on implementing two findings from this study: not to include an adjustment for seasonality\(^\text{79}\); and to include an adjustment for responsibility. The effect was to reduce the MFF for Inner London by 3% and distribute this around the rest of the country. (TAG saw this as being sound but raised a question of materiality\(^\text{80}\).)

86. ACRA recommended further exploration of three issues.

\(^{77}\) The rejected alternatives were merging of neighbouring PCTs, using variations on the existing smoothing techniques, and taking account of where Trusts are located by interpolation.

\(^{78}\) These were the introduction of weights, pooling, effects coding and inclusion of part-timers.

\(^{79}\) The study’s estimates seasonality were based on small numbers and not statistically significant.

\(^{80}\) ACRA agreed with the recommendation to make the adjustment at regional level on the basis that it had been recommended by HERU and agreed by TAG. Although the adjustment does not have a material effect, it was part of a package of adjustments that have been calculated and are available for use. And hence did not set a precedent for what is ‘material’ in resource allocation terms.
a. ACRA accepted TAG’s recommendation to include observations for city workers\textsuperscript{81}, who had previously been taken out of SSWD calculations because they were seen as atypical of the general labour market, having different levels of job responsibility. This made a material difference to the estimated SSWDs for City and Hackney PCT, but ACRA agreed that the adjustment for the impact of the City should be made for the whole of London. ACRA asked for more detail on the rationale for the inclusion of city workers and the extent to which their different responsibilities are accounted for by the job responsibility adjustment.

b. Whether to exclude doctors from the MFF adjustment\textsuperscript{82}.

c. Exploration of options for ‘smoothing’ (to remove ‘cliff edges’). ACRA agreed that the best method would be based on Travel to Work Areas, but unfortunately that was not possible with the data that were then available\textsuperscript{83}. ACRA asked for three options to be evaluated: baseline with and without acute trust interpolation (the current technique), and the exponentially distance decayed smoothing and interpolation (as recommended by HERU). ACRA agreed criteria to govern the choice of the preferred option: it should reduce the current MFF gradient and the scale of cliff edges\textsuperscript{84}; be technically justifiable, explainable and defensible.

87. The Specific Cost Approach led by Crystal Blue Consulting comprised three strands

a. Analysis of national data sets from the Healthcare Commission, which showed that trusts in high-MFF areas were both less productive and had higher costs per Full-Time Equivalent (FTE); but it was not possible to identify the direction of causation.

b. A micro study of 14 trusts, which found that the payroll dataset was of poor quality (with numerous recording errors, and a large number of categories). Analysis of these data suggested that trusts were underpaying with respect to the MFF. Analysis of the general ledger also showed trusts in high-MFF areas were both less productive and had higher cost per FTE; and found that spending on medical and nursing staff accounted for two-thirds of trust budgets. The issue here is that the MFF enables trusts to pay, e.g. nurses more than the national rates, but that does not happen with consequent problems of recruitment,

\textsuperscript{81}TAG’s view was: that GLM method should not exclude workers arbitrarily; City workers were the only highly-paid workers excluded from the GLM; and the question of their different responsibility ought to be handled by the new adjustment for that.

\textsuperscript{82}Both HERU and Crystal Blue had found that doctors are different. HERU have proposed that doctors are excluded. TAG, however, thought insufficient evidence was used to justify this proposal (declared 3 month vacancy rates).

\textsuperscript{83}Once the TTWA\textsuperscript{s} were updated using the ONS work, this option would be explored to see if it were a possible option in the short or medium-term.

\textsuperscript{84}HERU and Crystal Blue Consortium both found that the current MFF gradient was too steep, and the ‘cliff edges’ were too sharp.
retention and heavy use of agency staff. Analyses by Hall et al (2008) suggest that one consequence of this is higher mortality rates for patients admitted with Acute Myocardial Infarction (AMI).86

a. an econometric study showed that London trusts were smaller and hence failed fully to exploit economies of scale.

86. ACRA recommended a number of changes to how the MFF should be calculated in March 2008. These were: to add an adjustment for job responsibility; to include City of London workers in the calculation; to not apply the staff MFF to spend on doctors; to smooth the raw MFFs for PCTs and providers.

4.4. Primary care and integrated formula

89. In 1976-77, the Government introduced a system of cash limits on public expenditure that applied to HCHS but exempted ‘demand-led’ family health services. This included the remuneration and expenses of GPs, the costs of prescriptions written by them, together with all other pharmaceutical, dental and ophthalmic service costs. This created three components of expenditure on Family Health Services: costs of prescribing, and of General Medical Services (GMS), which were and were not cash-limited. GMS cash-limited (GMSCL) was also known as ‘Discretionary’ and GMS non-cash-limited (GMSNCL) was also known as ‘Non-Discretionary’. The three elements composed the ‘Global Sum’. In this Report GMSCL is called Discretionary GMS and GMSNCL called Non-Discretionary GMS. By 1999, prescribing costs were cash-limited but Non-Discretionary GMS was still exempt.

90. From 1999/2000 to 2001/02, two formulas were used for funding Discretionary GMS and for prescribing (see Annex 7). In anticipation of legislation, that abolished the Medical Practices Committee and gave Health Authorities the power to decide whether a GMS GP vacancy existed, a letter from the Minister of State for Health (John Hutton), in November 2000, asked ACRA to consider the development of a single formula for the whole of GMS for the purpose of target setting. This formula would cover both the GMS Discretionary and Non-Discretionary Components (the latter would remain exempt from cash-limits85). TAG was asked to consider key analytical issues for Discretionary GMS and to report back in summer 2001. A paper by Smith raised a number of issues that should be taken forward in the longer term. The membership of TAG was changed to include a GP who had been a member of the MPC. ACRA had considered a report in 1999 from a joint working group with the Medical Practices Committee (MPC), which had recommended a formula developed by Carr-Hill for setting fair share targets for GPs at HA level. There were three principal issues:

85 The focus of the revised formula would be on the distribution of growth monies in the existing cash-limited unified budget. In future Health Authorities with spend on the GMS Non-Discretionary Component above their ‘fair share’ on the basis of the revised funding formula would receive less growth in their unified budget allocation than they would have under the existing formula (their GMS Discretionary Component would however remain exempt from cash limits although the total budget for primary care is cash limited).
a. Whether to use the additional needs weighting developed by Carr-Hill for setting fair share targets for GPs at HA level to cover both Non-Discretionary and Discretionary GMS\textsuperscript{86}.

b. How to take account of those elements that are geographically skewed: such as payments for dispensing\textsuperscript{87} and ‘deprivation’\textsuperscript{88}.

c. How to take account of unavoidable cost variations\textsuperscript{88}.

91. ACRA considered a Report from the University of York (Carr-Hill and Rice, 2001) for a formula for Non-Discretionary GMS, which was based on limiting long-term illness and a combination of area level ‘deprivation’ variables (Jarman score) and derived from analyses of GHS data. ACRA accepted that this approach would represent a robust basis for a formula for the short term\textsuperscript{89}.

92. ACRA also considered a Report from the University of York for a workload formula for the new GMS contract. A number of issues were raised: defining core services to which the new formula would be applied; modelling at practice level might not pick up the effect of deprivation on consultation rates in estimating additional needs; the range and impact of factors on unavoidable costs would vary considerably across the UK (e.g., remoteness was likely to be an important issue for parts of Scotland and Wales); and dual registration would result in double counting of the population.

93. ACRA considered the application of the current MFF to Non-Discretionary GMS. TAG had endorsed two main conclusions: that the staff and premises MFF from the cash limited budget should be applied to non-cash limited expenditure (except that the staff MFF should not be applied to GP pay). In addition it was acknowledged that London has areas that have high living costs and are unattractive as places to work as GPs. TAG had recommended using specific adjustments rather than the MFF. ACRA’s views were:

a. for practice staff and premises, further work be carried out with a view to applying the existing MFF indices. GP pay was not the only factor in determining where GPs locate. Other factors were cited, including the availability of locums, where GPs were disinclined to do out-of-hours work, plus employment opportunities for partners. Although it would be difficult to influence where GPs decide to locate on a purely financial basis, ACRA accepted that resources would be absorbed in remedying the problem of under doctored areas.

---

\textsuperscript{86} And hence replace the existing GMS Discretionary formula. Both the new and old formulas were based on analysis of GP consultation rates from the General Household Survey (GHS).

\textsuperscript{87} Dispensing payments made to GPs in market towns and rural areas were for additional work, and were funded from the non-cash-limited pharmaceutical budget.

\textsuperscript{88} The market forces factor embedded in the GMS Discretionary formula takes account of the attractiveness of an area, and capital values of properties. For the GMS Non-Discretionary component, in areas of population mobility, high turnover of patients adds to GPs’ workload.

\textsuperscript{89} ACRA had, however, two reservations on the use of the GHS. First, the exclusion of high cost populations because they are institutionalized (TAG had raised similar concerns). Second, using consultations as a proxy for need was problematic. This was because up to 40% of all consultations were initiated by practices (and comprised screening and prevention programmes with increasing numbers seen by practice nurses which was not taken into account by GHS).
b. the development of interim specific adjustments was the next step, which required identifying data on the factors that influence GP location.

94. ACRA considered the need to produce a more robust MFF adjustment for Non-Discretionary GMS for use from 2003/04\(^{90}\). ACRA concluded that the impact on unified allocations (for HCHS and primary care), was very small and agreed not to commission further work at this time. It was acknowledged that any modelling or additional analysis to inform contract negotiations could be commissioned separately.

95. ACRA’s recommendations were accepted and the outcomes for the formulas used in making revenue allocations to PCTs for 2003/04 to 2005/06 (Department of Health, 2003 and 2005) were as follows:
   a. Population: 2001 population estimates from the 2001 census\(^91\);
   b. Prescribing: needs weightings for demography were based on analysis of total prescription cost data of 120 English practices over a one-year period; and for additional need from the AREA Report.
   c. Discretionary and Non-Discretionary GMS: needs weightings for demography were mainly based on Morbidity Statistics from General Practice - 4th National Study (MSGP4)\(^92\); for additional need from the study by Carr-Hill and Rice (2001); and for unavoidable variations in costs mainly from the HCHS staff index for the Discretionary component and compensation for the higher costs of recruitment and retention of GPs in deprived areas in the Non-Discretionary component\(^93\).

96. In 2004, ACRA considered two options for a formula to integrate primary care into a single unified allocation (ACRA (2004)01)\(^94\):
   a. Using the formula developed by Carr-Hill and Rice (2001) for setting GMS discretionary targets, and had been agreed by ACRA / TAG, but would need development to take account of market forces.
   b. Using the Global Sum allocation (payments) formula, which was based on analyses of data from the Health Survey for England (for

---

\(^{90}\) TAG had recommended the utilisation model based on practice expenditure should not be taken forward at present. TAG noted that the level of complexity involved in modelling a GP pay market forces factor in this way had been underestimated. TAG had requested preparation of a comparison of target allocations both with, and without, the GP pay MFF adjustment.

\(^91\) Adjusted to exclude prisoners, members of the UK armed forces, and members of foreign armed forces and their dependants.

\(^92\) The age/time weight therefore also makes use of estimates of average consultation times by International Classification of Diseases (ICD) chapter and a weighting for home visits derived from the General Medical Practitioners Workload Survey 1992/93: Joint Evidence to the Doctors and Dentists Review Body from the Health Departments of Great Britain and the General Medical Services Committee.

\(^93\) For full details see Department of Health (2003, pp. 37-38 and 41-42).

\(^94\) A third option of commissioning a new formula for primary care allocations was not feasible in the short term.
1998/2000), which was similar to the formula used for GMS Discretionary and Non-Discretionary components\textsuperscript{95}.

97. ACRA raised four sets of issues that needed to be resolved in developing an integrated formula:
   a. What population base to use?\textsuperscript{96}
   b. Formulas need to cover primary care other than GMS: pharmacy and dental services;
   c. Accounting for needs of the prison population; and
   d. Accounting for high-cost service providers.

98. ACRA considered two reports on a formula to integrate primary care into a single unified allocation and agreed with TAG’s recommendations that elements of the formula would be as follows:
   a. Demography: the formula would weight for age and sex using estimates from the General Practices Research Database (GPRD) with an adjustment for nursing homes\textsuperscript{97}.
   b. Additional need: to use the indicators from the global sum formula\textsuperscript{98}.
   c. Unavoidable costs: the staff index would be based on the same MFF as used in the HCHS element of the formula; the land and buildings index would continue to be calculated in the same way; the GP pay adjustment would be updated using the ID2004 data and Inland Revenue weights.

99. ACRA agreed with TAG that the new registrations and rurality adjustments in the global sum formula addressed unavoidable costs rather than need. TAG sought guidance from ACRA on what constituted a truly unavoidable cost. ACRA agreed that such issues as new registrations and rurality ought to be looked at in the context of the whole formula, and not form part of the new formula for primary care funding.

100. ACRA’s recommendation were accepted and the outcomes for the formulas used in making revenue allocations to PCTs for 2006/07 to 2007/08 (Department of Health, 2005) were as follows:
   b. Prescribing: as for 2003/04 to 2005/06 (see paragraph 96 above).

---

\textsuperscript{95} This was based on registered lists, and included an age/sex element with adjustments for rurality and nursing homes, which may not be appropriate for resource allocation purposes.
\textsuperscript{96} Using registered lists would reward those with high ‘list inflation’ and would result in inconsistency, with different populations being used in different parts of the formula. Using ONS estimates as a constraint on registered lists would disadvantage areas where the census undercounted populations.
\textsuperscript{97} GPRD was considered better than MSGP4. Although GPRD included nursing home residents, their consultation time would have been under-counted. The adjustment for these consultations was made by taking the proportion of those in care homes and updating the length of consultation time to the average length of consultation time (which is 10 minutes). (The main effect was on the over-85 year olds and would have an extremely small impact on Unified Allocations.)
\textsuperscript{98} The Jarman index, which had been used in GMSCL and GMSNCL formulas, was based on indicators derived from the 1991 census and could not be derived in the same way from the 2001 census.
\textsuperscript{99} Adjusted to exclude prisoners, members of the UK armed forces, and members of foreign armed forces and their dependants.
c. **Demography**: the formula would weight for age and sex mainly based on estimates from the GPRD with an adjustment for nursing homes;

d. **Additional need**: to use the Global Sum allocation (payments) formula;

e. **Unavoidable costs**: as for 2003/04 to 2005/06 (see paragraph 96 above).

101. ACRA considered papers on Practice Based Commissioning and Practice Based Budgets. These papers laid out what was within and outside the scope of ACRA’s and TAG’s work. It was outside ACRA’s remit to advise on issues of pace of change, risk pooling and the financial management of budgets. The usual ACRA/TAG process would provide advice on:

a. the development of an appropriate methodology for determining ‘fair shares’ of PCTs resources for indicative budgets at practice level; and

b. the extent to which that methodology achieves an equitable distribution of resources, based on need within a PCT.

102. ACRA considered a toolkit for calculating HCHS weighted capitation targets at practice level to support practice based commissioning that used the same national formulas used to calculate PCT allocations. ACRA questioned whether a formula that was designed for PCT allocations could also work at practice level. ACRA pointed out the dangers in having a proliferation of formulas at a local level, and agreed that the best approach would be to have a locally applied national formula, but that it should not be applied prescriptively.

103. ACRA considered a report that set out a new approach to setting practice budgets through a research project to investigate the development of ‘patient categorisation’ according to ‘health risk’ characteristics. This information could be used in resource allocation by the same principle as weighted capitation and to address the more ambitious agenda of assigning a resource to patient categories and the costing of pathways. ACRA identified a number of problems in developing such an approach with problems over data on individual risk, categorisation of risk (whether it is based on need or doctor behaviour); and translating risk scores into resources. And if it were possible to produce a formula based on ‘patient categorisation’, it would be necessary to decide how to reconcile practice lists to ONS estimates and handle the transition from moving from one formula to another.

### 4.5. Other matters

104. DH provided ACRA with briefing papers on:

a. The methods used by the Department for Environment, Transport and the Regions (DETR) for allocating the central grant to local authorities and the element for personal social services and DETRs’ Index of Multiple Deprivation.

---

100 ACRA agreed that there should be the same local discretion that currently exists for prescribing.

101 This would attempt to identify individual characteristics, for example, whether the individual is a smoker, their body mass index, to create a risk score. The individual risks scores could be aggregated to provide practice-level fair share allocations based on key individual characteristics of patients on the GP list.
b. linking performance improvement to resource allocation.
c. the key resource allocation issues arising from the NHS Plan.
d. NHS allocations in relation to rural and disadvantaged areas.
e. The work of the Capital Issues Group in developing the formula for capital allocations at PCT level\textsuperscript{102}.
f. overlap between PCT revenue allocations and Payment by Results.

105. ACRA asked DH to give information:
   a. In 2004, on the impacts at PCT level of mainstreaming of HIV/AIDS treatment and care and HIV prevention within 2002/03 allocations\textsuperscript{103}.
   b. In 2005, on the impact that ACRA’s work has had on the distribution of resources in relation to equity, and the impact on health and health outcomes\textsuperscript{104}.
   c. In 2006, on what amount of variation was down to the different element of the formula - age, need, MFF.

106. ACRA agreed to publication in 1999 and 2000 of papers on\textsuperscript{105}:
   a. \textit{A Brief History of Resource Allocation in the NHS 1948-98} (RAWP 4);
   a. \textit{A History of GP Distribution} (RAWP 5);

107. After 2000, there were no reports to ACRA of publications on resource allocation. DH publishes descriptions of methods of resource allocation in the series \textit{Resource Allocation: Weighted Capitation Formula}\textsuperscript{106}. Despite this, in May 2006, ACRA was concerned over the lack of clarity and understanding (in the NHS) about current allocations and suggested that the formula may need to be better explained to the NHS. For example, members were interested in what amount of variation is down to the different element of the formula - age, need, MFF, etc. ACRA saw the need for information to be published:
   a. In 2004, on ‘Payment by Results’ (PbR).

\textsuperscript{102} As some issues might have been referred to ACRA, it was agreed that it would be useful to create an ACRA subgroup with expertise on capital and to have a member join the Capital Issues Group.
\textsuperscript{103} ACRA expressed concern that a small number of PCTs would be significantly affected by mainstreaming of HIV/AIDS funding due to localised growth in the number of patients requiring HIV treatment. ACRA requested that further modelling on HIV/AIDS covering the next few years should be undertaken at PCT level, based on projections of the growth of the HIV population.
\textsuperscript{104} The results of preliminary analysis only were reported showing (as expected) that there was a slightly higher allocation per head for those areas with a high need index and a slightly lower allocation per head for those with a lower need index. As further analysis was difficult because of the problems with the data, it was suggested that we used results of analyses of all public expenditure by the Office for the Deputy Prime Minister (ODPM).
\textsuperscript{105} These are available from <http://www.dh.gov.uk/en/Managingyourorganisation/Financeandplanning/Allocations/DH_4108515>.
\textsuperscript{106} The first edition (NHS Executive, 1994) described the bases for allocations for 1995-96 (this described how the results of the York Study were used in the formula); the second edition (NHS Executive, 1997a) for 1997-98; the third edition (NHS Executive, 1999) for 1999-2000; the fourth edition (Department of Health, 2003) for 2003/04 to 2005/06; and the fifth edition (Department of Health, 2005) for 2006/07 and 2007/08.
b. In 2004, on changes to the formula being very carefully explained in booklets published in relation to ACRA’s recommendations (September 2004)

c. In 2006, on the MFF. The objective being to counter cynicism by communicating clearly its purpose, advantages of inclusion in the formula. This would explain how the staff MFF relates to the Cost of Living Supplement (COLS) and High Cost Area Supplements (HCAS) (April 2005); and the GLM method (27 September 2006).

108. In November 2005, ACRA Members raised concerns over the increasing number and geographical spread of organisations in deficit, and the extent to which lack of funding might be responsible for the deficit position. It was pointed out that two-thirds of the deficits are in around 6% of organizations and likely that local circumstances would have contributed to the deficit. It was accepted that the issue of NHS deficits was largely outside ACRA’s remit.

4.6. The impacts of different components on the formula

109. This section examines the impacts of different components on the formula as it applied to the allocations in 2006/07 and 2007/08.

110. The weights of the five different components were as follows: HCHS, 77.4%; Prescribing (described as ‘pharma in the Figures that follow), 13.2%; and Primary Medical Services (PMS), 8.8%; and HIV/AIDS, 0.6%. In these allocations for 2006/07, the English Language Difficulties Adjustment (ELDA) was outside this system of weightings, and accounted for just over £13m out of a total target allocation in excess of £64bn (Department of Health, 2005a). Figure 4 illustrates the shares of all six components.

111. Table 2 gives for each of the different indices for each component their ranges and ratio of the maximum to the minimum values of each index (with a single index only for HIV/AIDS). Figure 5 illustrates how these ratios compared (except for HIV/AIDS because of the extreme value of the ratio for that index). There were three kinds of generic indices: for age, additional need and the MFF. Figure 5 shows that for both HCHS and PMS, the index for additional need had a bigger impact than for age, which was in turn greater than the MFF. For prescribing, however, the index for age had a bigger impact than for additional need (there is no index for MFF). Table 2 and Figure 6 illustrate the impacts of the different indices on the formula by multiplying their ratios (of the maximum to the minimum values) by the weight of each component. This brings out clearly how (as expected) the separate and global indices for HCHS dominated the other indices.

---

The group agreed that it should be clearly communicated to the Service that any adjustment to the MFF for these payments, meaning that PCTs receive both, would be double counting and also that the MFF is providing the right level of funding to Trusts to cover both the direct and indirect costs of employing staff.
Figure 7 shows for HCHS areas with high weighting for age tended to have low weighting for additional needs and vice versa. This is what ought to happen in a weighting system as so much of health care resources are spent in the last six months of life, so, for that reason alone, weighting for age alone would underestimate needs in areas with life expectancy below average and vice versa\textsuperscript{108}. Figure 8 shows for HCHS that there was no relationship between the weighting for additional needs and the MFF. Hence it was vital to have three different indices for age, additional need and the MFF.

In the different components, however, we would expect each index of the same type to be strongly correlated. Figure 9, 10 and 11 show this to have been so for each index in each component: for age and additional needs across HCHS, Prescribing and PMS; and for MFF for HCHS and PMS.

\textsuperscript{108} This is examined by Sutton et al (pp. 70–72).
Figure 4: Weighting of different components in the formula

Source: Department of Health (2005a)

Key:
Pharma = Prescribing
PMS = Primary Medical Services
HIV = HIV/AIDS
ELDA = English Language Difficulty Adjustment
HCHS = Hospital and Community Health Services
### Table 2: Ranges and ratios of indices for different components

<table>
<thead>
<tr>
<th>Component (Weight) and Indices</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Ratio (Max/Min)</th>
<th>Product of Ratio &amp; Component Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HCHS (77.4%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Additional Needs</td>
<td>0.71</td>
<td>1.52</td>
<td>2.14</td>
<td>1.66</td>
</tr>
<tr>
<td>Age</td>
<td>0.85</td>
<td>1.25</td>
<td>1.47</td>
<td>1.14</td>
</tr>
<tr>
<td>MFF</td>
<td>0.90</td>
<td>1.24</td>
<td>1.38</td>
<td>1.07</td>
</tr>
<tr>
<td>Emergency Ambulance Cost</td>
<td>0.99</td>
<td>1.03</td>
<td>1.04</td>
<td>0.81</td>
</tr>
<tr>
<td>Adjustment (EACA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global</td>
<td>0.75</td>
<td>1.45</td>
<td>1.93</td>
<td>1.50</td>
</tr>
<tr>
<td><strong>Prescribing (13.2%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Additional Needs</td>
<td>0.84</td>
<td>1.43</td>
<td>1.74</td>
<td>0.23</td>
</tr>
<tr>
<td>Age</td>
<td>0.74</td>
<td>1.29</td>
<td>1.70</td>
<td>0.22</td>
</tr>
<tr>
<td>Global</td>
<td>0.79</td>
<td>1.49</td>
<td>1.89</td>
<td>0.25</td>
</tr>
<tr>
<td><strong>Primary medical services (8.8%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Additional Needs</td>
<td>0.83</td>
<td>1.34</td>
<td>1.48</td>
<td>0.13</td>
</tr>
<tr>
<td>Age</td>
<td>0.83</td>
<td>1.23</td>
<td>1.61</td>
<td>0.14</td>
</tr>
<tr>
<td>MFF</td>
<td>0.95</td>
<td>1.17</td>
<td>1.23</td>
<td>0.11</td>
</tr>
<tr>
<td>Global</td>
<td>0.83</td>
<td>1.33</td>
<td>1.60</td>
<td>0.14</td>
</tr>
<tr>
<td><strong>HIV/AIDS (0.6%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global</td>
<td>0.14</td>
<td>9.11</td>
<td>65.07</td>
<td>0.39</td>
</tr>
</tbody>
</table>

Source: Department of Health, 2005b

---

109 The lowest value is for Wokingham and the highest for North Manchester.
110 The lowest value is for Tower Hamlets and the highest for Bexhill and Rother.
111 The lowest value is for Torbay and the highest for Westminster.
112 The lowest value is for North Birmingham and the highest for Herefordshire.
113 The lowest value is for Wokingham and the highest for Tower Hamlets.
114 The lowest value is for South East Oxfordshire and the highest for Easington.
115 The lowest value is for Tower Hamlets and the highest for Bexhill and Rother.
116 The lowest value is for Bracknell Forest and the highest for Easington.
117 The lowest value is for Wokingham and the highest for North Manchester.
118 The lowest value is for Teignbridge and the highest for Westminster.
119 The lowest value is for Tower Hamlets and the highest for Bexhill and Rother.
120 The lowest value is for Wokingham and the highest for North Manchester.
121 The lowest value is for Durham Dales and the highest for Lambeth.
122 HCHS: Table 4a.9 column B for age-related need adjustment, column C for additional need adjustment, column D for MFF adjustment, column E for EACA. Prescribing: Table 4a.13 column B for age-related need adjustment; column C for additional need adjustment. PMS: Table 4a.19 column B for age adjustment; column C for additional need adjustment; column D for MFF adjustment. The index for HIV in 4a.27 is derived from 6 Tables (Tables 4a.20, 4a.2, 4a.22, 4a.23, 4a.24, 4a.25, 4a.26). That reported here is the global index which is the ratio of the HIV/AIDS weighted population given in Table 4a.27 divided by the crude population given in Table 4a.9 (2006-07 HCHS weighted population, column A).
Figure 5: Ratio of maximum to minimum of indices for HCHS, Pharma, PMS

Source: Department of Health (2005b)\textsuperscript{123}

Key:
HCHS = Hospital and Community Health Services
Pharma = Prescribing
PMS = Primary Medical Services
HIV = HIV/AIDS
MFF = Market Forces Factor
EACA = Emergency Ambulance Cost Adjustment

\textsuperscript{123} The details of the sources are as given for Table 2.
Figure 6: Ratio of maximum to minimum of indices for HCHS, Pharma, PMS, HIV weighted by spend on each component

Source: Department of Health, 2005b

Key:
HCHS = Hospital and Community Health Services
Pharma = Prescribing
PMS = Primary Medical Services
HIV = HIV/AIDS
MFF = Market Forces Factor
EACA = Emergency Ambulance Cost Adjustment

124 The details of the sources are as given for Table 2.
Figure 7: Indices for additional need and age for HCHS

Source: Department of Health, 2005b\textsuperscript{125}

Key:
Pharma = Prescribing
PMS = Primary Medical Services
HCHS = Hospital and Community Health Services

\textsuperscript{125} HCHS: Table 4a.9 column B for age-related need adjustment, column C for additional need adjustment.
Figure 8: Indices for MFF and age for HCHS

Source: Department of Health, 2005b\(^\text{126}\)

Key:
HCHS = Hospital and Community Health Services
MFF = Market Forces Factor

\(^{126}\) HCHS: Table 4a.9 column B for age-related need adjustment, column D for MFF adjustment.
Figure 9: Indices for additional need for HCHS, Pharma and PMS

Source: Department of Health, 2005b\textsuperscript{127}

Figure 10: Indices for age for HCHS, Pharma and PMS

Source: Department of Health, 2005b\textsuperscript{128}

Key:
Pharma = Prescribing
PMS = Primary Medical Services
HCHS = Hospital and Community Health Services

\textsuperscript{127} HCHS: Table 4a.9 column C.  Prescribing: Table 4a.13 column C. PMS: Table 4a.19 column C.
\textsuperscript{128} HCHS: Table 4a.9 column B.  Prescribing: Table 4a.13 column B for age-related need adjustment.  PMS: Table 4a.19 column B for age adjustment.
Figure 11: Indices for MFF for HCHS and PMS

Source: Department of Health, 2005b

Key:
Pharma = Prescribing
PMS = Primary Medical Services
HCHS = Hospital and Community Health Services

129 HCHS: Table 4a.9 column D. PMS: Table 4a.19 column D.
Figure 12: Indices for EACA and HIV/AIDS

![Graph showing indices for EACA and HIV/AIDS]

Source: Department of Health, 2005b\textsuperscript{130}

Figure 13: Unified vs HCHS weighted populations (000s)

![Graph showing unified vs HCHS weighted populations]

Source: Department of Health, 2005b\textsuperscript{131}

\textsuperscript{130} Table 4a.9 column E for EACA. The index for HIV in 4a.27 is derived from 6 Tables (Tables 4a.20, 4a.2, 4a.22, 4a.23, 4a.24, 4a.25, 4a.26). That reported here for HIV/AIDS is the global index, which is the ratio of the HIV/AIDS weighted population given in Table 4a.27 divided by the crude population given in Table 4a.9 (2006-07 HCHS weighted population, column A).
114. Figure 12 also shows the relative impacts of HIV/AIDs and the EACA to be large for different reasons: the EACA has a small ratio (of maximum to minimum values) but is part of a large component of expenditure (HCHS); and although HIV/AIDs accounts for a small amount of total expenditure only, its ratio is enormous. These indices, however, have a material impact on a few PCTs only as shown by Figure 12: one PCT only (Herefordshire) had an EACA index more than 1.06; 10 PCTs had HIV/AIDs index of more than 5. These indices were (as expected) negatively correlated: the rural PCTs benefited (slightly) from the EACA index, and mainly London PCTs benefited (slightly) from the HIV/AIDs index. HIV/AIDs accounted for only 0.06 per cent of total expenditure. If the EACA were removed the impacts on total allocations would have been maximum gains and losses of 0.5% (for Heart of Birmingham Teaching PCT) and 1.31% (for Herefordshire PCT).

115. As the HCHS component was so dominant, different generic indices were correlated, and the area specific indices accounted for relatively small sums of money, this raises the question of the materiality of the various complicated refinements to the formula for the different components. Figure 13 shows the scatter of observations for PCTs with target allocations based on applying the weightings for HCHS to the whole of PCTs’ expenditure, compared with target allocations based on separate weightings for HCHS, prescribing, primary medical services, pharmaceuticals and HIV/AIDs. Figure 13 raises the obvious question of what is gained by detailed weightings for the smaller components of expenditure, which made the formula more complex and hence less transparent.

4.7. Financial deficits

116. The Report of the Health Select Committee (2006, p. 42) on NHS Deficits concluded that ‘The increases in the underlying deficits incurred by PCTs and hospital trusts have many causes’ and observed that ‘A number of witnesses argued that there was a correlation between trusts’ deficits and the allocation of funding’.

117. From 1999/2000 to 2003/04 the NHS reported small surpluses (i.e. annual total expenditures were less than the money voted by the House of Commons), but, in 2004/05 and 2005/06, the NHS in total reported net deficits of £221m and £547m respectively. These deficits were a small percentage of total expenditure, but were concentrated in a small number of organizations.

118. Examining the geographical pattern of deficits at the level of PCTs shows that, in 2003/04, these tended to be distributed randomly but, in 2004/05, tended to be concentrated in a cone-shaped area above a line from Bristol to Southampton and below a line from Bristol to the Wash. The claim

---

131 Table 3.3a column A gives HCHC weighted populations and column G gives unified PCT 2006-07 weighted populations.
that the new formula developed by the AREA team, which gave an increased weighting for additional needs, was a cause of deficits in the NHS has been examined by DH’s chief economist (Department of Health, 2007, pp. 53-62). His analysis showed any link between deficits and the new formula to be spurious for seven reasons:

a. The first year the new capitation formula applied was in 2003/04, in which the NHS in total had a small financial surplus (of £30m). Deficits became a problem the following year.

b. Changes were also made for 2003/04 to the MFF, which tended to offset the change in allocations from the capitation formula: the MFF weight tended to be reduced (increased) for those PCTs receiving an increase (decrease) in target allocation.

c. The new formula was used for target allocations, but had minimal impacts on total allocations as there was a relatively slow pace of change to target allocations over the period 2003/04 to 05/06.

d. It is misleading to look at deficits in terms of providers only or PCTs only: as a provider could have been in surplus whilst its principal PCT was in deficit, or vice versa. Hence to make sense of deficits they need to be estimated across both for PCT economies as a whole.

e. In 2003/4 there is no evidence of any tendency for the revenues of PCT health economies in areas with relatively low weighting for additional needs to be any more likely to be in deficit.

f. One cause of the emergence of financial deficits in the NHS was the loss of local arrangements for planned financial support (known as ‘brokerage’) by SHAs. Comparison of SHA-administered support to PCT health economies in 2003/04 with the age/needs index shows no general tendency for local health service managers to move resources away from the needier areas suggesting the lack of a relationship between deficits and age/needs in 2003/04 [i.e. it is unlikely that a substantive relationship was hidden by accounting practices].

g. In 2004/05, each SHA experienced similar growth in allocations but there were sharply different rates of financial deterioration\(^{132}\); all PCTs received substantial growth (of at least 8.8 per cent); and statistical analysis shows no significant relationship between the scale of increases in PCTs’ allocations from 2003/04 to 2004/05 and deficits in PCT economies in 2004/05.

119. The Report concluded (Department of Health, 2007, p. 61) that:

The tendency for deficits arising in 2004/05 to be more prominent in both the South/East and the low age/needs index areas appears not to arise from shocks to the income side of PCT budgets propagated by the resource model, and instead these effects must arise from increases in the expenditure side of the budget for the South/East and less needy areas, or from

\(^{132}\) SHAs such as Essex enjoyed a comparatively high rate of resource growth but experienced a prominent rate of financial deterioration. In contrast, West Yorkshire received one of the smallest rates of resource growth and achieved an improvement in financial balance. Norfolk, Suffolk and Cambridgeshire received an average rate of growth of resource but also the second largest deterioration of financial balance.
2004/5 changes to accounting practice that impact in regionally uneven ways.

5. Commentary of the processes of ACRA and TAG

120. This commentary offers observations on:
   a. Interpretation of, and adherence to, ACRA’s terms of reference;
   b. the research that ACRA has commissioned in terms of quality and independence;
   c. the criteria used for measuring the robustness of ACRA’s recommendations;
   d. how objective and evidence based has ACRA’s decision making been, and what judgement factors have been applied;
   e. whether ACRA’s membership is independent, representative and appropriate;
   f. whether there is a clear demarcation between areas where ACRA is expected to offer advice to Ministers and where Ministers are expected to make their own decisions.

5.1. ACRA’s terms of reference

121. ACRA’s general terms of reference as set in 1997 were (ACRA, 1998, p. 5):
   a. To advise the Secretary of State for Health on the distribution of resources across primary and secondary care, in support of the goal of equitable access to healthcare for all;
   b. To develop and apply methods which are as objective and needs-based as available data and techniques permit.

122. Subsequently ACRA was asked in a letter from the Minister of State, in March 1999, specifically to:
   a. facilitate the transition from resident ONS populations to registered populations;
   b. develop a HIA;
   c. examine cost issues; and
   d. need-related issues.

123. The above review of the scope of ACRA’s work makes it clear that ACRA have fulfilled its general and specific terms of reference. Indeed the specific issues identified by the Minister of State’s letter of March 1999 have dominated ACRA’s work programme.

124. There is, however, a question over what falls within and is excluded from ACRA’s terms of reference. Over time ACRA’s remit has become restricted to the development of formulas and not their application, which involves decisions on the pace of change, setting budgets at practice level, and the interaction of resource allocation with reference to formulas and payments (under PbR and QoF). (Although ACRA was asked to recommend how the MFF ought to be applied for acute providers in PbR.) Another issue is ACRA’s scope over research and development of better formulas, which has
been restricted to opportunities that can produce changes in the short term only (which is discussed in the next section).

5.2. The research ACRA commissioned

125. ACRA has to decide the balance of work between that to be done inhouse by DH and commissioned externally as research. Table 3 outlines the work commissioned by ACRA since 1999. Of the four issues that have dominated ACRA’s work programme, the continued problems of ‘list inflation’ were examined in-house\textsuperscript{133}. The other three have each been subjected to external research. Exploratory work, which does not warrant the transaction costs of competitive tendering, has been commissioned directly. The major projects on developing new bases for accounting for unavoidable costs and additional need have been let following competitive tendering. ACRA has hence managed the mix of work appropriately to its nature. Since the mid-1990s, a small group of researchers have consistently won contracts for the work commissioned by ACRA, which makes this vulnerable to criticisms of being a cliquish activity. This is a well-recognised characteristic in contracting for such work\textsuperscript{134}, and the best way of opening up competition would be by funding a longer-term research programme into resource allocation (as suggested below).

\textsuperscript{133} A report to ACRA in 2007 on the Population Base for PCT Revenue Allocations Post 2007/08 recommended a two-phased approach: an academic study of list variation, including a review of the registration process and an analysis of records from sample practices; and the implementation of a national solution to enhance the data quality of information from registered lists.

\textsuperscript{134} This is described by Williamson (1974.pp 34-35) as ‘first mover advantage: i.e. those who win the first contract gain such know how that they are well placed to win the next; and he later points (Williamson 1985, pp 62-63) out that this can lead to a ‘fundamental transformation’ that limits effective competition.
Table 3: Work commissioned in developing the weighted capitation formula since 1999

<table>
<thead>
<tr>
<th>Year</th>
<th>Population base</th>
<th>Integrated formula</th>
<th>Need &amp; HIA</th>
<th>Unavoidable costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999</td>
<td>Population Base Group (PBG) established in 1999 and reported in 2000\textsuperscript{135}.</td>
<td>ACRA considered a Report from the University of York (Carr-Hill and Rice, 2001) for a formula for the Non-Discretionary Component.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td></td>
<td>ACRA considered an interim report on the 'bottom up' general service module scoping study commissioned from York University.</td>
<td></td>
<td>A report from the Institute of Employment Research at Warwick University on the review of the MFF, and was presented to ACRA for information only.</td>
</tr>
<tr>
<td>2001</td>
<td></td>
<td>ACRA tendered for a new small-area study to estimate weights in the capitation formula from empirical analysis, which was awarded to the AREA team from University of Glasgow/ISD Scotland.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{135} The Groups terms of reference were: to make recommendations to reduce list inflation, ensure that it remained at minimal levels and improve the accuracy of GP and health authority data at the point of registration.
Table 3: Work commissioned in developing the weighted capitation formula since 1999 (continued)

<table>
<thead>
<tr>
<th>Year</th>
<th>Population base</th>
<th>Integrated formula</th>
<th>Need &amp; HIA</th>
<th>Unavoidable costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td></td>
<td>ACRA considered a Report from the University of York for a workload formula for the new GMS contract.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td></td>
<td>A study by the Institute of Employment Research at Warwick University examined: the size and shape of zones; and contouring, to reduce ‘cliff edges’ between zones.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>For revenue allocations post 2007/08, ACRA established a working group was established (Chair by Dave Roberts) to take this approach forward, which reported in 2007.</td>
<td>Two studies were commissioned. 1) From the Health Economics Research Unit (HERU) at the University of Aberdeen to provide an update on the existing GLM approach to the MFF. 2) From a consortium comprising Crystal Blue Consulting, the Centre for Health Economics York and the City Health Economics to explore the value of the specific and recognisable cost approach.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
125. ACRA has sought to keep abreast of developments in other countries: and
commissioned, in 1999, a literature review of recent research on developments
of health risk adjustment or capitation methods in health care in developed
countries (Rice and Smith, 2001). ACRA were presented, in 2000, with the
outcome of the review of resource allocation for the NHS in Scotland. This
influenced the specification of the research in 2001 into estimates of additional
need from the AREA study, which developed an innovative approach of using
new variables in addition to those from the census and mortality. The latest
study to estimate need using the top-down approach commissioned by ACRA
sought to take account of criticisms of the AREA work (Stone and Galbraith,
2006). ACRA also commissioned a review by Rice of approaches in other
countries.

126. ACRA has also sought to develop innovative approaches to the
development of formulas in resource allocation:

a. This was essential in developing the HIA: the 1999 review of the
literature found no evidence from other countries to support a formula
on health inequalities. ACRA did receive an information paper on an
ESRC-funded project to estimate numbers within a designated
population who have a specific condition, which was complementary
to work on the bottom up health inequalities scoping study.136

b. ACRA continues to explore an alternative to the top-down approach
to resource allocation as an option for the long-term and has, in 2007,
commissioned a person-based approach in analyses of utilization.

c. ACRA explored the alternative of the specific cost approach to
estimating unavoidable costs, which was informative but did not offer
a sound alternative to the general market approach.

127. Although ACRA has encouraged different approaches to developing
formulas from those that have been used in the past, exploratory work
has shown little change to have been possible because of the data that
are routinely available. Williamson (1990) quotes Simon’s
observation (Simon, 1984) on the differences between the physical
sciences and economics when confronted with problems with data:

In the physical sciences, when errors of measurement and other noise
are found to be of the same order or magnitude as the phenomena
under study, the response is not to try to squeeze more information out
of the data by statistical means; it is instead to find techniques of
observing the phenomena at a higher level of resolution. The
 corresponding strategy for economics is obvious: to secure new kinds
of data at the micro level.

126. All work seeking to estimate need lacks data at the level of the
individual, and all work to estimate the staff MMF has lacked data at the level

136 This was based on the calculation and attribution of condition-specific prevalence rates using the
Health Survey for England (HSE). ACRA echoed TAG’s concerns that patient categories for some
conditions were based on very small numbers, a consequence being that the results did not show the
expected ‘gradient’ in inequality.
of the travel to work area. Expert researchers have been asked to revisit these constraints to see if they can squeeze more information out of standard datasets by statistical means within short timescales. Carr-Hill et al (1994a, p111) concluded that in estimating need, ‘… if an empirical basis is sought for identifying the determinants of utilisation, the only method likely to yield significantly more robust and credible results (than their study) is by the use of long-term cohort studies of individuals’. Such a study would be a shift in the paradigm of research into resource allocation by breaking out of constraints of existing limited sets of data.

127. ACRA’s work programme has, quite appropriately, been driven by the overriding objective of making feasible changes to formulas in the short term. The radical approaches that ACRA have explored for the longer term have been outside the work ACRA can currently commission as they involve pioneering research. Smith (2007, p. 140) observes that ‘in spite of the very large sums of public finance involved, the volume of UK research into formula funding is pitifully small’. It would seem to be beneficial to enable ACRA to take time out to explore possible directions in the longer term that are not subject to two constraints: the imperative of producing changes within a year, and inadequacies in existing data. There are two possible avenues worthy of exploration. First, to consider what might be possible with developments in data over the next five to ten years. Second, to develop research that could guide the collection of data to improve the bases of resource allocation. Another advantage of funding a research programme would be to widen the expertise available to bid for work on shorter timescales.

128. ACRA agreed (ACRA, 1998) two sets of criteria that would be applied in evaluating formulas for resource allocation (see Annex 9):
   a. Essential: technical robustness; transparency; objectivity; plausibility; freedom from perverse incentives; reliability of calculation.
   b. Desirable: comprehensibility to non-specialists; durability; practicality; clarity of contribution of indicators; flexibility; stability; materiality.

129. It is useful to have these multiple criteria spelt out as they capture very well what those who have designed formulas have sought to apply since the RAWP Report of 1976. All ACRA’s recommendations have to satisfy the criteria of practicality, durability, stability, reliability of calculation, and consistent applicability. Throughout there are tensions between: transparency, comprehensibility and technical robustness; and flexibility and materiality.

130. The methods recommended by RAWP (although criticised for their complexity at the time) now look to have been transparent and comprehensible to non-specialists. This is because they did not include a MFF and did not derive weightings for additional need from data on utilisation. The RoR methods that estimated weightings for additional need from analyses of small area variations in utilisation using elementary statistical methods (Ordinary Least Squares regression analysis) were comprehensible to non-specialists but were not technically robust. To satisfy the latter criterion requires more sophisticated statistical methods, which are neither transparent nor comprehensible to non-specialists. Estimates of the MFF have been made
using the general labour market approach, which appears to be the only approach that is technically robust. There are problems in translating indices of labour markets into what ought to be differences in labour costs in the NHS: the MFF has consistently been seen to lack face validity with questions over zones, ‘cliff edges’, the inclusion of medical and dental staff, and whether there should be a ‘cut off’. Empirical analyses of need and costs are always open to challenge. ACRA has sought to ensure the empirical estimates of each were produced by experts using methods that were technically robust.

131. ACRA has to make judgements on the materiality of refinements to the formula: there are particular problems where an element affects a few PCTs only, or where there may be a set of related issues none of which is material in its own right, but may have a material cumulative impact. There seems to be a natural history of development of formulas: the initial formulation is a simple formula driven by a few principal elements; this is followed by an attritional process of demands from the NHS for special allowances, which results in various elements are added; which is followed by a radical pruning to return to a simple formula and so on. Part of the problem here is developing a national formula that has to be applied consistently across PCTs with very different characteristics in terms of the needs for, and supply of, health care.

132. One of ACRA’s functions ought to be to provide a bulwark against different elements being added to the formula so that, as one chair of TAG described (Alan Meekings), it begins to look like a Christmas tree. One of the aims of the RAWP Report (Department of Health and Social Security, 1976, p. 8) was to keep the methods as simple as possible subject to the criterion of materiality. The working party ‘rejected many approaches which might have made the criteria more sensitive, but which on examination would have led to much greater complexity with little significant gain in the result’.

133. It may be useful to formulate rules of materiality in terms of the potential impacts of a new element in terms of both its total share of expenditure for England and range of impacts on PCTs. The key point is that any formula, however sophisticated its technical basis, can only offer a form of rough justice. Each of the four principal drivers of the formula is subject to margins of error: population size, and weightings for demography, additional need and the MFF. There seems little point in seeking refinements that are within these margins of error. There looks to be a good case for ACRA reviewing the materiality of the allowances for: the English Language Difficulties Adjustment (ELDA); the Emergency Ambulance Cost Adjustment (ECA); and HIV/AIDS. It would be helpful to enforce stronger criteria on materiality so that ACRA can focus on getting the big things right: better data on population; more satisfactory ways of accounting for unavoidable variations in costs; and economical ways of accounting for demographic and

---

137 As illustrated by the PCTs at the extremes of the different weighting elements for HCHS in Table 2: North Birmingham and Herefordshire (EACA); Tower Hamlets and Bexhill and Rother (age); Wokingham and North Manchester (additional needs); and Torbay and Westminster (MFF)
138 ACRA recommended in March 2008 removing ELDA on grounds of materiality and that the data that supports the adjustment is very old and cannot be updated without an additional data collection
additional need that are sensitive to differences across different types of health care. The last includes taking a strategic view on which common indices can be used across different types of care and where different indices are justified on grounds of materiality. The RAWP report sought economy and transparency with much less knowledge than we have now and for a formula that applied to RHAs and HCHS only. The demands on the current formula are greater because of its wider scope and application on a much smaller scale. Nevertheless we can use the knowledge we have of the materiality of refinements to simplify the formula so that it can be more transparent and hence easier to explain and more acceptable to the NHS and the public.

5.3. ACRA’s decision making

134. ACRA’s main recommendations for change have followed from:
   a. evidence produced from projects let by competitive tendering;
   b. a technical review of that evidence by TAG;
   c. consideration of technical and policy issues by ACRA.

135. DH publishes clear accounts of the bases for the formula used for resource allocation. But there has been nothing published on developments in resource allocation since 2002. It would be beneficial to have publications on: the rationale of the current formula and accounts of alternatives that ACRA has explored and why these have been rejected.

136. A series of policy decisions have been made either in the absence of an advisory group, or without reference to, or against the recommendation of, an advisory group. These include the following decisions:
   a. not to include a weighing for deprivation in the formula used between 1991/92 and 1994/95 (although this was recommended by the NHS Management Board’s Review);
   b. to apply a weighting for additional need to only 76 per cent of HCHS spend in 1995/96 (although this was criticised by the research team that estimated additional need and broke with previous practice);
   c. to include medical and dental staff in, to remove the cut-off from, the staff MFF;
   d. the removal of the HIA.

137. Some of the above decisions are questionable. The first two have been reversed and ACRA recommended reversing the third and fourth in March 2008. This strongly suggests the value of ACRA and the need to justify policy decisions made without ACRA’s support. There has, however, been no regular report to ACRA on the extent to which ACRA’s advice has been accepted by Ministers in changes to the formulas. In 2001, ACRA asked DH to report on the outcomes of TAG and ACRA’s recommendations. These were that ACRA’s recommendations for the 2001 had been adopted with changes in

139 These recommendations followed on new analyses of the MFF and a new basis for the formula for funding PCTs for the objective of equal access in which the elements of the AREA formula that were intended to capture unmet need have been removed. The removal of the HIA may have been justified on grounds of materiality, but this raised the question of whether its rationale, of reducing health inequalities, remained an objective of the formula.
respect of the Health Inequalities Adjustment and review of the staff MFF. There have been no such reports since. It would be helpful for DH to publish annual reports on ACRA’s deliberations, recommendations, and Ministers’ decisions.

5.4. Membership of ACRA and TAG

The role of TAG is to advise on technical issues in resource allocation and ACRA on policy issues. Each group includes experts and members from the NHS and DH and, as each is advisory, Ministers make decisions on what is normally described as matters of policy. There is inevitably overlap of roles between the two advisory bodies and DH. Although TAG does not seek to address policy issues, ACRA does consider both policy and technical issues. There is merit in having two groups, but the chair of each needs to guide discussion to recognize the distinctive contribution of the other: thus TAG to be reminded that policy issues will be considered by ACRA, and ACRA that technical issues have been considered by TAG.

ACRA is chaired by Mr David Fillingham (Chief Executive, Bolton Hospitals NHS Trust) and two officials from DH finance division are the secretariat. Other members were:

- Three from DH (two economists, one from finance)
- Five from the NHS (two from SHAs, one from a PCT, one from an acute trust, and one from the information centre)
- Four from general practice (from the BMA, a practitioner, and two academics)
- Three other academics (with expertise in econometrics, social policy and population statistics).

The Technical Advisory Group (TAG) is chaired by Ms R MacDonald (Chief Executive, Bath and North East Somerset PCT) and two officials from DH finance division are the secretariat. Other members were:

- Five from DH (three economists, one from finance, and one from the prescribing support unit);
- Five from finance in the NHS (three from acute trusts and two from PCTs);
- Four academics (with expertise in econometrics, operational research and public health); and
- Analysts from the Scottish Executive Health Department, and the BMA, and from DH.

ACRA and TAG operate on the tradition of voluntary input from members. For academics, this offers an important opportunity to relate their research to major issues of public policy. For others, the demands of these bodies have to be fitted alongside the heavy demands of their day jobs. As the NHS is the key group that ought to have confidence in the formulas, it is appropriate that each body is chaired by NHS chief executives who are respected by their peers.

140 but located in the regional public health group based at Government office Northwest.
5.5. Practice in other countries

Rice and Smith (2001) surveyed approaches to capitation and risk adjustment in 20 countries and pointed out the common interest in this system for financing purchasers of health care in countries with different models of finance: social insurance (the Bismarck model) or a NHS (the Beveridge model). Van de Ven and Ellis (1999) emphasise the fundamental importance of risk adjustment in the development of fair competition between sickness funds and Chernichovsky and Van de Ven (2003) outline developments in five countries that use the Bismarck model in Europe. Rice and Smith (2001) point out that most countries use some form of risk rating in addition to demographic risk (i.e. additional need) derived from analyses of empirical data. Rice has undertaken a recent review commissioned by ACRA, which included the development in Wales of what has been described as an epidemiological approach, which uses morbidity data based largely on analysis of the Welsh Health Survey (WHS). This approach was recommended by the Townsend Report (NHS Resource Allocation Review, 2001) and uses the heroic assumption of a one-to-one relationship between reported levels of morbidity and the allocation of resources. As Rice observed: ‘For example, an area which had 5% of the total injuries due to accidents, as measured by the WHS, would receive 5% of the amount spent nationally on A&E, irrespective of its share of the Welsh population’.

Methods of capturing additional risk are constrained by access to data: with aggregate data only being available in most countries. Individual level schemes were found in: Stockholm County (which was proposed to be extended to Sweden), Alberta, the Netherlands and New Zealand. Rice and Smith (2001) saw as the Stockholm model as ‘the most satisfactory way of setting capitation payments because it minimises the ecological problem associated with more aggregate data’. This point is made by Diderichsen et al (1997) in their comparison of systems of capitation in England and Sweden. Van de Ven et al (2004) describe the development of risk rating for hospitals and prescribing using data on utilisation by individual patients (based on both Diagnostic Cost Groups (DCGs) computed from hospital diagnoses only and Pharmacy-based Cost Groups (PCGs) computed from outpatient prescription drugs). Smith (2007, pp. 81 to 86) describes the methods of risk rating used for Medicare that use data at the level of the individual. This literature identifies new opportunities for developments of formulas in England using data at the level of the individual. Smith (2007, p. 99) concluded that, in the absence of such data, small-area analysis is a pragmatic and cost-effective alternative.

---

141 The five countries are Belgium, Germany, Israel, the Netherlands and Switzerland
142 They point out that at the time of their survey: France, Israel and Japan used age alone; and Germany and Switzerland used age and sex. Smith (2007, pp. 134-35) points out, however, that the Scottish government rejected revising the capitation formula using results from empirical analysis. And the (flawed) empirical results of the RoR were not followed by other UK countries or English RHAs, unlike the methods recommended by the RAUP Report (Sheldon et al, 1993; Bevan and Spencer, 1984)
144. The international surveys of risk rating say nothing on systems of governance in the development of formulas. I have been able to find out about approaches to governance in Scotland, New Zealand, and the Netherlands and outline these arrangements here. (I was unable to find information on current governance arrangements for resource allocation to the NHS in Wales and Northern Ireland).

145. The NHSScotland Resource Allocation Committee (NRAC) was established in February 2005 with a remit to 143:

a. improve and refine the Arbuthnott Formula for resource allocation for NHSScotland;

b. advise on possible formulaic approaches to the parts of health expenditure not currently covered by the Formula (e.g. primary care dental, pharmaceutical and ophthalmic services);

c. keep under review the information available to support existing elements of the Formula and consider the inclusion of new data (e.g. ethnicity);

d. consider in the light of the pilot exercises adjustments to the Formula for unmet need; and

e. consider any relevant issues that are referred to it.

146. NRAC includes non-executive and executive members of NHS Boards in Scotland, one GP, one academic, one co-opted member and is supported by officials 144. NRAC has core criteria to be used in considering the merits of different formulaic options. The equity objective is that of access to health services. The other criteria are similar to those used by ACRA (see Annex 9): practicality, transparency (although avoiding over-simplification of details which might add precision), objectivity, relevance (defined as making explicit where hard information is being used about one aspect of a service to make some assumption about an area where information is less good or absent), face validity, stability, responsiveness, and evaluability (capable of being tested against the objective of increasing equity of opportunity of access). The interesting omission is the criterion of materiality and addition of evaluability.

147. I am grateful to Nicholas Mays (New Zealand Treasury) for this account of the processes of updating and revising capitation formulas in New Zealand. The system of health care in New Zealand is organised on the basis of 21 District Health Boards, which are responsible for planning most services and delivering hospital services, and were created in reforms that abolished the separation of purchaser and provider (Devlin et al, 2001). DHBs are funded by a capitation formula 145, which is reviewed every five years. In the interim, there have been annual updates and unavoidable technical adjustments to deal principally with two things: new population projections which are updated annually (NZ has a census every five years) and addition to District Health

---

143 Details are given at <http://www.nrac.scot.nhs.uk/index.htm>

144 from the Scottish Executive Health Department and Information and Statistics Division in NHS National Services Scotland.

Board (DHB) budgets of new types of funding. The annual ‘technical’ adjustments are analysed and proposed by the Ministry of Health and put to a DHB advisory group comprising nominees from each ‘region’ (groups of DHBs). These are usually directors of purchasing and/or finance and are chosen by the DHBs themselves. The group is strictly advisory: the final decisions are taken by Ministers on the advice of Ministry of Health officials. The principal role for the group (which is not a standing group -- membership changes from year to year) is to provide peer review and consensus support for changes (the group rarely votes).

148. The process of the five yearly review is expected to consist of three stages:
   a. Consultations with DHBs through workshops and a questionnaire, guided by an advisory group (including technical ‘experts’).
   b. A policy paper setting out areas needing to be reviewed, the rationale, and a work plan for analysis.
   c. Presentation of proposals for change based on analysis to an advisory group (likely to be DHB representatives and technical experts). The issues that are likely to be put forward by DHBs include: the higher costs faced by small, remote DHBs to provide a full range of services to small populations; taking adequate account of the poorer health status of disadvantaged populations; mental health, and how to integrate separately identified ‘new initiative’ funding into the formula.

149. The practice in New Zealand offers an interesting contrast with England, in that it is based on a regular five-yearly review that begins with consultation with the NHS. I am unaware of any similar opportunity for consultation with the NHS in the various reviews of the formulas in England, and it might be worthwhile to consider this.

150. I am grateful to Frans van Zon (who is responsible for risk equalization in the Ministry of Health in the Netherlands) for this account of the processes of updating and revising risk adjustment in the Netherlands to compensate health-insurance-companies for differences in health risks in their populations. The formula for risk adjustment is revised each year to take account of changes in the system of health care (new developments, opportunities, costs and policies) and improvements to the model using data

---

146 The way this worked in 2007 was as follows. Insurance-companies received a premium-subsidy for every insured person with costs estimated to be higher than average for three categories of costs: hospital care, production-independent; hospital care and specialist care, production-dependent; outpatient care. Production-independent costs for the budget year (t+1) were estimated as the historical costs per capita of the year t-1 (revalued for inflation). Production-dependent costs and costs for outpatient care were estimated for each insured person using characteristics age/gender, socio-economic factors (like for instance wage-earner, self employed, disablement insurance benefit or bad housing district), costly (chronic) disease (indicated by medicine use (pharmacy cost groups) are by clinical use (diagnostic cost groups) all measured in t-3. The insurance-company received the premium-subsidy as a product of the number of persons insured, their characteristics and the average costs (in t+1) per characteristic. Production-independent costs were fully adjusted subsequently for real cost development, production-dependent costs were adjusted for about half, and no adjustment was made for costs of outpatient care. The degree of these adjustments indicate the risk the insurance-company bore. There were also subsequent corrections for movements between companies of the insured, errors in estimates of total costs and for the insured with extremely high costs in one year.
supplied by the insurance companies. These developments are funded by a continuing research programme. The Minister of Health is responsible for, and makes decisions on, risk-adjustment. The Minister is advised by three advisory bodies and ad hoc committees of technical experts. There are three bodies that provide advice on research\textsuperscript{147}, implementation\textsuperscript{148} and policy\textsuperscript{149}. The arrangements in the Netherlands are similar to those in England, the principal difference being that there is a continuing research programme, which has its own governance arrangements.

6. Conclusions

151. The work done so far on this review leads to three principal conclusions.

a. All attempts to account for need and unavoidable costs will always be essentially contested. In 1976, the Government was content to accept RAWP’s elegant and economical approach to accounting for additional need based on a single index (the SMR) and a heroic assumption (with a weighting of one). Since then Governments have sought an evidence base for both additional need and unavoidable costs from analyses of empirical data. This research has been conducted in an exemplary way: the contracts were let though competitive tendering, the studies were done by experts using state-of-the-art techniques, their work was subject to challenge from experts, and from 1997, governed by a policy and technical group.

b. There are serious constraints on how further analyses can improve formulas given limitations in the basic data. Even the basic building block of population statistics at the level of general practices is insecure: we still lack a good explanation for the material differences between estimates of population from censuses and registered lists. And no single body is responsible for the accuracy of data on registered lists. In estimating additional need, we lack data at the individual level on measures of risk and utilisation. In estimating the staff MFF we have lacked data on travel to work areas. More fundamental approaches to developing a bottom-up approach to resource allocation are impractical without a transformation in the data that are available routinely. But it would be naïve to believe that even if such data were to become available this would enable us to banish controversy from the formulas used in resource allocation.

c. The process that has developed since 1997 with two standing advisory groups on resource allocation has worked well (and offers what RAWP

\textsuperscript{147} which meets eight times a year, whose membership is from the department of health, the health insurance board, insurance-companies (and their association) and the bureaus who do the research.
\textsuperscript{148} which meets ten times a year, whose membership is from the department of health, the health insurance board, and the association of insurance-companies.
\textsuperscript{149} which meets four times a year, whose membership is from the department of health, the health insurance board, and the association of insurance-companies.
called for 30 years ago). These groups have: considered the scope of
developments in the light of Ministerial priorities; set priorities; tried to
develop new ways of accounting for the HIA, additional need and
avoidable costs; been focused on making changes for each year with
best use of data and methods. Hence ACRA has identified what had to
be done to provide a sound basis for resource allocation each year and
also to map out options for the longer term.

152. There is hence much to be proud of in the development of capitation
formulas. That does not, however, mean that what we have is perfect.
Although the terms of reference of my review did not ask for
recommendations, it is appropriate to end with a few suggestions for
improving the processes through which capitation formulas could be
developed in future:
   a. To clarify the objective of formulas: are they to promote equity of
      access for equal risk or reduce health inequalities, or both?
   b. To agree criteria for materiality of refinements to the formula.
   c. To expand the membership of ACRA and TAG to include
      representatives of other services, such as a Chief Executive of a mental
      health care trust.
   d. To enable ACRA and TAG to benefit from access to expert advice on
      specific issues such as labour economics.
   e. To consider formalizing a process of five-yearly reviews that begins
      with formal consultation with the NHS.
   f. For DH to publish:
      i. the rationale of the current formula and accounts of alternatives
         that ACRA has explored and why these have been rejected;
      ii. annual reports on how ACRA’s recommendations have and
         have not been implemented; and
      iii. reports on the impacts of resource allocation on inequalities in
         use of services and in health outcomes.
   g. To widen ACRA’s remit so that ACRA could commission longer-term
      research in resource allocation and in this way increase the expertise
      available bidding for short-term contracts.
7. References


Carr-Hill *et al* (1994c) *Exploring the possibilities of deriving a capitation formula for guiding the setting of GP fundholder budgets: Results of a study commissioned from the University of York by the NHS Management Executive*.  


Carr-Hill (1999? ) formula developed by for setting fair share targets for GPs at HA level?


<http://cep.lse.ac.uk/pubs/download/dp0843.pdf>


RAWP 1: *History of the Staff MFF*.

RAWP 2: *Population Data For Allocations*.

RAWP 3: *The Exeter dataset and attribution*.

RAWP 4: *A Brief History of Resource Allocation in the NHS 1948-98*.

RAWP 5: *A History of GP Distribution*.

RAWP 6: *The Years of Life Lost Index and the Health Inequalities Adjustment*.


