The Hospital Revenue Allocation Formula:

The development of the National Health Service since 1948 has seen increases in overall expenditure persistently accompanied by considerable differences between regions in the standards of provision of health services.

The N.H.S. was thus initially faced with a wide variation in the standards, largely due to the unequal regional distribution of income and charity, and with the problem of removing this variation in order to provide similar treatment "to secure improvement in the physical and mental health of the people."[1] (For the purposes of this discussion equality of provision will be broadly interpreted as equal access to equal-quality treatment for equal risk groups regardless of the region they live in).

The difficulty of removing these variations in standards lies partly in the method of funding used to allocate revenue in the hospital service from 1948 until quite recently. Regions received an annual sum sufficient to maintain their existing services (with an allowance to counter cost inflation) plus a margin for any special requirements accepted by the Ministry of Health as necessary or desirable. Meanwhile, hospital building has carried on to up-date the capital of the hospital service, and in part to reduce the differences in the regional availability of beds. The slow pace of the hospital building program in the 'fifties, due to economic pressures on government expenditure levels, also contributed to the continuation of inequalities in regional standards and expenditure. The extent of the continuing differences has been outlined elsewhere,[2] and a simple comparison of expenditure per head is sufficient to indicate the order of magnitude involved. (Fig 1)

---

1. The National Health Service Act 1946.

2. Griffiths, D. A. T., Inequalities and Management in the National Health Service, The Hospital, July 1971.

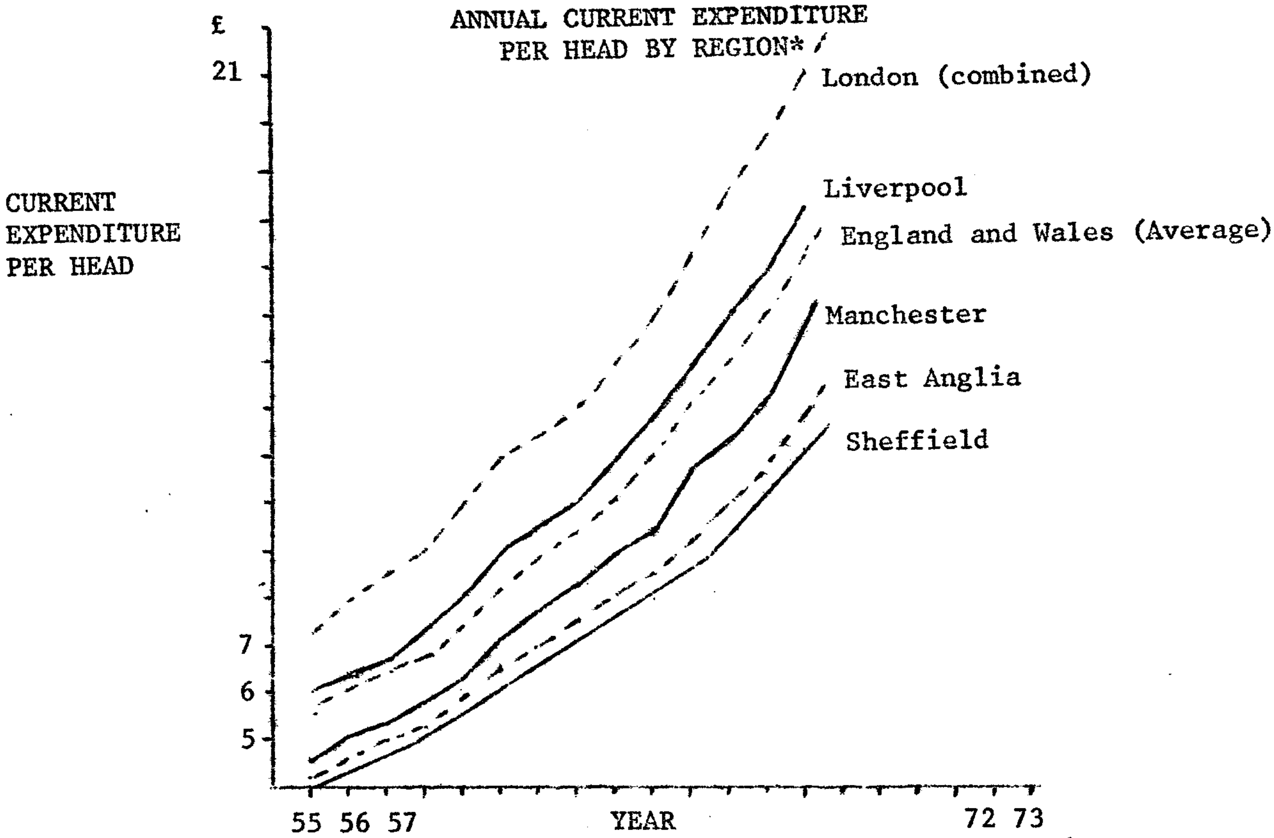

Fig. 1

In order to reduce and eventually eliminate the discrepancies between regions, the D.H.S.S. has introduced a standard formula to be used to calculate the allocation of funds to each region. Revenue is to be allocated according to three separate regional factors; the population, the bed-stock, and the case-flow, as measures of need, capacity, and throughput respectively.

Population: The formula takes the forecast population for the region, broken down by age and sex, and weights each group by the appropriate national occupancy rate, from HIPE data. This weighted population is then adjusted to take account of imports and exports of patients across regional boundaries and to teaching hospitals in the region (which are separately funded) to reach a final adjusted population figure. The total annual allocation to R.H.B's is then divided on the basis of these adjusted populations, and the allocations indicated are compared with the

---

*From Griffiths. op. cit.

previous year's actual allocation to arrive at the financial adjustment required on population grounds.

Beds: The formula takes the average daily total of occupied beds for each region by specialty and multiplies it by the estimated average cost per annum of each specialty bed, to obtain the allocation required to finance the region's beds at average cost. Adjustments are made for out-patients by adding total specialty visits multiplied by the average cost per visit, and also for geriatric and chronic sick beds because of what is seen as the need to improve these areas of the hospital service. The resulting allocations are then all adjusted proportionately so that the overall total equals the current year's allocation for the hospital service. This allocation is then compared with the previous year's allocation to obtain the variation required on the basis of bed-stock.

Cases: The formula uses case-flow for each region from SH3 data for the previous year, by specialty, and multiplies each specialty case by the appropriate national average cost per case. This gives the required regional allocation. Again, corrections are made for M.I. and M.H., as well as outpatients, and then, as with the beds measure of allocation, all the totals are adjusted proportionately so that the national total equals the revenue total to be allocated. The regional totals were then compared with the previous allocation to show the adjustment required on case-flow grounds.

The formula draws the three strands together by adding the three sets of variations together. Any combination process of this kind must use arbitrary weights for each factor, according to its relative importance in the formula - even simple addition of the three implies equal weighting; so, since it regards population as the more important factor, the D.H.S.S. has accepted a weighting of 1/2 for the population adjustment, 1/4 for the beds' adjustment, and 1/4 for the case-flow adjustment. The aim of the introduction of the formula is the removal of inequalities of allocation in the hospital service, but the size of these means that any attempt to remove them in a single year would considerably disrupt the hospital service. For this reason, the elimination process is to take place over a ten-year period, with one tenth of the required adjustment being implemented in the first year, one ninth in the second, and so on, until the remaining adjustment is fully implemented in the tenth year (A simplified calculation is shown in the appendix). (Adjustments for each year are calculated with the up-dated elements in the formula, and are implemented with respect to the previous year's allocation and not that in the initial period of the formula.) Beyond the tenth year after implementation of the formula, 1981/2, the formula's main use will presumably be to counter the effects of small annual changes in the regional populations.[3]

Criticism of the formula is not simple, because of the seemingly direct connection between the aims of the formula and the means introduced to attain them. However, an examination of the inter-regional consequences of the formula may serve to highlight some of the underlying complications in its present definition.

The inital document on the formula[4] indicates the reasons for including population, beds and cases. The weighted population is regarded as a measure of the medical needs of each region, the

---

3. The formula will continue to take account of inflation and adjustments such as London weightings, but these are comparatively insignificant since they would necessarily be included in any and every allocation scheme.

4. D.H.S.S. circular RHB Chairmen 3/70

bed-stock as a measure of the ability of each region to meet this need, and the case-flow as ".. an indication of the use of facilities necessitated by circumstances not covered by the other elements of the formula."[5] The document also comments on the need to improve the formula by gradual refinement, suggesting the elimination of the bed-stock element as soon as possible. While no reason is provided for the elimination of this part of the formula, it appears to this author at least that the use of the bed-stock element is at present uncertain and that its removal may indeed yield an improvement. The uncertainty arises for two reasons; because of intercorrelation between the 'Beds' and 'Cases' elements in the formula, and because of the difference between the stated aim of using the bed-stock and its actual use in calculations.

The intercorrelation is the obvious relation between the case-flow of a region and its bed-stock. Whilst it is typically the case that the two do not move proportionately, because relatively higher-bedded regions tend to operate a longer stay per case, it is clear that the two will provide very similar measures of the region's size, and an indication of its present allocation and provision of service. Thus, the argument for including both elements in the formula must rest on the belief that they provide different indications of the region's positon. The bed-stock element measures the capacity of the region while the case-flow, apart from being a similar index of capacity, also highlights any patterns of morbidity which might require corrections to the allocation based on the adjusted population alone. The adjusted population provides a measure of need based on average morbidity for particular age and sex groups; differences of morbidity

---

5. Ibid. para. 6

between the region and the national average will affect the medical requirements of the region, and may be indicated by an analysis of previous case-flows. However, the pattern of morbidity revealed in the regional case-flow is not a direct measure of regional morbidity since it is biased by the admissions policy and referral practices of the region's general practicioners and hospital staff. These in turn are affected by the capacity of the region, introducing a further correlation between bed-stocks and case-flows. In view of these complications, it may well be necessary to reconsider the inclusion of both elements in the formula, with a view to establishing ~~that~~ whether both in fact add useful information to the calculation.

In the calculation process itself, a further reservation attaches to the bed-stock element in the formula. If the regional bed-stock is to measure the ability of a region to meet its medical needs, which are comparatively short of beds, in order to give a compensating adjustment to pay for more intensive, and therefore more costly, treatment. This would facilitate a faster turnover of cases and counter the effects of the bed shortage until such time as the region's bed-stock can be altered by the hospital building programme. However, the formula in fact provides more money to comparatively higher bedded regions on the basis of the cost of running this bed level in the previous year. Clearly, this does indicate regional capacity and the likely cost of running it, but it fails to take account of the ability of a region to meet its needs, since it makes no allowance for relatively under-provided regions. The result of including occupied bed-stock in the formula will be to continue, rather than eradicate, a part of the difference in regional standards of hospital service. As has been pointed out elsewhere,[6] one effect of the formula's present structure will be to soften the impact of the other adjustments on the relatively well-provided regions in the N.H.S. But given the existence of a ten year adjustment period, any required softening of the blow could be simply achieved by altering the length of this period rather than by the use of a contradictory element in the formula itself. We therefore appear to be justified in concluding that the removal of the bed-stock element in the formula would be an improvement in the formula itself, and in its efficacy.

An attempt to convert it to a measure of bed needs by taking, for example, the reciprocal of bed-stock as a basis for allocations would also fail to produce unambiguous conclusions, since the use of beds utilised in such a measure would lead to a situation where a fall in the utilisation rate of a region would lead to an increase in its allocation, and vice versa.

An alternative method of evaluating the inclusion of a measure of the bed-stock in the formula is to consider the consequences of its omission. To simplify this, we will assume the existence of two regions with identical population numbers and morbidity, and with identical case-flows through their hospitals, but with one region having more beds than the other (Both total beds and utilised beds).

Omitting the bed-stock element from the formula would lead to these regions having the same theoretical allocation. Obviously, the higher bedded region operates a higher length of stay per case, on average, since it uses more beds for the same case-flow. However, if the regions have the same allocation, they are obliged to have the same average cost per case in order to keep within their allocations. We therefore have two regions, one with a longer stay but lower daily cost

---

6. Griffiths. op. cit.

per patient than the other. In the absence of an accurate production function for hospital services, we are unable to comment on the relative efficiency of either treatment. If, however, we assume that both regions bring about identical improvememts in the health of their patients, then there is no case for anything but equal funding for the two regions, since the provision of medical care in the two is identical, and is achieved at identical cost. Although it may be argued that this is an obvious consequence of an extreme assumption, it is nonetheless the case that we have no scale on which to compare the two treatment practices other than cost; and whilst discharge may not occur at the same state of health in every region, there is little evidence to suggest that it occurs at very different states of health; except perhaps between regions with large differences in the extent of non-hospital care available.[7] If variations in treatments with the same average cost are felt to be serious, then standardisation on cost grounds alone will fail to achieve the equality of provision required, and some limitation on clinical freedom will be required. Once again, there is no role for the bed-stock in any of these considerations, and so the conclusion is that the bed-stock measure, in its present form, should be removed from the formula.

In terms of the original formula, we now have the regional allocations based on adjusted populations and case-flows in each region. These elements appear to be correct in that they take account of the needs of each region and the rate at which these are being met. We will now examine the efficacy of the formula in terms of its ability to achieve equal provision, and also its power to enforce this, once achieved.[8]

The implications of the formula with population and case-flow

---

7. These complications are omitted for greater simplicity, but remain an important factor in any allocation process.

8. We will again abstract from the adjustments required for inflation in order to simplify the discussions.

elements only can be analysed by consideration of the equilibrium solutions produced in terms of expenditure, case-flow and costs per case. Fig (2) provides an illustration of the expenditure levels for a[9] given population and various case-flow levels.

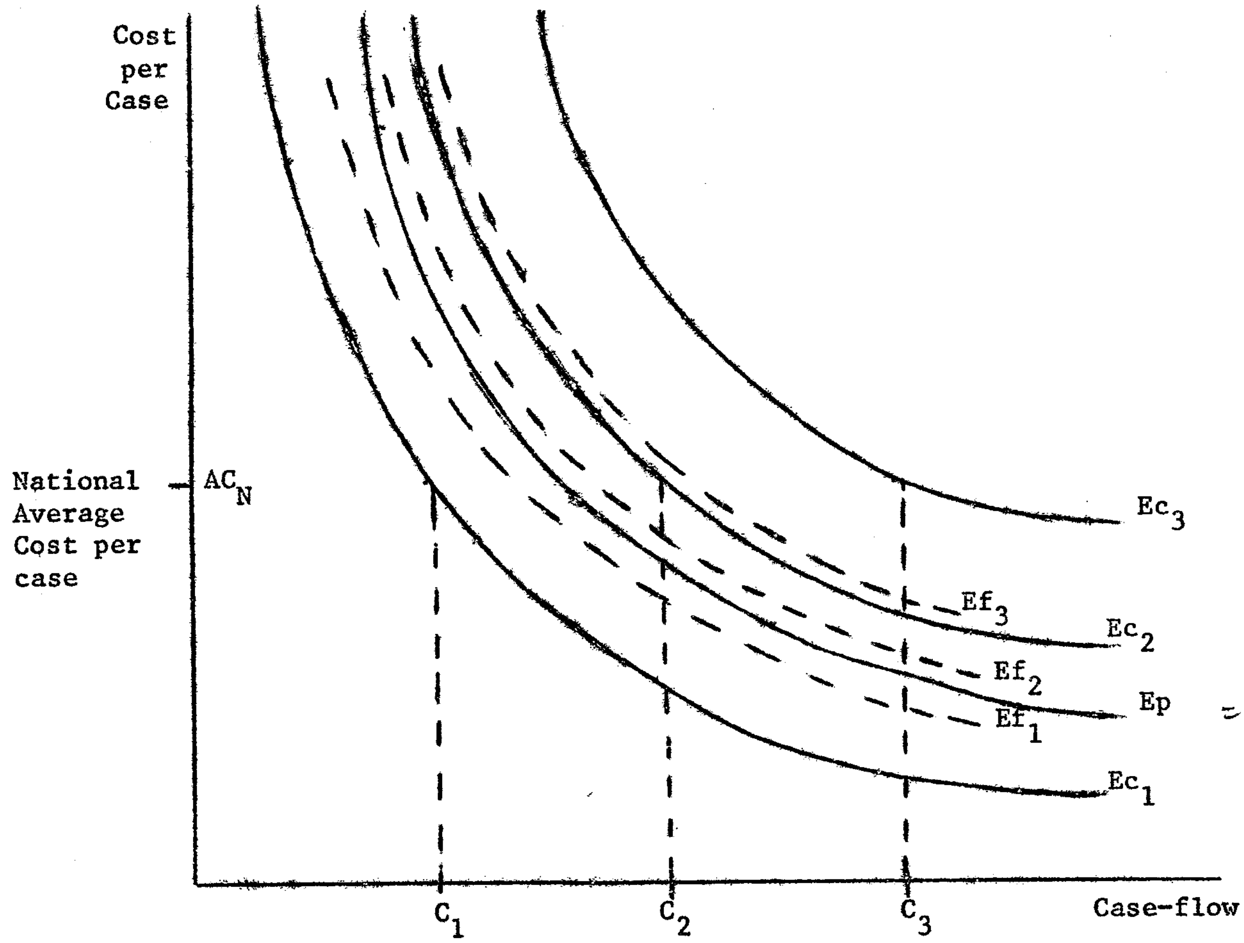

Fig. (2)

Curve Ep represents a constant expenditure for the region; Ep is the expenditure or allocation for the region required on grounds of the adjusted population only. For each case-flow level C_1, C_2, C_3 an expenditure level. (Ec_1), (Ec_2), (Ec_3) is indicated by the case-flow element in the formula, from the national average cost per case level, AC_N. Given the original weights for each element in the formula, whereby the

9. We abstract from complications such as the specialty mix, and assume the existence of one composite case-flow and average cost. Also, the effects of one region's cost on national average cost are ignored, to simplify the analysis.

population factor receives twice the weight of the case-flow factor, the combined expenditure curves are equal to two thirds of (Ep) plus one third of the relevant (Ec) for the region. The region then adjusts its cost per case and case-flow to the new expenditure levels, (E_{f1}), (E_{f2}) or (E_{f3}), with the result that the following year's allocation is altered if case-flow is changed by the allocation. The equilibrium expenditure situation is one where the region has a case-flow and cost per case such that it is in a position where the population and case-flow adjustments to its previous position cancel each other out. This is perhaps most simply shown by example.

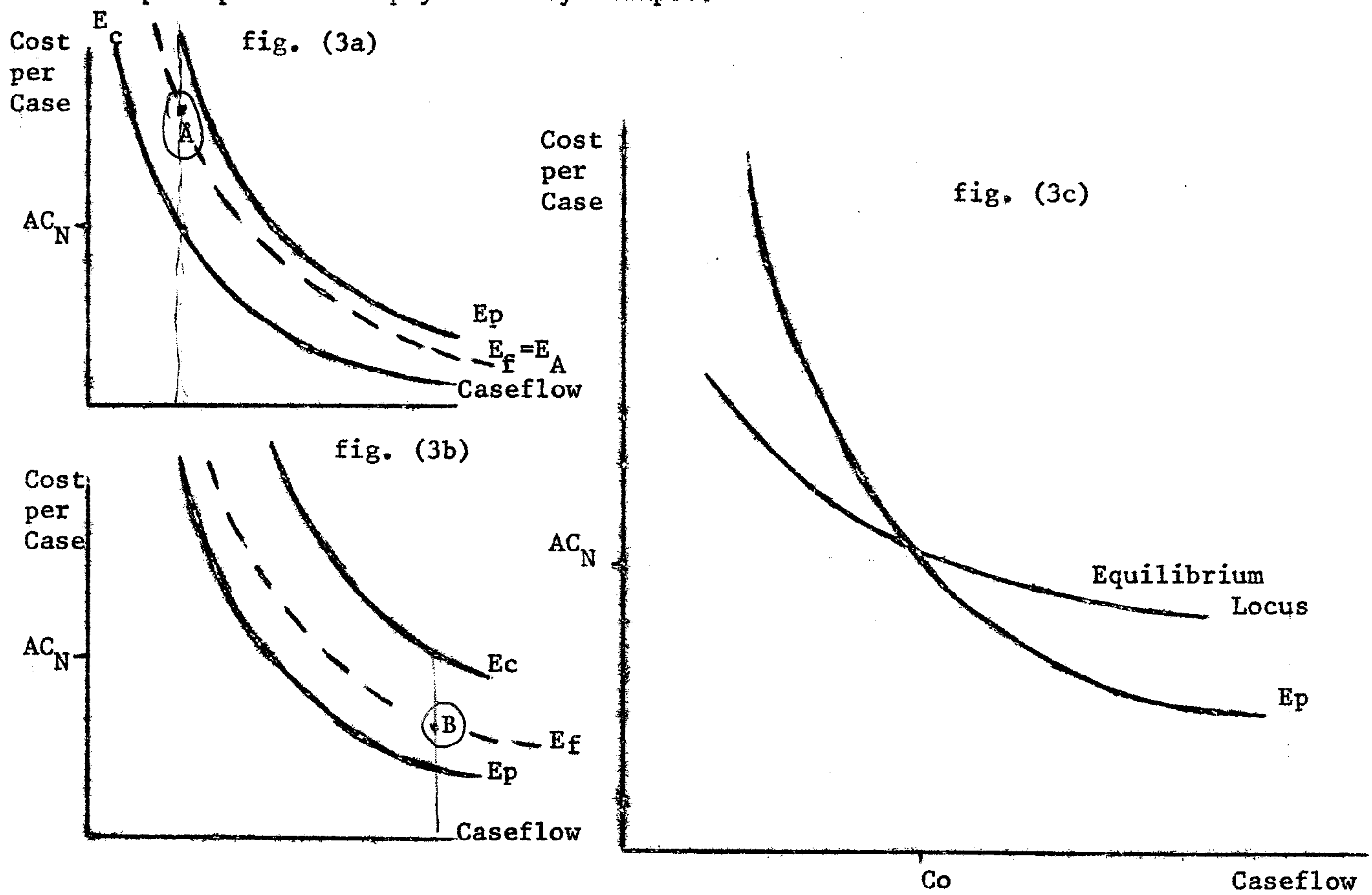

Figure (3a) shows a region with population-based allocation Ep and cost per case and case-flow shown by point (A). The cost per case is above average, and so the case-flow allocation curve Ec, calculated from average cost, lies below and to the left of (A). (A) has been chosen to be exactly on the expenditure curve (E_f) where (E_f) is the formula allocation and equals (1/3)(Ec) plus (2/3)(Ep). Thus, a region at (A) will be given an allocation (E_f) the same as it currently receives, since its current allocation (E_A) must also pass through (A) and have an identical shape to (E_f) for constant expenditure along its length. (A) is therefore an equilibrium position for a region for as long as its population remains constant, since it receives an allocation sufficient to maintain its cost per case and case-flow at current levels.

Figure (3b) shows the same region, with allocation (Ep) as before, but now at point (B). (B) lies above (Ep) but below (Ec), the case-flow based allocation, and has been positioned so as to lie exactly on the (E_f) curve generated by (Ep) and (Ec). It is again an equilibrium situation in that a region with that cost per case and case-flow would continue to receive an allocation sufficient to maintain them.

Figure (3c) shows the locus of equilibrium points such as (A) and (B). The locus cuts (Ep) at average cost per case because a region treating Co cases at average cost would have expenditure equal to (Ep); the formula would indicate an allocation of (Ep) on population grounds and an identical allocation on case-flow grounds, namely that required (and currently received) to treat Co cases at the average cost per case. Once the region, by virtue of its treatment practices or policy, reaches a combination of average cost per case and case-flow that lies on the equilibrium curve, it continues to receive the allocation sufficient to maintain that combination. Any change in this combination must occur within the given allocation, and so moves the region off the equilibrium locus. It is the process by which non-equilibrium regions are returned to equilibrium combinations that constitutes the main effect of the formula. But the significant fact revealed by the equilibrium is that equilibrium is possible at a range of expenditure levels, and costs per case.

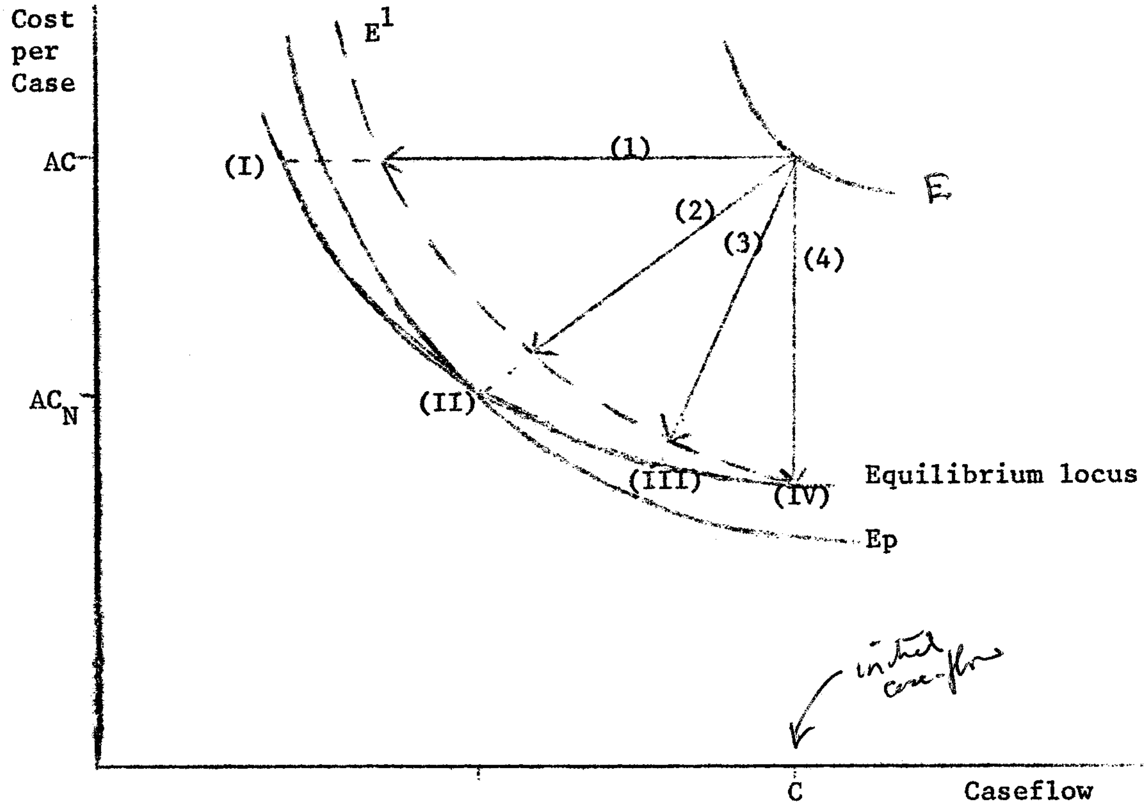

Fig. (4)

NB This is also leeway in varying case-flow by discharge/readmission of same people (i.e. many short episodes of treatment)

Figure (4) shows the (Ep) and equilibrium curves for a particular region which, at the introduction of the formula has case-flow C, cost per case AC, and total expenditure (E). Typically, this is the position of the relatively well-funded regions at present. The formula applied to its case-flow and population indicates an allocation (E^1) for the subsequent period. The final impact of the formula is now dependent on the reaction of the region to the reduction in its revenue - that is, the extent to which economies are achieved by reducing the case flow and the cost per case. Four possible reactions are shown in fig. (4). These each represent the choices made by the medical staff in the face of budget restrictions and, as such, are equally possible, though the reactions which make at least some reduction in both case-flow and cost per case may appear more plausible. Reaction (1) achieves the full economy by cutting case-flow, and maintaining cost per case. (2) and (3) both involve cuts in both elements whilst (4) involves solely a reduction in cost per case, with case-flow maintained at its previous level. Each reaction takes the region down to the required expenditure level, but each also has a different effect on future allocations. (1), (2) and (3) all involve a reduction in case-flow, and so will lead to further reductions in the formula-based allocation. If each reaction function indicates a constant trade-off between case-flow and cost per case that is maintained as a proportionate relation, then the region will, through time, move closer and closer to equilibrium at (I), (II) and (III) for reactions (1), (2) and (3) respectively.[10] Reaction (4) achieves equilibrium directly at (IV) since its case-flow, and therefore its allocation, remains unchanged after the adjustment in the first period of the formula's use. Thus (I), (II), (III) and (IV) represent the equilibrium results of reactions (1), (2), (3) and (4) respectively, and are part of the equilibrium locus. These are simple examples of the effects of the formula. Any region's initial position could be used as a starting point, and any reaction introduced to find the particular path to equilibrium. Since we are unable to predict the reactions of regions to economy, we are obliged to conclude

---

10. The curve is reached by a series of decreasing steps towards it and so, mathematically, is never reached, since the size of the step is reduced as the distance to the curve falls.

that the introduction of the formula will bring regions to the equilibrium locus but cannot determine the final size of their expenditure since the equilibrium locus is consistent with a range of expenditure levels.

We can now evaluate the formula in terms of its effectiveness in achieving 'equality of provision' between regional hospital services by examining its likely effects on regional expenditure, cost per case, and case-flow.

The simplified framework of equilibrium suggests that divergences between rich regions with a high case-flow and high cost per case, and poor regions with a low case-flow and low cost per case will be reduced by the formula, since, assuming the two regions had identical populations, both would move towards the same equilibrium locus, and closer expenditure curves. Regions which continue to operate an above average cost per case are penalised, and regions with below average cost per case rewarded, so that, in equilibrium, the latter have the greater revenue. In this respect, the formula appears to be succeeding. But in fact, as its outcome depends on the reaction functions of regions, it is unlikely to achieve complete equalisation of provision in its present form. Thus, two regions with identical populations would be unlikely to move to the same point on the equilibrium locus, and differences in their funding, due to differences in case-flow, would be likely to continue.

We have highlighted one difficulty of the present structure of the formula, namely its inability to completely eradicate differences in allocation. However, since these allocations are only a proxy for measuring the provision of hospital care, we must examine the situation detailed above to see if it is possible to suggest an equality of provision from unequal funding. We again assume the existence of two regions with identical population numbers and morbidity, both in equilibrium with respect to the allocation formula, but with region (A) treating more cases than (B), at a lower cost per case. If region (A) is simply more efficient, with the quality of treatment identical in the two regions, then clearly the provision of hospital care is better in (A) than (B). However, if the quality of treatment has been reduced in (A) to cut costs, then we are confronted with a trade-off problem, the choice between better treatment for fewer people in (B), and poorer treatment but for more people in (A). The trade-off rate is essentially a choice for the consumer of health care, and is one that could only be made by policy-makers in the light of considerable electoral evidence. The absence of a production function for health services also prevents us from comparing the two regions from the standpoint of efficiency.

Thus, whilst the formula has achieved equilibrium, and can do no more to achieve equality of financial allocation, no other simple criterion exists to facilitate a comparison of the provision of care between regions given unequal expenditures on care. We may conclude that in terms of moving towards an equal allocation of resources, the formula is a move in the right direction, since it moves allocations from richer to poorer regions, but lacks the necessary detail to achieve complete eradication of differences in funding.

To consider the ability of the formula to sustain equality of allocation, once achieved, we will again assume the existence of two regions (A) and (B) with identical population numbers and morbidity, identical case-flows, and also receiving identical allocations after the formula has been fully implemented (That is, both regions are at the same position on the equilibrium curve relevant to their population).

Let us further suppose that in some subsequent period, (t), region (A) becomes more efficient and treats more cases than (B). This situation is shown in fig. (5). (B) shows the equilibrium position they both had reached, and (A_t) shows region (A)'s new position.

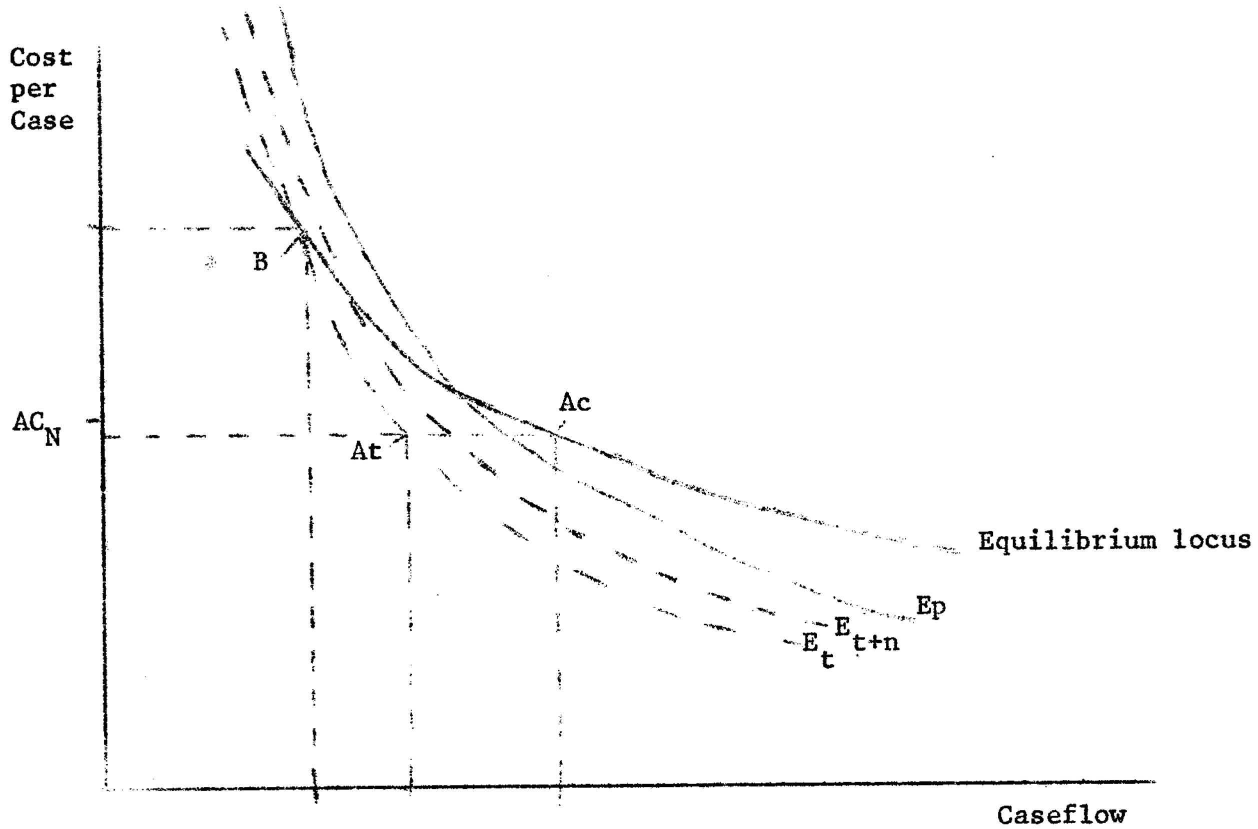

Fig. 5

(A) has moved from (B) along a constant expenditure curve to (A_t). Now, in a future period (t + n), (where n is the time required to process and aggregate hospital data) data from period (t) will be used to recalculate regional allocations. The application of the formula will suggest that region (A) requires additional funds to meet its higher case-load, and will raise (A's) revenue accordingly, from (E_t) to (E_{t+n}). This in turn permits further extensions of case-flow, depending on the reaction of region (A), which may be able to maintain its lower cost per case. By the same step-wise process, (A) will eventually return to the equilibrium locus, and is likely to be below and to the right of the position of (B) (except in cases of extreme reaction to increased revenue, namely a rise in costs per case with case-flow reduced!). (A) now treats even more cases than (B), and the formula has compounded a discrepancy that arose because of an increase in the efficiency of (A). In this way, the formula has produced an undesirable effect because of its treatment of efficiency.

The increase in efficiency by region (A) is rewarded by increased revenue. In this way, the gains from the increased efficiency go solely to region (A) when in fact we may feel that, in the interests of horizontal equity of provision, the gains should be shared out throughout the hospital service. Whilst it might appear that such a measure would discourage efficiency, it is surely more undesirable to disadvantage one group of people relative to another purely because of the failure of administrators to match a rise in efficiency elsewhere, especially as this shortcoming may be a reflection of their capital or staffing levels rather than of their own abilities. Thus, the present structure of the formula mishandles efficiency gains by allocating more funds to more efficient regions. Whilst this will improve the overall efficiency of the hospital service, its effects on equity are counter to the objectives of the formula.

The example above highlights the major difficulty in the present structure of the formula, which is that it uses data from past performance to assess present needs, but ignores the past costs of the performance in assessing the required current allocations. If a region treated its cases satisfactorily at below the national average cost per case in previous years, and given that no changes occur in the organisation or treatment practices of the region, there is no reason why it should not continue to do so, and therefore no need for altering the allocation to it other than to take account of inflation or population changes. Funding on a basis of national average cost per case takes no account of the variations in regional cost per case that are bound to arise as a result of differences in treatment practices, efficiency, and the age and quality of existing capital between regions. Unless the quality of treatment can be shown to be significantly lower in regions with a lower cost per case, there is no reason to raise their allocations on case grounds. Furthermore, it cannot be assumed that raising the allocation to regions deemed to have a lower quality of treatment will raise this quality, since its use may well be left in the hands of the medical staff who have lowered the quality in order to raise the case-flow and reduce the waiting time of patients. In short, the absence of a clear production function for hospital services again prevents us from assessing the relative desirability of a shorter stay, the trade-off between waiting time and treatment quality, or indeed of one treatment method over another. However, we are not prevented from drawing attention to the logical flaw of using previous performance as a measure of future performance while failing to use previous costs per region as a measure of future costs. Nonetheless, it is not valid either to resort to the other extreme, that of taking the previous allocations as constant except for population adjustment. Such a scheme would have the obvious fault that it would fail to provide incentives towards efficient use of regional resources. It would also enshrine past allocations as the rule and therefore fail to relate to the objectives of equality of provision.

The answer to the dilemma as to which allocation rule to choose is to choose neither, since the two are equally unsatisfactory, but to resort to a form that removes the influence of historical accidents in past performance and uses instead some objective measures relating to the needs of each region. The weaknesses of the present formula arise because it is only partially based on the objective needs of the region through the use of adjusted population. The other factors depend too heavily on the practices, efficiency, and capital of each region to escape the consequences of historical accident. The formula has the merit that it makes most of the allocation process explicit, but nonetheless fails in its attempt to be a comprehensive method for resource equalisation. However, the analysis above suggests that it may achieve a considerable reduction in the extent of unequal funding, and further refinement may not be required until the gaps in allocations have been reduced.

An alternative approach to the allocation problem may well be to analyse the population, bed-stock, and case-flow, as in the current formula, but to use them in an alternative framework. The essence of the population element in the formula at present is that it provides an indication of medical needs, and on such a basis compares the required with the actual allocation. It may well be possible to use the beds per region as an alternative index of current provision, rather than need, and on that basis compare the beds required (as calculated by applying national average bed provision to the adjusted population), with actual bed-stocks to obtain the necessary bed-stock adjustment. The formula could treat case-flows in a similar manner, again deducting the difference between actual and required case-flows. The various elements might then be combined by taking population-based allocations as the norm, and adding compensating sums to counter the effects of bed-shortages, and case-flow deficiencies until such time as these are remedied by the building programme. Furthermore, examination and standardisation of the specialty bed proportions and staffings would be a further step towards solving the difficulties of comparatively underprovided regions. These are put forward as possibilities for a formula, rather than as a firm structure. The present formula is a valuable step towards equalizing the provision of revenue to hospital regions, but to be fully successful, any such formula must be more closely connected to regional needs than is the current model.

Appendix

Table 1

National Hospital Allocation	£1,000,m	(1)
National Population	50m	(2)
National Average cost per bed	£3,000	(3)
National Average cost per case	£100	(4)
Regional Allocation (in previous period)	£150m	(5)
Regional Population	5m	(6)
Regional Bed-stock	40,000	(7)
Regional Case-Flow	1,100,000	(8)

Allocation on Population grounds = (6)/(2) x (1)
= 5m/50m x £1,000,m.
= £100m.

Previous Allocation = £150m

∴ Population-based Adjustment = -£50m

Allocation on Bed-Stock grounds = (7) x (3)
= 40,000 x £3,000
= £120m

Previous Allocation = £150m

∴ Bed-based Adjustment = -£30m

Allocation on Case-Flow grounds = (8) x (3)
= 1,100,000 x £100
= £110m

Previous Allocation = £150m

∴ Case-flow based Adjustment = -40m

Combined Adjustment = ½ x Population adjustment
+ ¼ x Bed-stock adjustment
+ ¼ x Case-Flow adjustment

= ½ x -£50m + ¼ x -£30m + ¼ X -£40m
= -£25m - £7.5m - £10m
= - £42.5m

∴ Adjustment in Year 1 of formula = -£42.5 x 0.1
= -£4.25m