How many patients should A GP Have? A review of policies, Concepts and data

J.R. Butler

H.S.R.U. Report no 37

October 1979

"We recommend that before a maximum or minimum list size is adopted, considerable research on this important question should be undertaken."

Report of the Royal Commission
on the National Health Service

For a summary of this report, readers are recommended to turn to the following sections:

CONTENTS

contd...

ACKNOWLEDGEMENTS

I am very grateful indeed to my colleagues in the Health Services Research Unit, Ken Dawes, Rose Knight, Robert Lee, Myfanwy Morgan and Michael Warren, for their constructive criticisms of an earlier draft of this report. They naturally bear no responsibility for the imperfections and shortcomings of this final draft.

The clerical work associated with the production of this report was done meticulously by Lynn Browne and Isobel Butler, and I am, as always, indebted to them.

John Butler
October 1979

INTRODUCTION

This report has been written in response to the question posed by the Department of Health and Social Security: what is a reasonable list size for a general medical practitioner to provide an adequate level of care? The report does not contain the results of any original research; rather, it attempts to draw upon the existing literature in order to build up a coherent basis from which further research might proceed. Some possible topics for future investigation are set out at the end of the report, but the execution of those investigations will naturally depend upon the Department's views about their importance and value.

The report tries to do two principal things. First, it attempts to locate the question within a framework of ideas that generates practical questions for empirical investigation. Although (for reasons discussed below) the question about a reasonable list size for general practitioners has important implications for several aspects of health services policy, it is couched in terms that do not readily indicate the sort of information needed to answer it. Further questions need to be considered, and further concepts defined, before material can be assembled in ways that might lead to an answer. The report tackles the task of clarifying the meaning (or meanings) of the question in a number of different ways. First, it reviews some of the more influential opinions that have been offered in the past twenty-five years about the size of list that has been regarded as reasonable; second, it examines critically the foundations upon which these opinions appear to have been based; and third, it summarises the conceptual approaches to manpower shortage that are to be found both in official reports and in the academic literature. No uniquely 'correct' interpretation emerges from these varying sources, but, based upon them, a framework is proposed that is intellectually defensible, is related to issues and concepts that appear to be of considerable contemporary interest, and is productive of specific questions for empirical investigation. It naturally has many inadequacies, but it is hoped that it offers a reasonably coherent basis upon which future studies might build.

Having identified a set of empirical questions from the framework, the second principal aim of the report is to assess the extent to which they can be answered from the existing research literature. Many studies of general practice have been published, and although they are of variable quality, they merit careful re-examination in the light of a coherent conceptual framework before embarking upon new fieldwork studies. About half of the report is taken up with this review of the research literature, and it goes some way towards

answering the detailed questions and also identifying gaps that may be worth filling through future studies.

In spite of the difficulties involved in making sense of the notion of a reasonable list size for general practitioners, it is an important notion with several implications for policy. First, the judgement about a reasonable list size carries obvious and substantial implications for the future supply of medical manpower. As will be seen, post-war opinions about the number of patients for whom each GP can provide an adequate standard of care have varied from about 4,500 to 1,800 or less, and this represents a substantial range in the target supply of family doctors. The Department's recent discussion paper on medical manpower during the next twenty years illustrates the magnitude of this difference (DHSS, 1978). A reduction in the average list size of unrestricted principals in Great Britain in 1977 from the actual figure of 2,275 to a target of 2,000 would have required an additional 2,800 practitioners, and a further reduction to 1,800 would have required 5,500 extra principals. The increased cost of such reductions would be reflected not only in the training and servicing of the additional doctors but also in the enhanced levels of capitation payments required to maintain existing levels of income with fewer patients per doctor. Conversely, an average list of 4,500 could have been sustained with 12,300 fewer principals than were actually practising in 1977. Although there would be obvious savings of expenditure on doctors' incomes and expenses if average lists rose to this level, there would be offsetting additional expenditure on the other categories of staff that appear to be needed to enable doctors to cope adequately with this number of patients.

The implications, however, are wider and more complex than this, for it will be argued that the judgement about a reasonable list size cannot properly be made without reference to the context in which care is provided. The number of patients for whom a GP can provide adequate care will, for example, be heavily influenced by the range and content of the care he is expected to provide. Various proposals have been made in recent years to extend the content of the GP's role in child health, rehabilitation, screening, the staffing of community hospitals, the follow-up of hospital outpatients, the care of early-discharge inpatients, etc.; and whilst few of these proposals have yet had a substantial impact on the workload of most doctors, the possibility of significant extensions to the GP's role must influence, and be influenced by, opinions about an acceptable list size.

In the context of the individual practice, the judgement about a reasonable list size should also take a conscious account of the scope for substituting the doctor's time by other, less costly resources. The development of primary

health care teams is an obvious example of the way in which the careful and sensitive delegation of tasks from the doctor to other health care workers can release additional time for the doctor to extend his list size or to use in other ways. Urgent questions still remain about the most efficient size and composition of the team (DHSS, 1978; Marson, et al, 1973), but as answers begin to evolve, the judgement about a reasonable list size for the team will carry important implications for nursing and other resources as well as for medical manpower.

Yet another policy issue to which the concept of a reasonable list size is relevant is that of the distribution, as well as the supply, of general practitioners. The average list sizes used by the Medical Practices Committee as the basis for the classification of practice areas have changed little since 1952, in spite of major changes in the scope and organisation of general practice, and the average list size that is normally used to indicate a shortage of doctors (2,500) has not changed at all. Among the reasons for this conservatism are the undesirability of frequent changes in the criteria for classifying areas, especially when they affect remuneration, and the impracticability of lowering the criterion of a designated area in the absence of any substantial increase in the total supply of GPs. An additional reason, however, may simply be an uncertainty about the rational basis for either increasing or decreasing the average list size at which areas become designated. Because there has been little systematic analysis of what a reasonable list size actually means, it is understandable that 'under-doctored areas' have been defined in relation mainly to the national average, rather than to any intrinsic judgement about the point at which standards of care are seriously threatened or doctors are unreasonably over-burdened. Yet if the arguments came to be accepted of those who claim that list sizes of 4,000 are not unreasonable, the designated areas would disappear at once. The matter is particularly important in view of the Department's expressed desire to 'remedy persistent shortages of personnel in locations where they occur by encouraging a better distribution of manpower' (DHSS, 1976, para. 3.6), and of the establishment of a working party to review, inter alia, the criteria by which under-doctored areas might be identified (Review Body, 1978, para. 42).

Finally, it will be argued that the concept of a reasonable list size inherently involves questions about standards of care in general practice. Notwithstanding the absence of any agreed definition of good standards, it seems reasonable to suppose that some element of compromise must exist between the number of patients for whom a GP provides care and the standard of care that he is able to give. All else being equal, the GP with a list of 1,500

patients will be capable of providing more and/or better services to his patients than one with a list of 3,000: he may, for example, see his patients more frequently, or he may spend more time in consultations with them and carry out more (or more thorough) diagnostic or treatment procedures. In either case it might be hypothesised that the provision and perhaps the overall quality of care will be higher in the practice with the smaller list. It is this hypothesis that appears to underlie much of the current concern within the medical profession about the current inadequacy of manpower resources. The Secretary of the BMA has noted the inability of doctors to 'provide for their patients the standard of service they would wish and for which they have been trained' (British Medical Journal, 1978a, p.841), and the constraints of time were explicitly identified as a source of low standards in a motion debated at the 1978 annual representative meeting of the BMA (British Medical Journal, 1978b, p.449). In proposing the motion, the speaker was reported as saying that in the health service money was time: money could buy time to see patients and their needs, and everyone was 'sick and tired of conveyor-belt medicine'. It does not, however, necessarily follow that a substantial reduction in list sizes is a reasonable policy objective, for it will be argued that smaller lists do not necessarily lead to higher standards of care, and in any case a realistic definition of a reasonable list size should embody some assessment of the point at which it ceases to be worth using additional resources to secure further gains in standards of care. The cost of producing and servicing the 5,500 additional principals required to reduce average list sizes in Great Britain to 1,800 may simply not be regarded as justified by the resulting gains in standards of care. Such judgements are ultimately matters of social policy, but they bring together important issues of quality, resources and output around the theme of the reasonable list.

In sum, there are significant policy issues that impinge upon the concept of 'areasonable list' and that justify an attempt to define and measure it. The Royal Commission on the National Health Service (1979) appeared to find it sufficiently important to recommend that, before a maximum or minimum list size is adopted, considerable research should be undertaken on an optimum range of list sizes, and a similar case has been put from within the profession. An editorial in the Journal of the Royal College of General Practitioners in 1972 observed that 'the number of patients a doctor can look after properly is a key index of care. As the basis of medical care ... is fast centring on the primary physician, the population that he can care for now interests both doctors and governments ... The national average list size has always been about 2,500, ... and as a result, the status quo of the list size has come to

to be accepted, almost unquestioned. Much of the current organisational planning for and in general practice is being devoted to methods of improving the delivery of care with the implicit assumption that 2,500 patients per doctor is about right. Elaborate costings, carefully contrived attachment ratios, plans for future vacancies, teaching requirements, and a host of other projections are being constructed; yet all rest on this one fundamental assumption. Surprisingly little work has been done to test if 2,500 is indeed the optimum number. How many patients can a general practitioner look after?' (pages 491-2).

Five years later the Journal (1977) returned to the theme. 'Despite the fact that the number of patients that an average general practitioner can properly look after is one of the most crucial parameters in the whole field of primary medical care, remarkably little work has been carried out on it' (page 3). This report attempts to set out a basis for such work.

TRENDS IN OPINIONS ABOUT A REASONABLE LIST SIZE

There has been no shortage of opinions and pronouncements over the last twenty-five years about a reasonable list size for general practitioners, and it therefore seems sensible to begin an exploration of the meaning of the concept by turning to the sources of these pronouncements. The aim in this section is to identify the sources and to examine trends in the opinions. The next section then looks more critically at the foundations upon which these opinions seem to have been based.

Although a number of opinions have been expressed by official committees and influential groups, they are difficult to summarise because of inconsistencies in the handling of related but distinct concepts. One distinction to be observed is between a reasonable and a maximum list size. General practitioner principals are not permitted to have more than 3,500 patients on their list (or an additional 2,000 if an assistant is employed), although as will be seen, suggestions have been made from time to time that this maximum should be reduced. However, although the maximum permitted list size is rarely regarded as a reasonable or desirable or target list size, some reports and commentators have failed to distinguish clearly between them.

A second distinction to be observed is between a reasonable list size for individual practitioners and a reasonable average list size for the country as a whole. Average and individual list sizes will coincide only in the Utopian circumstances of a perfect distribution of practitioners in relation to population. Where imbalances exist between the distribution of doctors and patients, an average list size will conceal a proportion of individual practitioners with actual lists in excess of (as well as below) the average. It is important under these circumstances to specify whether the target list size is the average or the individual, and to quantify the magnitude of the discrepancy between them. An example of this is found in the Eighth Report of the Review Body on Doctors' and Dentists' Remuneration (1978), which quoted the claim made by the profession about the excess size of the present average list. 'They (the profession) told us that, in their view, no individual doctor should be responsible for more than 2,000 patients: this implied a target average list size of 1,700 in the future' (para. 43, emphases added). It is, however, unusual to find the distinction made as clearly as this.

With these reservations in mind, this section summarises opinions that have been offered over the last twenty-five years about a reasonable list size for GPs. It is, plainly not exhaustive, but it does attempt to include the opinions of individuals or groups that might be expected to carry weight and authority.

Medical Practices Committee

Statutory responsibility for determining the number of practitioners required in an area (and hence, by implication, the list size that is indicative of the requisite supply of manpower) was placed by the 1946 National Health Service Act upon the Medical Practices Committee (MPC). Section 34(2) of the Act required the Committee to 'secure that the number of medical practitioners undertaking to provide general medical services in the area of different Executive Councils or in different parts of those areas is adequate.' The Committee responded in 1949, following the first submission of information on manpower needs by executive councils, by classifying districts as 'needy', 'open', 'doubtful' or 'closed'. The classifications appear to have been based upon the average list sizes within districts, the qualifying list size for a 'needy' area being an average of 3,000 or above. In 1952 the General Medical Services Committee expressed its concern to the MPC about the excessively high list size used as the indicator of the need of a district for more doctors, and later that year, following the Danckwerts award, the Committee revised and regularised the criteria for the classification of districts (or practice areas as they were now called). Under the new system areas with average lists in excess of 2,500 were regarded as designated and doctors were strongly encouraged to initiate practice in them. An Initial Practice Allowance was introduced to ease the financial difficulties of doctors setting up new practices in these areas, and in 1966 further financial incentives were made available. An average list size of 2,500 still remains the basic (though not the only) criterion for designating a practice area.

Willink Committee

The Medical Practices Committee is concerned principally with the average list sizes of areas, not with the actual list sizes of individual doctors. By implication, the combination of the maximum permitted list (3,500) and the average list size for the designation of an area (2,500) defines the upper limit of the acceptable range of individual lists. This was substantially the view taken by the Willink Committee (Central Health Services Council, 1957). In estimating the shortfall in the availability of GPs, the Committee regarded the current (1955) national average list size of 2,283 as 'not unreasonably high' (para. 34), but it was concerned about the wide variations between different parts of the country, and the slow rate of improvement in manpower distribution. 'A problem of maldistribution thus remains, and we thought that we ought to budget for an increase in the total number of general practitioners large enough to effect an early and material improvement. A realistic estimate of

likely expansion over the next few years would be the number of additional principals required to reduce the average number of patients per list to 2,500 in non-rural areas, and similarly to 2,000 in rural areas' (para. 34). The Committee commented that, in choosing these two figures, it took note of the Medical Practices Committee's standard of 2,500 as the criterion of an under-doctored area, but in rural areas this was reduced to 2,000 because of the wide scatter of patients. On this basis, the extra number of GP principals needed in England and Wales at July 1955 was calculated to be 600 (assuming that they were optimally distributed). Interestingly, the Committee appears to have given no thought to the means by which these extra principals would be attracted to the most needy areas. The Committee also refused to recommend any reduction in the maximum permitted list size (3,500), even though representations to that effect has been made by several organisations. The ostensible reason given for this by the Committee was that it lay outside its terms of reference, but it did quote with approval the conclusions in the Cohen Report (Central Health Services Council, 1954) that the upper limit of 3,500 patients, with an additional 2,000 where an assistant is employed, 'fairly reflects present needs and conditions'. The Willink Report commented that 'none of the evidence we have received leads us to believe that this is less true now than it was at the time the (Cohen) Committee drew up its report' (para. 37).

Gillie Committee

A different view about the maximum list size was taken six years later by the Gillie Committee (Central Health Services Council, 1963). Although noting the impracticability of any immediate reduction in the maximum size of a doctor's list, the Committee emphasised that 'without exception, every principal we met has regretted the impossibility of achieving a satisfying standard of work with the present maximum under existing working conditions' (para. 119). However, in spite of favouring a reduction in the maximum list size for individual doctors, the Gillie Committee appeared to imply that the current average list size was, if anything, on the low side. After noting the variations that inevitably exist between list sizes, the Committee commented that 'the average list is, however, only about 2,300 compared with a maximum of 3,500, and there is clearly scope for a better distribution of doctors' (para. 120, emphasis added). The Committee concluded that the ideal number of persons in the care of each doctor, and the maximum that it is reasonable to fix, must be kept under review.

Charter for the Family Doctor Service

A further review was indeed made by the British Medical Association (1965)

in its Charter for the Family Doctor Service. The wording of the Charter suggested a concern by the BMA with both a maximum and also a target list size for individual doctors. With regard to the maximum, the Charter commented that 'it is not in the interests of the patients that any doctor should have to care for a list of 3,000 or over' (para. 12). The Charter's comments about a target list size were a little less precise. 'There must be a reduction in the excessive number of patients for whom many doctors have to care. As more doctors enter general practice the maximum size of lists will be progressively reduced ... It is difficult to predict, but we would not regard a maximum list of 2,000 as an unreasonable target' (para. 3(vi)). A similar, but rather clearer, position was adopted by the Association in endorsing the view expressed in the GMSC's New Charter Working Group in 1979 that 'the average list size for GPs should be reduced to 1,700 patients (consistent with a maximum of about 2,000)' (British Medical Journal, 1979, page 565).

National Institute of Economic and Social Research

The opinions expressed in the 1965 charter were substantially repeated a year later by Paige and Jones (1966) in their wide-ranging study of the health and welfare services in Britain in 1975. They suggested that the maximum permitted list size should be reduced to 2,500 patients, quoting in support the Charter's view that it would not be unreasonable to aim for a target list of 2,000. Apart from observing that 'most people would regard it as a reasonable maximum for a proper family doctor service', Paige and Jones offered no analytical justification for their suggestion. In arriving at a judgement about a desirable future average list size for GPs nationally, Paige and Jones took account of the consequences of reducing the maximum list size to 2,500, and also (though they did not indicate how) of the increased demands on GPs resulting from an ageing population and the growth of community care for the handicapped and mentally ill. The result of these calculations indicated that a reduction would be needed in the average number of patients per GP in Great Britain 'from the 1962 level of about 2,050 to about 1,775 by 1980' (page 133). The total supply of doctors required to meet this level of demand was estimated as 33,700, or an increase by 1980 of 35 per cent over the number in 1960.

Todd Commission

In 1968 the Royal Commission on Medical Education repeated the exercise of estimating the future supply of and demand for medical manpower. The

Commission presented both long- and short-term estimates. The short-term (1966-75) estimate of the demand for general practitioners was based on a different assumption from that of Paige and Jones. Whereas the latter had, by implication, regarded the average list size in Great Britain in 1962 as too high, the Todd Commission regarded it as about right in the short-run. After noting that the lowest post-war ratio of population to principals had occurred in 1961 (2,180:1), the Commission assumed that this ratio 'provides a reasonable guide to what it should be in 1975'. (The discrepancy in the 1961-2 population/doctor ratios between Paige and Jones and the Todd Commission arises from the fact that, although both sets of data relate to Great Britain, the former includes assistants and trainees, whilst the latter is restricted to principals only.) The Commission estimated that, in order to achieve this target, and also to restore the relationship between the number of assistants and trainees and the number of principals that had obtained in 1953-5, an additional 3,250 principals and 1,700 assistants/trainees would be required between 1965 and 1975. If this increase could be achieved, and with some improvement in the distribution of practice sizes, the Commission noted that 'it should be possible to ensure that the maximum patient lists of principals should not be greater than 2,500 (2,000 in rural areas)'. And it added: 'it has been suggested by the Ministry of Health ... that upper limits of this order are desirable, and these levels are consistent with other evidence submitted to the Commission' (Appendix 12, Annex para. 9).

In approaching its estimate of the long-term need for manpower, the Todd Commission acknowledged that the technique it had used in making its short-term estimate (that is, extrapolating on the assumption that the lowest achieved population/GP ratio since 1948 was a reasonable target for the future) was unsatisfactory. Instead, the Commission relied heavily upon long-term trends in the ratio of <u>all</u> doctors to population, and the consistencies contained within them. It first plotted the growth in this ratio in Great Britain between 1911 and 1961, and found a remarkably regular annual growth of about 1¼% per year. The Commission then continued the same extrapolation into the future, and found that in 1975 it passed 'very close' to the short-term estimate for that year already produced by the earlier procedures. On the basis largely of this evidence, the Commission concluded that 'the needs of the future will not be met by an annual growth rate of less than 1.5% in the doctor-population quotient' (para. 338). Although this long-term estimate made no distinction between general practitioners and hospital doctors, an application of the 1.5% cumulative growth rate to the ratio of principals to population would result in a decline in average list sizes from the short-term

target of 2,180 in 1975 to about 2,021 in 1980 and 1,738 in 1990. The Commission presumably accepted this as a reasonable target.

Harvard Davis Committee

Three years after the Todd Report, the Report of the Harvard Davis Committee on the Organisation of Group Practice (Central Health Services Council, 1971) was published. The Report is relevant in the context of a maximum or desirable list size because of its discussion about the effect which a well-constituted group practice might have upon the number of patients for whom a GP can adequately care. The Harvard Davis Committee, after reviewing the evidence, rejected the view that group practice increases the number of patients for whom the GP can accept responsibility. Indeed, the addition of a nurse and health visitor may actually add to the doctor's workload by virtue of the hidden needs they may uncover. The Committee concluded that the introduction of group practice tends to redistribute the workload, thereby enabling the doctor to spend more time with the individual patient. The majority of general practitioners, in the Committee's view, 'consider that the time they are able to give to each patient is inadequate' (para. 45), and hence the advent of the team is to be viewed as a gain in the quality of medical care rather than an increase in the number of patients to whom care can be given. In the light of this, 'we think it is unlikely that a general practitioner would be able to look after very many more persons than the present average list size of 2,500 persons' (para. 46). Accordingly, the Committee regarded an optimum group practice as one consisting of five or six doctors, together with the nurses and supporting secretarial staff, and responsible for a population of approximately 15,000 people.

Review Body on Doctors' and Dentists' Remuneration

The Review Body has several grounds for interest in the concept of a reasonable list size. One is that it has, on occasions, seen it as part of its duty to recommend awards that will encourage desirable patterns of recruitment to the profession. In many of its reports, the Review Body has commenced the chapter on general medical services by evaluating trends in population/doctor ratios, and the 1970 report endorsed the view of the Todd Commission that a further 500 doctors were needed each year in general practice (para. 63). By implication, the Review Body accepted the Commission's judgement that existing list sizes should be substantially reduced in the long run. The Review Body is also concerned with securing appropriate financial

rewards for GPs when workloads become 'excessive', and has on occasions regarded list sizes as one among a number of indicators of the volume of work. However, the Review Body has never (as far as we can trace) explicitly defined a reasonable list size, partly perhaps because it has never actually been necessary, and partly perhaps because the Review Body has tended to adjudicate arguments from outside rather than initiate new opinions from within.

The accounts given in the Review Body's Reports of the evidence it has received make it clear that, over the years, the British Medical Association has consistently regarded the workload of general practitioners not only as being unreasonably high with existing list sizes, but also as growing, thereby justifying the recruitment of more GPs and an increase in fees and allowances. In 1966 the Review Body commented that 'there remains the very strong impression, certainly in the minds of the doctors themselves and fixed strongly in our minds by their representatives, of a growing burden on doctors ... to a point that many doctors are beginning to find insupportable' (para. 68). In 1968: 'They (the profession's representatives) pointed out that the civilian population, expressed as a ratio to the number of GPs, had risen to 2,207 in 1966 compared with 2,172 in 1965 and 2,053 in 1962. They stressed that, quite apart from the rise in population, additional demands were being made on GPs as a result of the rapid technical advances in medicine' (para. 53). In 1970: 'The BMA argued that, irrespective of the size of lists, there was greater pressure on general practitioners because of changes in hospital policy and in medical techniques' (para. 29). In 1975: 'The profession has suggested that this difference (between the rate of manpower expansion in general practice and in the hospital service) is too great, and that there is a need to increase the number of general medical practitioners. They would like to see the average list size reduced from the present level of about 2,350 patients to 2,000' (para. 36). And in 1978: 'The profession have described to us the ways in which the workload in general practice has changed and how, in their view, it has increased in recent years ... They told us that, in their view, the present average list size of 2,294 was too large ... and that no individual doctor should be responsible for more than 2,000 patients' (para. 43).

The views of the Health Departments, as recounted by the Review Body, have been more circumspect. Whilst acknowledging that there is no agreed optimum list size (Review Body, 1975, para. 36), the Departments have generally been less willing than the profession to concede the unreasonable size of existing lists. In 1966, for example, the Departments were reported to have argued that 'a doctor may up to a point be able to deal with more patients without loss of efficiency' (para. 70). In 1968 the Departments thought that the

deterioration in doctor/population ratios 'was not sufficiently marked to cause any significant increase in the average workload falling upon general practitioners' (para. 54). In 1975 the Health Departments informed the Review Body that current manpower targets, if they were met, would reduce the average list size to around 2,250 patients - still in excess of the BMA's target.

The Review Body itself, as judged by its comments and recommendations, has generally interpreted a decline in average list sizes as a desirable trend, and an increase as undesirable. It tended to accept the BMA's contention that workloads were increasing in the 1960s, but it has been more sceptical in the 1970s. In its 1966 Report, for example, the Review Body noted that 'the unattractiveness of general practice is primarily a matter of conditions and workload: doctors find that they are not able to practise good medicine in the conditions and under the strains of general practice. The remedies for these problems lie (inter alia) in more doctors ....' (para. 77). And in 1970: 'Our conclusions are that the evidence, including the figures for average list size, suggest that the workload of general medical practitioners has been growing in recent years, principally because of the new techniques in medicine and the extra burdens referred to (by the BMA)' (para. 35). By 1972, however, the Review Body felt that the trend in list sizes was 'improving to some extent' and did not justify 'any exceptional measures' (para. 64); in 1975 the Review Body felt it had 'no clear evidence to show that the workload of the average general medical practitioner has increased materially over the past few years' and that it was therefore 'not in a position to judge whether an accelerated reduction in the list size as suggested to us by the profession is justified at present' (para. 37); and by 1978 it was noting that 'we have no material evidence to show that the overall level of workload has changed to a significant extent over recent years' (para. 44).

Opinions of doctors

Several of the sources quoted above make reference to the views of the BMA about a reasonable list size, and they also quote the feelings of ordinary doctors about the pressures to which their workloads subject them with their current list sizes (see the Gillie Report, para. 119 and the Harvard Davis Report, para. 45). In addition, evidence is available from surveys of general practitioners about the range of opinions within the profession concerning a reasonable list.

Cartwright (1967) reported the results of a survey carried out in 1964 of 422 GPs in England and Wales in which doctors were asked to state the ideal number of patients they felt they could look after under present practice arrangements. Almost half of the doctors (46%) identified a list of 2,000-2,499 as ideal; 29% selected a list of 1,500-1,999; 14% regarded 2,500-2,999 as ideal; and the remainder opted for list sizes of either less than 1,500 (8%) or more than 3,000 (3%). The mean ideal list size was about 2,100 although this is only a crude estimate calculated from the publish ed data (page 16). Cartwright repeated the survey in 1977 (Cartwright and Anderson, 1979), and although the results published so far have not contained the replies to the corresponding question, they do indicate that rather more GPs regarded themselves as overworked in 1977 than had done so in 1964 (27% compared to 20%).

In 1973 the Consumer's Association conducted a survey of GPs based upon the earlier work of Cartwright (Lavers, 1978). No details are given of the population or the response rate; all that is known is that the results are based upon the replies of 112 doctors. The question about an ideal list size appears to have been replicated from Cartwright's survey. The results showed that 39% of the GPs thought lists of 1,500-1,999 were ideal; 33% selected 2,000-2,499 as the ideal; 14% regarded a list of 2,500-2,999 as ideal; and the remainder opted for list sizes of either less than 1,500 (10%) or more than 3,000 (5%). The mean ideal list size, calculated from the published data (page 162), was about 2,000. Taking the results at their face value, there appears to have been a slight downward revision between 1964 and 1973 in GPs' views about an ideal list size, although as Lavers points out, over half the respondents in each survey selected an ideal list in excess of the BMA's target (2,000), and the choices were made without conscious reference to the financial implications of them.

Mechanic (1974) carried out a postal survey of a random sample of GPs in England and Wales in 1966, and received 814 replies (a response rate of 60%). No direct question was asked about ideal list size, but respondents' perceptions about various aspects of their work have been classified according to their actual list sizes (pages 98-99). The proportion of doctors reporting a 'very serious problem' with the number of patients in their practice was 7%, 8% and 9% respectively among doctors with less than 1,500, 1,500-1,999 and 2,000-2,499 patients, but rose to 17% among those with lists of 2,500-2,999, 27% among those with patient loads of 3,000-3,499, and to 38% of those with lists of 3,500 or more. These results indicate that subjective feelings about an unreasonable workload (as indicated through perceptions of the severity of the problem of large lists) may be conditioned by the average list size

existing in the country. Fewer than one in ten of doctors with individual lists below the average regarded their list size as a serious problem; but the proportion increased to two-fifths of doctors with list sizes of 1,000 or more above the national average. Such data are consistent with those of Cartwright and Lavers, although they are not directly comparable.

Two further expressions of opinion by general practitioners, based upon detailed studies of workloads in general practice, are important because of their dramatic divergence from the general trend of opinion. Fry (1972) reviewed a span of 21 years' work in a two-man general practice, located in a middle-class South-East London suburb, and supported by a nurse, health visitor, midwife and secretary-receptionist. The practice contains some 9,000 patients, giving an average list for the two doctors of almost twice the national average. In his paper, Fry contended that 'in this particular practice it is possible for two practitioners to provide sound care for a population of over 9,000 patients in ways that, apparen tly, are satisfactory to both patients and doctors' (page 527); and he commented on the implications of this for the future supply of medical manpower. 'The major question is how many general practitioners are needed in the future? Have we perhaps a surfeit now? Should we be trying to induce more and more young doctors to enter general practice? These are important national and public issues .... The results merit urgent national studies to test the hypothesis that perhaps there are already enough general practitioners' (pages 527-8).

Marsh and McNay (1974a) reported the results of one year's detailed recording of the workload of a general practitioner in the Teesside conurbation in 1972. Working as one of five partners in a team comprising two state-registered nurses, four receptionists, two filing clerks, one administrative secretary and one research secretary, the GP provided care for 3,137 patients during the year. The analysis of consultation patterns during the year 'showed that even in this area of high morbidity and mortality the workload was very small' and that 'by delegating work to a team of trained para-medical workers, by increasing the proportion of personal medicine, and by engaging the co-operation of his patients, the general practitioner reduced his workload considerably, without any apparent reduction in standards of care' (Marsh and McNay, page 315). Elaborating these results, Marsh and Kaim-Caudle (1976) commented that 'the average list size of 2,400 patients may well become too small to occupy the time of the established general practitioner ... Some doctors will want to spend time teaching new entrants to general practice, and others will wish to pursue clinical or operational research. Some may wish to undertake more work in hospitals or increase their commitments in preventive

care and health education. But there is also the distinct possibility that in future general practitioners will consider it reasonable to have longer lists than are acceptable at present and indeed need these in order to satisfy their clinical interests. The manpower requirements of general practice in a team setting will be different from those in the past for all members of the team, but for doctors themselves it seems that the requirements will be lower than is generally thought at present' (page 66).

Fry's and Marsh and Kaim-Caudle's arguments have not gone unchallenged, and McGregor (1973), Bain and Haines (1974) and Price (1974) among others have reaffirmed the seemingly more widespread view that a list of 2,400, far from being too small to satisfy a GP's clinical interests, is still unreasonably large. The counter-argument has been put with notable vigour by Tudor Hart (1971), dismissing Fry's arguments as 'dangerously complacent'. Tudor Hart's argument is that the volume of real morbidity in areas of the country such as the Welsh mining valleys is so great that any improvements in the efficiency of general practice must be used to increase the standards of care, not to increase the numbers of patients on GPs' lists. Commenting, for example, on the claim that large list sizes could result from increases in productivity through rationalisation and devolution, Tudor Hart wrote: 'Of course, much devolution and rationalisation of this sort is necessary, not to cope with rising numbers but to make general practice more clinically effective and satisfying, so that people can be seen less often but examined in greater depth. If clinically irrelevant work can be devoluted or abolished, it is possible to expand into new and valuable fields of work such as those opened up by Balint and his school, and the imminent if not actual possibilities or presymptomatic diagnosis and screening, which can best be done at primary care level and is possible within the present resources of NHS general practice.' (page 408).

Summary

Once the Medical Practices Committee decided in 1952 that an average list size in an area of 2,500 signalled the need for more practitioners, the figure of 2,000-2,500 has repeatedly been selected as a reasonable list size for a general practitioner to provide an adequate level of care. The Willink Committee (1957), explicitly basing its opinions on the MPC's standard, thought that a reasonable average list was 2,500 in urban areas and 2,000 in rural areas. The Gillie Committee (1963) appear on textual interpretation to have regarded an average list of 2,300 as being, if anything, on the low side. The BMA, in its Charter (1965) and in repeated evidence to the Review Body, has recommended a list of 2,000 for each practitioner as a not unreasonable

target. Paige and Jones (1966) explicitly endorsed the Charter's recommendation, and their recommended target average list size (1,775) appears on the low side only because it included assistants and trainees as well as principals. The Todd Commission (1968) regarded an average list of 2,180 as a reasonable guide to what it should be in the short-term future. The Harvard Davis Committee (1971) described an average list of 2,500 as optimum, provided the practitioner was located in an appropriate group setting. The views of doctors themselves, elicited in sample surveys, indicate an average list of about 2,000-2,100 as the ideal size, and as actual lists exceed the national average, a growing proportion of GPs express a serious concern about the numbers of patients for whom they have to care.

Against such consistency of opinion, views have recently been expressed by some general practitioners, based upon detailed analyses of their own practices, that lists of 3,000 or even 4,500 can be managed perfectly well, with no diminution in the quality of care and to the satisfaction of both doctors and patients, by doctors working within the context of a well-organised team. At present, however, such views do not seem to be widely shared among the profession.

In contrast to opinions about a reasonable list size, views about what the permitted maximum list size should be, where these have been explicitly stated, have generally declined over the period. The Cohen and Willink Reports in the 1950s saw no justification for reducing the existing maximum of 3,500 patients, but the BMA (1965) argued for a reduction to 3,000, and Paige and Jones (1966) for a further reduction to 2,500. The Todd Commission (1968) wanted a maximum of 2,500 in urban areas and only 2,000 in rural areas. In fact, the maximum list size permitted in the NHS has remained unchanged at 3,500 since 1952.

THE BASES OF OPINIONS ABOUT A REASONABLE LIST SIZE

In reviewing the post-war history of medical manpower planning in Britain, Maynard and Walker (1978) have documented the discrepancies between the forecasts made by various committees and commissions of future manpower needs and the actual supply of doctors in the forecast years. They find that, in the case of general practitioners, the actual supply has consistently fallen short of the forecasted requirements, whilst in the case of hospital doctors supply has exceeded forecasted requirements, sometimes by very large amounts. In commenting on the reasons for these striking discrepancies, Maynard and Walker talk about the mechanistic approach that has commonly been adopted in the past. 'Planners have been mesmerised ... by ratios - ratios that are partly the product of history; ratios that are largely the product of unsystematic thought about 'best' practice; ratios that are aggregated averages themselves; ratios too, that have been merely plucked from the air ... One "guestimate" has irregularly been substituted for another' (p.179).

This section seeks to examine this assertion in relation to the opinions expressed about a reasonable list size. On what have the opinions been based? Have they reflected any systematic analysis of what the concept of a 'reasonable list' might mean, or have they, as Maynard and Walker imply, been merely 'plucked from the air'? In examining such questions as these, the section begins to explore the assumptions, arguments and methodologies that have been employed in the past in operationalising the concept of a 'reasonable list'. It will be seen that, often, the bases for the opinions have been insubstantial and inadequate, but it is nevertheless important to examine them as carefully as possible in order to distinguish potentially useful ideas that might subsequently be incorporated into a sensible framework.

The section proceeds by examining in turn the basis of each of the opinions described in the previous section.

Medical Practices Committee

The early annual reports of the Medical Practices Committee offer some indication of how the Committee approached its statutory duty of ensuring 'an adequate number of medical practitioners in the areas of the Executive Councils'. Initially the Committee relied very heavily upon the judgements of the ECs themselves, and since the EC reports are never published, there are no means of knowing the basis on which these judgements were formed. However, the MPC was concerned from the outset not to appear to the medical

profession to be unreasonably restrictive, and it therefore emphasised in its first Annual Report (1949) that it required the Executive Councils to furnish 'the most cogent argument and the fullest information' that the number of practitioners was adequate, in order to maintain the greatest possible freedom of choice of doctor for the public, and to preserve for doctors the right to practise in any part of the country save in the exceptional circumstances described in S.34(3) of the 1946 Act.

Two years later, in its third Annual Report (1951), the MPC discussed the question of whether a local body really was better suited than a central one to judge the adequacy of the medical manpower in an area, and decided that it was not. The reason given by the Committee was that 'a local body can only judge the position from a local point of view and by comparing different parts of its own area; but it has no means of comparison with similar areas all over England and Wales' (page 1). This comment is instructive for the insight it offers into the Committee's apparent dependency upon comparative ratios in judging the adequacy of medical manpower. The implication is that an adequate judgement cannot be made by looking solely within an area; much depends upon comparison with the supply of manpower elsewhere. Yet no justification was offered in the Report that the supply elsewhere was either more or less adequate than in the original area. The Committee appears to have been using the comparative ratios between areas more as a means of securing an equal distribution of the existing stock of manpower than of determining the adequacy of the supply of manpower in each area.

A similar method of argument was evident in other topics discussed by the Committee in the early years of its life. In its sixth Annual Report in 1954, for example, the MPC talked about the signs of a saturation point within a few years in the total number of general practitioners in the NHS (implying that list sizes might actually fall below a reasonable average), yet the main evidence used to support this view was the growing numbers of GPs and the declining average list size in recent years. The implication seems to be that the ratio of doctors to population had been about right a few years earlier, and that any increase in that ratio would tend to be surplus to requirements. There was no justification of the view that the earlier supply of doctors had been adequate and sufficient. Or again, in its first Annual Report (1949), the MPC noted the tendency for doctors in industrial areas to have larger lists than those in residential or urban areas. On the basis of this evidence the Committee commented that 'it may be necessary at some later stage to consider whether or not some special action should be taken to encourage more practitioners to start in what may be regarded as the less attractive industrial

areas' (page 2). Here, the ratios are compared geographically rather than across time, as in the previous illustration. The Committee seems to have been saying that more doctors were needed in the industrial areas because the average list sizes there were higher than elsewhere, though once again there was no justification that the supply of doctors in these other places was adequate and sufficient.

It would be wrong to suggest that the Medical Practices Committee has arrived at its judgements about an adequate supply of manpower exclusively through the use of comparative ratios. It has been noted that, at the beginning of its work, the Committee had regard to the desirability of maximising the freedom of patients in their choice of a doctor and the freedom of doctors in their choice of a practice location. Other factors that the Committee has taken into account include the commitments of doctors in other parts of the NHS and outside it, the proportion of people receiving private primary medical care, the type of practice predominant in an area (urban, rural, coastal, industrial), local topography and conditions of travel, the dispensing responsibilities of doctors, the number of temporary residents signed on, and others. Most of these factors have in common a potential capacity for limiting the amount of care the doctor can give to his NHS patients, and as such they may rightly be used to modify a strict ratio approach to the matter of an adequate supply of manpower. The restriction in their utility arises from a very imperfect understanding of how far they limit the doctor's capacity to care, and whether the consequences of that limitation are regarded as sufficiently serious to warrant the extra resources necessary to compensate for them. In the case, for example, of two areas with comparable average list sizes, in one of which all doctors have sessions in the local hospital and in the other none, it would seem sensible to ask, before deciding whether any additional manpower was required in the former area, what differences occur between the areas in the amount and quality of care delivered to patients, and whether such differences as may exist are regarded as sufficiently important to justify the addition of more manpower. Expressed like this, it is apparent that a fair part of the MPC's difficulties is not of its own making, for as will be seen, there is very little information of this kind available.

Willink Committee

The Willink Committee (1957) also touched upon the output or productivity of general practitioners, but only briefly and cryptically. 'We gave considerable thought to the likelihood of any change in the average number of items of service given by a general practitioner each year to each of his patients.

The evidence we were able to secure on this point was, however, meagre and conflicting. In the absence of any discernible trend we decided to make no additional allowance one way or the other for this possibility' (para. 36). The paragraph is interesting because it shows the relevance that the Committee attached to the output of services in judging manpower requirements. Even though an absence of data apparently prevented the Committee from incorporating the notion of output into its final judgement, the paragraph indicates the importance of looking not just at numbers of doctors, but also at what they actually do. The paragraph is, however, clouded and ambiguous. The Committee may have had in mind the ways in which organisational and staffing changes in general practice could improve the efficiency of the GP's work, thereby enabling him either to provide more services to the same number of patients, or the same number of services to more patients. An analysis of this kind would explain the relevance of measuring output to the problem of defining manpower needs; but the paragraph is insufficiently clear to be sure that this is what the Committee had in mind.

Apart from this, two other considerations are discernible in the Report that appear to have influenced the Willink Committee's views about the future needs for general practitioners. First, the wider scatter of patients in rural than in urban areas indicated to the Committee the need for differential target average list sizes between such areas (para. 34). The argument is important in indicating how considerations of geography may affect the capacity of doctors to provide services, and therefore justify a different list size in rural and urban localities, but the argument has no direct relevance to the basic question of what the reasonable list size should normally be in either location. Second, the Committee noted the forecasts made by the Government Actuary of the growing proportion of elderly people in future years, and also the evidence discussed by the Cohen Committee and the General Register Office on 'the extent to which the need for medical attention varies with age' (para. 35). In the light of this, the Committee increased its estimate of the number of GPs needed in the future by 75 per year, but it gave no indication in the Report of how this number was reached, or whether the expanding needs of the elderly could be met more efficiently in other ways.

Gillie Committee

The Gillie Committee (1963) touched upon several matters that are pertinent to the judgement about an appropriate list size, although it is not clear from the Report exactly how the Committee used these insights. First, the Committee made explicit reference to the views of doctors themselves:

the conclusion that the maximum permitted list size should be reduced appears to have been based centrally upon the impossibility expressed by doctors of working to a satisfactory standard under existing conditions. Second, the Committee noted the variability between doctors in their capacity to provide services and hence (by implication) in the number of patients for whom they can provide adequate care. 'The number of patients a doctor can look after depends on many variables including his equipment, methods, rate of work and personality ... Doctors are so different in their individual characteristics of energy, health and enthusiasm that their lists are bound to differ' (paras. 119-120).

A third distinctive feature of the Gillie Committee's analysis was the introduction of the notion of standards of service. 'Co-operation between family doctors in group practices and local authority field workers enables a greater range and depth of work to be achieved, but reveals more sources of demand within the existing number of patients. The effect is therefore to improve the service rather than to lighten the load of work' (para. 122). This paragraph, though not particularly clear, is important in suggesting a relationship between the supply of resources, the demand for care, and the standard of service. An increase in the availability of resources (in this case, the joint resources of doctors and local authority field workers) leads to some increase in demand, but also enhances the scope and quality of service that can be offered. The Gillie Committee argued that the additional capacity resulting from a greater volume and efficiency of resources should be used to improve standards, not to lighten workloads or to increase list sizes; but in principle the extra capacity could be used for either of these purposes, with obvious consequences for the definition of a reasonable list.

A fourth distinctive element in the Gillie Report is the reference it made to a reasonable list size as a cost-benefit judgement. Again, the central point having been made, it is not elaborated; but the making of it is important. 'As professional opportunities for increasing the value of the family doctor's work become greater, with benefit to the patient and a saving (often concealed) of cost to the community, so the ideal number of persons in the care of each doctor, and the maximum that it is reasonable to fix, must be kept under review' (para. 122). Like other statements elsewhere in the Gillie Report, and in other reports, the meaning of the paragraph is not entirely clear. A possible interpretation is that, by expanding the scope and value of the work of family doctors, the additional benefits to patients might outweigh the extra costs of supplying more GPs. The Committee did not specify the areas in which the scope and value of general practice might be enhanced, but they are not

difficult to find. On this interpretation, an ideal list size is one where benefits and costs are equated at the margin; that is, where the benefit to be derived from an extra item of service is judged to be equal to the additional cost of providing that item. But even if the Committee did not have quite as precise an interpretation as this in mind, the reference to benefits and costs in the context of the need for flexibility in judging the ideal and the maximum list size is evidence of a considerable progress in thought since the Willink Report six years earlier.

In sum, the Gillie Report touched upon several important ideas in discussing future manpower needs in general practice, and in so doing it established a more rational approach to the question of a reasonable list size than the circularity inherent in the use of comparative ratios. However, like the Willink Committee before it, the Gillie Committee must have found itself lacking the information necessary to translate these ideas into firm recommendations, for having discussed them, the Committee did not draw upon them in any consistent way in reaching a judgement about the maximum and ideal list size.

Charter for the Family Doctor Service

The Charter for the Family Doctor Service (British Medical Association, 1965) contained no real justification for its view that a maximum list of 2,000 patients would be a reasonable target. On the basis of certain assumptions and data contained in the Charter it has been roughly calculated that a maximum list of this size might increase the time available for an average surgery consultation by about 30%, but the Charter did not follow the lead given by the Gillie Committee in discussing whether the benefits resulting from this additional time would be regarded as justifying the costs of the extra doctors needed. A fuller statement of the BMA's case can be constructed from accounts in the reports of the Review Body on Doctors' and Dentists' Remuneration of the evidence submitted to it by the Association. These are examined later in this section.

National Institute of Economic and Social Research

As noted above, Paige and Jones (1966) relied explicitly upon the views expressed in the Charter in support of their observation that a list of 2,500 patients would be regarded by most people as a reasonable maximum. In discussing a desirable list size for the future, Paige and Jones concentrated much more on the demand than on the supply side of the equation. They pointed, for example, to the growing number of old people in the community, to the growth of community care for the elderly, the handicapped and the mentally ill,

and to the inadequacy of existing standards. 'There is no doubt that we shall want a big expansion in the domiciliary services and in day care, but much of it will be needed to help those who are at present getting too little help or none. There are many elderly people managing in their own homes without domiciliary help who ought not to have to do so, and many mentally disordered persons living in private households only at an excessive cost to other members of their families' (page 32). Paige and Jones said that they allowed for this pattern of need in reaching their estimate of future manpower requirements, but they did not indicate how. There was no attempt to quantify these additional demands, or to calculate the number of additional GPs required to produce the services needed to satisfy them.

Todd Commission

It was noted above that, in making its short-term (1965-75) estimates of manpower needs, the Todd Commission (1968) regarded the population/doctor ratio that had existed in 1961 (2,150:1) as being a 'reasonable guide' to what it should be in the near future. However, as with previous attempts to forecast manpower requirements on the basis of comparative ratios, the Commission's Report contained little systematic justification for its belief in the adequacy of the base (1961) ratio. It talked, for example, about the need to bring the numbers of general practitioners 'up to the desired standards' (para. 318), but the only apparent ground for assuming that existing standards were less than desirable was 'the rise in recent years in the average number of patients for each principal' (para. 327). In fact the Commission did offer some justification, but it took the form less of explaining why the lower ratio was better than of pointing out the support that existed for this viewpoint. 'It has been widely accepted that lists of patients in some practices and parts of the country are too high. It is difficult, however, to establish a firm level of "need". Conditions in practices vary widely for geographical, environmental and social reasons as well as by the extent of ancillary help' (Appendix 12, Annex para. 8). The problem still remains, although a widespread consensus about the excessive size of practice lists would indicate at least a subjective experience that must be accorded some weight. Again, however, the Todd Commission, like the Willink Committee before it, encountered the problems of a paucity of operational research which could serve as a guide to the establishment of need, or to an assessment of the effect of practice conditions on the services provided by doctors.

In making its long-term estimates, the Todd Commission, like Paige and Jones, attempted to identify the factors that might cause future changes in the

demand for doctors, and hence, by implication, changes in the desirable ratio of doctors to population. The Commission noted, for example, that the chronic diseases of old age would tend to raise the demand for medical care, and that psychiatric services would also be extended. Against this, however, more attention would be paid to the promotion of health through such measures as screening for pre-disease conditions and regular medical examinations. The Commission also thought (rather surprisingly) that the changing age structure of the population over the next thirty years would 'lead to a lessening in the growth of medical care requirements' (para. 343). After reviewing these factors, the Commission expected the demand for medical services to be 'at least as great in the next thirty years as it has been in the last' (para. 335), and that accordingly 'provision should be made for the doctor-population quotient in Britain to continue to rise at a rate no lower than in the past' (para. 338). On the supply side, the Commission thought that the organisation and efficiency of the doctors' work would probably improve, especially with the growth of group practice based on health centres, but nevertheless 'in our firm opinion the needs of the future will not be met by an annual growth rate of less than 1.5% in the doctor-population quotient' (para. 338).

As in some of the other Reports discussed here, there is a gap in the Todd Commission's argument between the identification of factors likely to increase the future demand for doctors and the specification of the additional number of doctors needed to meet it. Having listed the various factors, they appear to have been used as no more than background factors in shaping the decision about the requisite future increase in manpower supply. The Commission explicitly acknowledged this. 'These concepts (of need and demand) cannot be defined quantitatively. In using them we do not imply that there is some absolute or optimum level of health services which can be measured and towards which we should aim. We do not believe that the health services in Britain are close to a stage where, on the assumption that all economic and social barriers were removed, they could meet all the demands likely to be made on them' (para. 333). Like Willink and Gillie before it, the Todd Commission lacked the detailed information necessary to make explicit links between future demand and supply, and it also seems to have ignored the role of standards of care. As a result, the Commission appears to have relied very heavily upon the extrapolation of past trends in doctor-population ratios in judging future manpower requirements, and to have used the material on future changes in demand as corroborative evidence that the judgement was broadly of the right order.

Harvard Davis Committee

The thinking set out in the Harvard Davis Report (1971) seems to owe more to Gillie than to Todd. The Committee took the view that a general practitioner cannot properly provide care for more than about 2,500 patients, although the arguments rehearsed in the Report actually appear to support the case for a lower list size than this. The Committee explicitly rejected the view that group practice increases the number of patients for whom GPs can provide care. Their reasons for this (echoing, but not quoting from, the Gillie Report) were first that the addition of community and practice nurses tend to generate new work rather than to relieve the doctor of some of his existing workload, and second that the majority of GPs consider that the time they are able to give to each patient at present is inadequate. (Although the Committee provided referenced evidence for many of their assertions, they offered no empirical support for this one.)

Having made the negative point that a well-organised group doesn't increase the number of patients for whom care can be provided, the Harvard Davis Committee went on to make the positive point that 'the advantages .... lie much more in the ability of the group to provide a higher quality of medical care in the community' (para. 46). Again, this appreciation of the way in which standards control the tension between supply and demand is evocative of Gillie: an increase in the supply of resources should be used to enhance the quality of care that is given, not to increase list sizes or to lighten workloads. The Committee assumed that the development of the team would enable the doctor to delegate some of his work to other members, thus allowing him to spend more time with patients requiring medical care and hence improving the quality of care he could give to them.

Like the Gillie Committee, Harvard Davis then took the argument one stage further by offering an economic justification for the pursuit of higher standards of care, namely the possibility of savings elsewhere, particularly in the hospital service. 'If we can achieve this object (of a higher quality of medical care in the community), it would justify economically the provision of a more generous staff/population ratio than exists at the moment' (para. 46). Again, the basic point having been made, it is not elaborated, and it is a matter of speculation as to thinking which underlay it. It is nonetheless. an important point. If cost was no object, there would presumably be total agreement that standards should be raised to the highest level that is technically capable of achievement, with obvious and gross consequences for the supply of manpower; but in reality there must come a point where further increases in

standards are not worth the cost of achieving them. The argument of the Harvard Davis Committee implies that this point exists when the marginal costs of similar benefits are equal in primary and hospital care. In other words, it is justifiable to add more resources to primary care (through the development of group practice teams) until the point is reached where further resources fail to secure corresponding savings in the hospital sector. However attractive this argument might be in theory, in practice insufficient data exist about the substitutability of group practice for hospital care, and about the marginal costs in each context, to base a judgement about ideal list size upon it. Like earlier committees, therefore, the Harvard Davis Committee, having identified an interesting approach to the problem, was obliged through lack of information to resort to an informed guess in specifying the target list size in a group practice context.

Review Body on Doctors' and Dentists' Remuneration

The evidence submitted to the Review Body shows that the British Medical Association has consistently regarded list sizes as excessive, and has argued that no individual doctor should be responsible for more than 2,000 patients. The general ground on which this argument has been advanced is that the workload generated by existing lists is unreasonably high, and increasing. Thus the nub of the Association's case (as documented in the Review Body Reports) has been the enumeration of factors which have tended to increase the demand for the services of family doctors. These have included: in the 1966 Report, the increased health-consciousness of people, the growth of certification for eligibility for benefits, the increase in stress diseases, and the advent of more elaborate diagnostic tests and more powerful drugs (para. 69); in the 1968 Report, the rapid technical advances in medicine (para. 53); in the 1970 Report, the trend towards earlier discharges from hospital (para. 29), the rising number of claims for sickness benefit (para. 30), and the change in moral attitudes towards sex and drugs (para. 31); in the 1975 Report, the growing use of better (but more time-consuming) investigative techniques, and the increased use of ancillary staff (para. 37); and in the 1978 Report, the emphasis on screening and preventive medicine, the growth of health care teams, and the increasing burden of management and administrative tasks (para. 43).

The basic argument pursued by the BMA is not inconsistent with those in the Gillie and Harvard Davis Reports: that, for all the reasons listed, the increasing workload generated by the average doctor's list cannot be handled without resort to unacceptable compromise. Either the doctor works excessively long hours, or he limits the number of patients he sees, or he compromises on the quality of care he gives by reducing the amount of time spent with each

patient. Unlike Gillie and Harvard Davis, however, the BMA has not apparently tried to justify the cost of the additional doctors, for example by pointing to possible net savings in other parts of the NHS or by claiming that the social (and possibly economic)value of the better standards of care would justify the cost. It is, moreover, implicit in the BMA's argument that, by reducing list sizes to 2,000, the workload could be accommodated by the average GP without resort to any unacceptable compromise. However, the Association has largely failed to document this assumption, for example by demonstrating that significant variations in output do occur between doctors with different list sizes, or that as list sizes increase above 2,000 patients the quality of care worsens appreciably.

It is on precisely these points that the Health Departments have tended to resist the BMA's case, claiming either that list size is a poor indicator of workload, or that, up to a point, higher list sizes do not inevitably involve an increase in work or a reduction in standards. It has been noted, for example, that the Department's evidence to the Review Body in 1966 contained the argument that improvements in the organisation of general practice might enable GPs to care for more patients with no loss of efficiency. In subsequent years the Departments tended to argue that the rise in average list sizes during the latter part of the 1960s had not been sufficiently great to cause a significant increase in the workload falling on the average general practitioner, and could not therefore be used to justify a claim for extra payment based upon additional work.

The Review Body itself stated the BMA's basic argument very clearly in its 1968 Report. 'Excessive workload is to some extent the counterpart of manpower shortage. We do not doubt that ... the burden on some GPs continues to increase. It would clearly be wrong to encourage GPs to have exceptionally large lists beyond their capacity to provide adequate treatment ... In general, we think that when the patients cared for by a doctor become exceptionally numerous the burden be rewarded by the corresponding increase in capitation fees' (para. 57). Without actually selecting an optimum list size, the Review Body makes the general point of principal that such a list would be one that did not place an 'excessive workload' on the doctor beyond his capacity to provide 'adequate treatment'. Moreover, the Review Body seems in this quotation to have accepted the existence of a positive correlation between workload and list size, although it has been noted that in more recent reports the Review Body has shown greater scepticism towards the BMA's claim of ever-increasing workloads. However, the Review Body's general statement of principal lacks substance because of its failure not only to quantify the key

terms of 'excessive workload' and 'adequate treatment', but even to suggest how their definition might be approached. Unless the ways are identified in which workload becomes excessive and treatment inadequate, the guidelines are lacking for any informed judgement about what constitutes a reasonable list size.

Opinions of doctors

Cartwright's (1967) study offered no direct evidence of the reasons for the respondents' choices of ideal list sizes although some information was given on the correlates of the choices. Doctors were more likely to select smaller (< 2,000) lists as ideal if they currently had a nurse employed in the practice, a finding that is explained by Cartwright in terms of their appreciation of the role of the nurse in extending the range of care available to patients rather than increasing the number of patients for whom care can be given. Such an argument was also advanced in the Gillie and Harvard Davis Reports, and was supported by an earlier study by Cartwright and Scott (1961). Doctors with actual lists below 2,000 were less likely than the rest to report that they enjoyed general practice 'very much', but the principal factor in enjoyment appeared to be the degree of congruity between actual and ideal list sizes. 'Sixty per cent of those whose present list size was in the same group as their ideal enjoyed general practice very much, compared with 55 per cent of those who looked after 500 more patients than they thought was ideal, 48 per cent of those who looked after 1,000 more than their ideal, and 32 per cent of those who looked after less patients than they thought ideal' (page 162).

Mechanic's (1974) survey, as noted, carried no direct question about ideal list size, but it did show a positive correlation between the doctors' actual list sizes and their perceptions of difficulty in coping with them. Doctors with lists above 2,500 were markedly more likely to report 'very serious problems' than those with smaller lists. Mechanic's comment upon the nature of these problems is worth repeating at some length. 'Doctors respond to large practices not by continually increasing their work-day, but by practising at a different pace and style which is particularly frustrating and uncongenial. They feel deprived not only in terms of the hours they devote to their patients, but more importantly in terms of the amount of work and effort they must pack into this period of time. Such a pattern of work requires them to practise on an assembly line basis which diminishes the unique satisfactions possible in a general practice .... All of the aspects of hurried practice - spot diagnoses, inability to provide enough time for patients, failure to do an adequate examination or undertake needed action - were related to size of practice. The size of practice variable not only encompasses numbers of patients, but also

the manner and pace of the doctor's work, and it exerts an influence on his entire outlook' (page 100).

The data underlying this comment show a remarkably consistent increase in reported dissatisfaction with various aspects of work as list sizes rise. The larger the number of patients for whom the doctors were responsible, the more likely they were to report a 'very serious problem' over the need for rapid diagnosis, the time availa ble for each patient, the time and effort required in the practice, the effect of time pressure on clinical practice, and the time spent on the social aspects of medical practice (Table 1, pages 98-99). These results clearly support the general argument that the B.M.A. has consistently put to the Review Body that large lists generate a volume of work requiring GPs to practise standards of medical care that they recognise to be unsatisfactory; and it is interesting that Mechanic also found that GPs with large lists were more likely than those with small lists to have submitted their undated resignations to the British Medical Guild in 1965. It is, however, difficult to use Mechanic's data as the basis for a rational judgement about a reasonable list size. First, the data merely show that, as list sizes increased, proportionately more doctors expressed concern about their standards of practice: they give no indication of the point at which further gains in standards of practice might be considered to be outweighed by the cost of producing the extra doctors needed to achieve those gains. Even among the doctors in Mechanic's survey with list sizes of less than 1,500 (that is, substantially below the B.M.A.'s recommended individual maximum list size), about 10 per cent reported a 'very serious problem' with most of the facets of their work. Second, there are no yardsticks by which to evaluate the significance of the doctors' perceptions of a 'very serious problem'. There is likely to be considerable variability among a random group of GPs in their perceptions of what constitutes a serious problem, and there is no indication of the effect or outcome for the patient when the doctor feels that his work is suffering through pressure of time.

Fry's (1972) claim that a GP can provide scund care for some 4,500 patients in ways that are 'satisfactory to both patients and doctors' was based on a review of 21 years' work in his own practice. The pattern of work in the practice had changed substantially during this period of time. The average number of surgery consultations and home visits per patient per year declined from 3.3 and 0.6 respectively in 1951 to 2.1 and 0.1 in 1971. The average number of surgery consultations per doctor per day declined from 40 in 1951 to 30 in 1971, and of home visits from 9 to 2. A decline in consultation rates occurred among patients with rheumatic, gastro-intestinal, cardiovascular,

central nervous system, respiratory and dermatological conditions. There were increased rates of attendance for obstetric and gynaecological care, and for immunisation. No marked changes were noted for upper respiratory infections, psychiatric disorders, ear, nose and throat and urological conditions. Hospital referrals per 100 patients declined from 10.5 in 1951 to 4.0 in 1971; radiography referrals declined from 6.9 to 5.5 and pathology referrals increased from 5.7 to 6.2. Fry attributed the decline in consultation rates to changes introduced in 1963, particularly the introduction of a full appointment system, the attachment of a health visitor, and the employment of additional secretary-receptionists; to 'active and positive efforts to reduce unnecessary work'; and to better methods of care. No information was given about demographic or social changes in the practice population that might have influenced the trends.

The consultation rates given by Fry indicate that his practice has coped with an uncommonly large list by achieving a very low rate of consultation. In comparison with Fry's surgery consultation rates, the two National Morbidity Surveys reported average rates of 3.75 in 1955-6 and 3.01 in 1970-1, with the regional variation in the latter study ranging from 2.5 in the West Midlands to 3.6 in the North West (General Register Office, 1958; Office of Population Censuses and Surveys, 1974; Crombie et al, 1975). In 1971, Fry's surgery consultation rate of 2.0 per patient per year was the third lowest of 14 practitioners listed in the RCGP's 'Present state and future needs of general practice', and his long-term average consultation rate of 2.8 between 1949 and 1972 was the lowest (in many cases by a large margin) of 15 practitioners for whom records for at least ten years were available (Royal College of General Practitioners, 1973). Studies published since 1970 (summarised on page 76 of this report) show that Fry's consultation rate remains uncommonly low.

Fry does not indicate whether his low rate of consultation per patient per year means that he sees fewer of his patients each year than the average GP, or whether he sees approximately the same number but on fewer occasions each. In either case, he is presumably coping with a larger-than-average list size by seeing fewer patients, rather than by reducing the average time spent with each patient seen. His claim, then that he was providing a 'much better service' for his patients in 1971 than in 1951, and that patients were satisfied with the standard of care provided, needs further substantiation. There is no information about the characteristics or needs of those patients who were seen less frequently by Fry than they might have been in a practice with a higher consultation rate, and no indication is given of the outcome (in terms of patient satisfaction as well as clinical outcome) of this style of practice. It is possible, moreover that the dissatisfied patients have, over the years,

transferred to other practices. In brief, Fry appears to have demonstrated the possibility of handling a list of 4,500, but not the desirability or reasonableness of a list of this size. An editorial in the Journal of the Royal College of General Practitioners (1972), based upon Fry's paper, posed some pertinent questions. 'Are patients receiving adequate care in the home? What pressures are placed upon those who request home visits? What kinds of conditions are brought to the surgery? Do any patients suffer? Do patients like it? What kind of consultation occurs? What work is done, how much is delegated, and what is referred? ..... This paper raises more questions than it answers' (pages 492-3).

The claim by Marsh and Kaim-Caudle (1976) that a list of 2,400 may well become too small to occupy the time of an established general practitioner was based upon the detailed analysis of one year's work by Marsh as one of five GPs in a practice team in Teesside (see also Marsh and McNay, 1974a, 1974b). Marsh and his practice team provided total care throughout the year for 3,137 patients. The data given by Marsh and McNay (1974a) indicate that the average number of surgery consultations with the doctor was 1.9 per patient during the study year, and the average number of home visits was 0.4. These consultation rates are very similar to those reported by Fry for 1972, suggesting that Marsh likewise coped with a large list by seeing fewer patients than the average G.P.

Marsh, however, is more explicit than Fry about the distribution of work between the doctor and other members of the team. Whereas Fry's paper fails to clarify whether the reported consultation rates refer to the doctor only or to the whole team, Marsh makes the distinction clear. The figures quoted above are for the doctor only; in addition to these, there was an average throughout the year of 0.6 contacts per patient with the practice nurse and 0.2 with the health visitor. In total, therefore, the average number of contacts with the team by each patient during the study year was 3.1, a figure very close to the average consultation rate for doctors found in the 1970-1 National Morbidity Survey (Office of Population Censuses and Surveys, 1974). One possible interpretation of these figures is that the nurse and health visitor are used as substitute resources for the doctor, especially in follow-up work after the initial consultation, thereby enabling the team to care for more patients in total, rather than to provide a better or more extensive pattern of care for a lesser number of patients. The use of the team in this way is in marked contrast to the function of the team envisaged by the Harvard Davis Committee.

Like Fry, Marsh has apparently demonstrated the feasibility of a GP, supported by a nurse, health visitor and secretary-receptionist, providing

care for 3,000 or more people, but he has not unambiguously established his claim that his style of work does not entail any reduction in the standard of care. 'We would like to be able to dismiss the thought that the general care provided for the patients might have been poor, but as yet no acceptable and defined standards of quality of clinical care in general practice have been worked out' (Marsh and McNay, 1974a, page 317). Some evidence is available, most of it supporting the author's belief that adequate standards were followed. A GP with a list of 3,137 patients, each of whom consults, on average, 2.3 times a year, has fewer consultations than one with 2,500 patients each consulting an average of 3.0 times a year. For the same total number of hours worked, the former doctor can actually spend rather more time on each consultation than the latter, although, as will be shown later, this does not necessarily mean that he is providing a better standard of care. Marsh also found that the proportions of patients admitted to inpatient departments and referred to outpatient departments during the course of the study year were, respectively, two-thirds and one-third of the national average. Of the investigations carried out by hospital diagnostic departments, 42 per cent were found to be abnormal. These results indicate that the doctor retained as much responsibility for his patients as most GPs, and was not relying excessively and trivially upon the hospital to enable him to cope. Against this, only about 40 per cent of consultations were initiated by the doctor himself, which is rather lower than those noted in other studies (Richardson, et al, 1973; Williams, 1970), and may be indicative of a lower standard of continuing care. Much depends upon the kind of follow-up care given by the nurse and health visitor.

Summary

The aim in this section has been to draw out as clearly as possible the considerations that seem to have shaped the expressions of opinion about a reasonable list size for GPs that were summarised in the preceding section. The results suggest that Maynard and Walker's (1978) observations about 'guestimates' being plucked from the air are rather harsh, for many and varied arguments have been deployed in support of the chosen ideal. However, it is clear that there has been little consistency in the use even of similar concepts, and there has also been a marked lack of empirical evidence of a kind that would enable these concepts to be converted into actual figures. The lack of relevant data may be one reason why so many of the opinions about a reasonable list size have tended to concentrate within a fairly narrow range, for in the absence of persuasive evidence to the contrary, it is sensible to keep within the boundaries of conventional wisdom. It may be relevant in this context

that two of the most extreme opinions reviewed in this report, those of Fry (1972) and Marsh and Kaim-Caudle (1976), are supported by more detailed evidence than any of the others.

One important consideration, running through many of the arguments, is that of standards of care. Although the concept of 'standards' is very undeveloped in many of these sources, the view is clearly expressed that, as list sizes increase, standards decline to the point where they are frankly inadequate. A major category of evidence in support of this view is the feeling expressed by doctors with large lists that they are often obliged to work at standards lower than those to which they have been trained and at which they would ideally wish to practise.
This is, plainly, an important argument. There are several possible general grounds upon which the judgement about a reasonable list size might be based, and to select that of standards is to declare an important belief about the purpose of the exercise. However, it is clear from the evidence reviewed in this section that considerable difficulty has been experienced in translating the general argument about standards into specific judgements about the point at which lists become so large as to pose an unacceptable threat to standards of care. Little work has been done on the definition and measurement of standards in general practice, and very little appears to be known about the ways in which standards vary with list size. Without a firm empirical basis of this kind it is difficult to progress beyond the general argument that, because doctors with large lists feel themselves to be practising at sub-optimal standards, those lists are unreasonably large.

A second consideration, evident in some of the arguments reviewed in this section, is that of the needs and demands of the population. It is, so to speak, the other side of the coin about standards, for it reflects the input or workload coming into a practice rather than the output of services. The general argument here is that list sizes are unreasonably large if they result in significant unmet needs or demands among the population, even though there may be a perfectly adequate standard of care to those patients whose needs are met. Again, this is plainly an important argument, particularly at a time when the pattern of needs and demands in the community seems to be changing quite quickly; but, as with the argument about standards, there has been little empirical evidence available by which to judge the point where list sizes become unreasonably large on this criterion. The most careful evidence (imperfect though it is) is probably that assembled by Fry (1972) and Marsh and Kaim-Caudle (1976), indicating that even list sizes above 3,000 can be handled in ways that do not lead to a major back-log of unmet needs. However,

a more systematic investigation of the relationship between needs, demands and list size would seem to be needed before an empirically-based judgement could be formed.

A third theme running through some of the arguments (though not always expressed in precisely this language) is that of efficiency. It is recognised in some of the sources that there is no unique way of producing the outputs of primary care, for different mixtures of resources may be used in different ways to provide the range of services that constitute primary care. Even when confronted with similar patterns of need in their practice populations, and when providing a similar quality of care, doctors may still vary in the number of patients with whom they can cope by virtue of their variations in efficiency. The organisation of the practice and the employment of other members of the primary care team are two obvious variables affecting the efficiency with which doctors produce their services and hence the number of patients to whom they can offer a specified standard of care. It is not too clear, however, how these insights about the relationship between resources and services have influenced the judgement about a reasonable list size. One possible implication (which is hinted at but not really developed in the literature) is that a certain level of efficiency is a necessary precondition for judging a list size to be unreasonably large. If, for example, a doctor could increase his list, with no diminution in his standards of work, by improving the efficiency of his practice organisation, it would be difficult to argue that his list is unreasonably large. However, as the evidence reviewed in this section indicates, the notion of efficiency in general practice requires careful handling. Forms of practice organisation which enable doctors to accept responsibility for a greater number of patients may incur costs that outweigh the savings from the consequent reduction in medical manpower. A full primary care team, for example, may actually be a more costly type of organisation even though it enhances the efficiency of the <u>doctor</u> by enabling him to accept more patients onto his list.

The conscious acknowledgement of costs represents a fourth consideration in the judgement about a reasonable list size, although it is by no means evident in all the sources reviewed in this section. The central argument is that reductions in average list sizes are costly, and that in the real world these costs have to be justified by the additional benefits they produce. There are two distinct points to be made here. First, there is the empirical question of the ways in which reductions in list size are related to improvements in services - for example by enhancing the quality of care or reducing the amount of unmet need. Although, as noted above, some generous assumptions have been made about this relationship, much more careful investigation is needed of the

precise ways in which benefits accrue from reductions in list sizes before a firm judgement can be made about the range of list size that is compatible with an acceptable volume and quality of service output in differing circumstances. The second point is that even if reductions in list sizes do produce concomitant benefits, not all the additional benefits that may technically be capable of achievement will actually be regarded as justifying the extra costs involved. If the costs of reductions in list sizes could be ignored, it would be impossible to resist the argument that no GP's list should exceed 1,000 or possibly even 500 patients - provided, of course, that reductions to these levels produced some identifiable gains in benefits. However, precisely because reductions in average list sizes do involve additional costs, the judgement about a reasonable list size must reflect an assessment both of the benefits and the costs of achieving it. By concentrating on the anticipated benefits of lower lists and ignoring the costs, many of the sources reviewed in this report are seemingly out of touch with the real world. They make the general point that, if list sizes were to be reduced, the quality of care might increase and the amount of unmet need might diminish, but they do not indicate by how much it is worth reducing list sizes in order to secure these benefits. Only the Gillie and Harvard Davis Committees explicitly referred to a cost-benefit type of argument, and in each case the argument was cast primarily in economic terms by emphasising the possible savings elsewhere that might flow from the input of more resources into primary care. In principle, however, the argument might also be conducted in social terms by weighing the social benefits of better care against the costs of achieving them. This would be a perfectly proper argument, although the failure to express the costs and benefits in a standard unit (such as money) would necessarily introduce value judgements into the equation. There would be nothing unusual about this, however, for a large part of health service planning involves judgements of this kind, and it is presumably one of the tasks of planners to supply (and if necessary defend) such judgements.

MODELS OF MEDICAL MANPOWER REQUIREMENTS

The previous two sections of this report have focused upon the opinions offered in recent years about a reasonable list size for general practitioners in Britain. An attempt was made to identify not only the target list sizes themselves but also the considerations that appear to have shaped the choice of those targets. As the summary to the preceding section showed, a number of important concepts and arguments are discernible in this area of the literature, but they have often been unsupported by empirical evidence, and they have not been welded together into a coherent set of ideas.

This section has a similar aim to the preceding one, but it seeks to broaden, and to some extent organise, these concepts and ideas by moving beyond the specific problem of a desirable ratio of population to GPs in Britain towards the more general problem of the adequacy of the supply of doctors in modern societies. All countries have to decide whether they have enough, too many or too few doctors, and the literature on medical manpower identifies a number of more or less coherent models of how such decisions are - or might be - reached. This section aims, in a necessarily selective way, to present some of these models and to discuss some of the strengths and weaknesses of them. There is inevitably an overlap of ideas with the previous section, for many of the arguments summarised there have been discussed and implemented elsewhere; but by making the models rather than the sources the focus of attention, it is hoped that a more organised picture will emerge of the possible foundations upon which policy judgements might be based about the adequacy of the supply of medical manpower in general and GPs in particular.

Surplus and shortage in the market

Where the services of doctors are bought and sold in a free market, the concepts of a shortage or surplus of doctors have precise meanings. There is a shortage of doctors when the demand for their services exceeds supply at current prices, and a surplus when the supply of their services exceeds demand at current prices. Imbalances of this kind are theoretically redressed through changes in the price of services. In fact it is doubtful whether any health care system allows the totally unfettered operation of free market forces. In various ways, both the suppliers and consumers of medical care can be protected from the full economic consequences of their behaviour, and probably all governments have taken the view, in varying degrees, that the social implications of health care are such that access to it cannot be allowed to depend exclusively upon the ability to pay the full market price. In health care systems (such as the NHS) where direct pricing has been abolished, the potential demand for care is very high indeed, and the equilibrium between supply and demand ceases

to be a precise indicator of the adequacy of supply. As long as health care carries no direct price to the patient, the demand for many health care services will exceed the resources that are made available to provide them, particularly primary care services where the nature of the output is heterogeneous and the doctor can control a large element of the demand for his services.

Comparative ratios

Assessments of the adequacy of the supply of medical manpower have often been based upon comparisons with other places or other times, the assumption being that the compared ratio of doctors to population is correct or desirable, and that surpluses or shortages exist whenever that ratio is exceeded or not attained. Examples have been given of the use made of comparative ratios between areas and across time by the Medical Practices Committee and the Todd Commission (see pages 19 and 24), and the method has also been used in other countries. A study sponsored by the Canadian Royal Commission on Health Services (Judek, 1964) was based upon a procedure very similar to that adopted by the Todd Commission. Judek assumed that the doctor/population ratio in Canada in 1961 was 'right', and that the average number of visits to the doctor by each person per year was also 'correct'. By forecasting future population growth, using projections of net migration and natural increase, Judek calculated the number of doctors that would be needed in future years in order to maintain the 1961 doctor/population ratios and consultation rates. He then calculated the future supply of doctors in the absence of any change in policy, and by subtracting this from the number of doctors 'needed' he was able to derive a measure of the future 'shortage' of doctors. These were quite large. By 1991, for example, Judek estimated that the need for doctors in Canada would exceed the supply by 30 per cent, and, apparently in large part on the basis of these findings, federal legislation was introduced in 1966 to accelerate the output of trained doctors (Migue and Belanger, 1974).

A similar story is reported from the United States by Reinhardt (1975). In 1966 the U.S. Public Health Service predicted that between 400,000 and 425,000 doctors would be needed by 1975. The higher of these two estimates was based upon the belief that, by 1975, the doctor/population ratio for the country as a whole should be the same as the best ratio actually attained by the four major regions in 1966. The lower estimate reflected the number of doctors needed if the nation as a whole was to achieve the staffing patterns characteristic of the comprehensive pre-paid group practices. Using the latter norm, the Public Health Service identified a national shortage of 50,000 doctors in 1969, a figure that subsequently found its way into the

1970 Manpower Report of the President and thence into the 1970 Report of the Carnegie Commission on Higher Education, where it was used to support the recommendation that the number of medical school entrants should be increased from the 1970 estimate of 10,800 to 15,300 by 1976 and about 16,400 by 1978.

As these examples show, the use of comparative ratios tends to displace rather than clarify the selection of a reasonable list size or doctor/population ratio, for the need remains to justify the compared ratio as a desirable target. Why was the average list size in Great Britain in 1961 assumed by the Todd Commission to be the right future target in the short-term? Why was the 1961 doctor/population ratio in Canada assumed by Judek to contain, as Migue and Belanger (1974) put it, 'absolutely no shortage or surplus'? Why was the best regional ratio in the United States in 1966 assumed by the Public Health Service to be the appropriate norm which the nation should strive to emulate by 1975? Clearly, such questions merely relocate, rather than eliminate, the need to justify the chosen ideal.

There are other difficulties in the use of comparative ratios (Lave, et al, 1975). One is that, by regarding the best achieved local ratio as a general future target, there will always be further improvements to be made, and an uncritical position may be adopted in which more doctors are invariably regarded as a 'good thing'. Reinhardt (1975) provides an interesting example of this type of inflationary thinking in his account of the recent history of manpower forecasting in the United States. The 1971 Comprehensive Manpower Act ensured that, on conservative estimates, the doctor/population ratio for the nation as a whole would, by 1980, have surpassed the best regional ratio of 1966, thus achieving the level of manpower regarded as necessary by the Public Health Services in 1966. By 1971, however, the basis of the PHS projections of manpower needs had shifted away from the best regional ratio in 1966 to the best state ratio in 1971, thereby ensuring that supply would continue to fall short of 'needs'. As Reinhardt puts it, 'If the time path of the actual physician supply in past years is compared with requirements projected earlier for those years, it will be noted that actual supply typically has come close to or even exceeded projected requirements. Oddly enough, this fortunate turn of events has never been a source of satisfaction, for in the meantime the definition of requirements has been changed and new manpower forecasts have been issued, each pointing to an existing or impending physician shortage. The problem of the "doctor shortage" appears to be one incapable of solution' (page 52). The (mis)use of comparative ratios is not the only factor in causing this state of affairs, but it does appear to be a contributory one. Stevens (1971) suggests one reason why this has been allowed to happen. Commenting on the assumption

in the Bayne-Jones (1958) and Bane (1955) reports that the existing ratio of doctors to population was a minimum social requirement to be maintained in the future, Stevens observes that 'while it could just as well be argued that there had been too many physicians in 1950, or that physician productivity had increased sufficiently to make an increase in the number of physicians unimportant, the findings were politically persuasive' (page 365). Stevens goes on to note that the (then) current concerns over the supposed scientific supremacy of the Russians, the mounting interest in the social provision of health services, and the much publicised deficiencies in health services for the elderly, all made it seem reasonable to suppose that the number of doctors should be increased. 'There were few who were likely to cavil over the need for more rather than less physicians.'

A more technical difficulty in the use of comparative ratios as the basis for future manpower forecasts lies in the assumptions they embody about the nature of the demand for, and supply of, health care services. On the demand side, the total demand for health care services is a function not only of the size of the population, but also its characteristics. In Great Britain, for example, women consult their GPs more frequently than men, old people more frequently than young people, and the widowed more frequently than those who are single or married (OPCS, 1978). If therefore, the level of demand that is satisfied by a specified doctor/population ratio is regarded as a reasonable indicator of what it should be in the future, account must be taken not only of projected changes in the size of the population, but also in its demand-related characteristics. Whilst this may be possible for basic characteristics such as age and sex, it is clearly more difficult for other, equally important variables.

On the supply side, as Fein (1967) has stressed, it is important to 'distinguish physician manpower from physicians' services' (page 4). Doctors have no inherent value; what is valued is the services they produce, and in accepting any particular doctor/population ratio as reasonable or ideal, there is the implication that the services produced by those doctors are also in some sense reasonable or ideal. By calculating the number of doctors required at some future date to achieve the desired ratio, the assumption is necessarily made that the productivity of doctors is constant - that is, that the same number of doctors will be needed in the future that are needed in the present to render a given number of services. A similar assumption is made in comparing ratios between areas as well as over time. In fact a large body of literature attests to the unreasonableness of this assumption, and indicates the dangers of making it (see, for example, Rafferty, 1974). Doctors do vary in their capacity to produce services, for reasons of age, motivation, style of

work, hours worked, geographical location, the availability of supporting services and personnel, and so on; and a failure to take account of these variations may lead to policy proposals that foster inefficiencies.

Reinhardt (1975) offers an interesting example of this. Although the example is drawn from the United States, the implications may be equally applicable in Britain. Reinhardt presents data for three census divisions in America with different endowments of medical manpower: New England (with 161 active doctors per hundred-thousand population in 1970), East-North Central (115), and East-South Central (95). Without enquiring any further into the productivity of the doctors in each division, the policy might be advocated of raising the manpower levels in the less well-endowed divisions to that enjoyed in New England. Such policy advocacy would be consistent with the approaches adopted in the past by the Bayne-Jones and Bane Committees, the Public Health Service and the Carnegie Commission, amongst others. Yet Reinhardt provides further information about the pattern of work in these three divisions which indicates the presence of countervailing considerations. The average number of doctor-visits per person per year was very similar in each division, and in fact was slightly higher in the worst endowed division (East-South Central) than in the other two. This suggests that the productivity of the doctors was inversely related to the doctor/population ratio of the division in which they worked: the higher the ratio, the lower the productivity of individual doctors. Some confirmation of this is found in the data on the time spent in patient care: the total number of hours of direct patient care, and the average number of visits per hour of patient care, both increased as the doctor/population ratio of the division decreased. However, the financial consequences that might be expected to flow from this, namely that doctors in New England would enjoy proportionately lower incomes than those in the other two divisions, were not apparent, for the reduced output of services was compensated to a large extent by higher fee schedules in New England than elsewhere. Moreover, the possible argument that the health care needs of the population of New England were greater than those elsewhere, and that the smaller average number of visits per hour was therefore justified by the greater complexity or intensity of care that was required, was not sustained by the available indicators of need. Of the 'direct' indicators, the infant mortality rate for both white and (especially) non-white children increased as the doctor/population ratio declined, and the same was true of such indicators of social disadvantage as low income, low levels of educational attainment and poor housing amenities.

For all the difficulties inherent in using macrodata in this way, not least its total failure to take any account of the quality (as opposed to the sheer volume) of care that was given by the doctors in each location, Reinhardt's example at least indicates the need to take account of variations in productivity in comparing doctor/population ratios between areas or across time. At face value, the data imply that, although New England clearly had more doctors per unit of population than the other divisions, they may have been working less efficiently (by producing fewer services at a higher unit cost) and may not for that reason be a valid exemplar for emulation elsewhere. Reinhardt summarises the point thus: 'Just what is being proposed when the physician-population ratio of the most richly endowed region or state is proffered as the culturally relevant standard of physician density for the nation as a whole? Is it proposed that all Americans should enjoy the level of health care enjoyed by residents of the most highly endowed region? Or is it suggested that the comportment of physicians in the most highly endowed region be a standard for all American physicians? If the latter - and by proceeding in terms of physician-population ratios one inevitably offers that prescription - then (the data) warrant at least the suspicion that by aiming for the highest prevailing physician-population ratio one may inadvertently accept inefficiently organised and unnecessarily costly medical practices as a national standard. And that inefficiency receives official blessing if public health manpower policy responds passively though conscientiously to whatever dire predictions emerge from this forecasting methodology' (pages 58-9).

In spite of all these difficulties, there is one sense in which the use of comparative ratios opens up the possibility of a more analytical approach to the problem. As some of the reports from the Medical Practices Committee seem to imply, the average list size from another place or time may be selected as a reasonable target because of a general feeling that 'things were better then' than in the present. This approach opens up the possibility of more analytical questioning. In what precise ways were things felt to be better? Is the feeling supported by empirical evidence? Is the future desired improvement best accomplished by reducing list sizes to their former level, or in other ways? Is the desired improvement worth achieving at all? Does it actually matter that things are not as good now as they once were? The material summarised in the preceding section indicates that, in Britain at least, such questions have not commonly been posed.

Professional estimates of population needs

The use of comparative ratios concentrates heavily upon the supply of manpower, and has little to say explicitly about the need for it. There is

presumably the assumption that the target ratio, if achieved, would satisfactorily meet the need for health care among the population concerned; but the examples discussed above do not reveal much systematic concern with the conceptualisation and measurement of need. Studies attempting to make direct measurements of need have broken out of the circularity inherent in the use of comparative ratios, for by quantifying the amount of clinical need in a community and calculating the supply of doctors required to meet it efficiently, an independent justification is offered for the desirable doctor/population ratio. Thus, doctors would be in sufficient supply if they could cope efficiently with the volume of clinically defined morbidity in their practice areas.

The archetype of this approach is the classic study by Lee and Jones (1933). There were four main stages: first, the frequency or occurrence of different disease states in the community was measured; second, a panel of medical experts determined the amount of services required to diagnose and treat each state; third, estimates were made of the average number of services rendered each hour by doctors; and fourth, the opinions of doctors were elicited about the average number of hours that it was reasonable to expect doctors to spend each year in patient care. After doing the necessary arithmetic, Lee and Jones estimated that the required number of doctors in the United States was 135 per 100,000 population, compared with the actual availability of 126 at the time of the study. There was, they concluded, a shortage of doctors.

Perhaps because of the daunting practical problems involved in a study of this kind, the Lee-Jones study was not repeated until 1972 (Schonfield, et al, 1972). Schonfield limited his study to primary care, excluding psychiatric and obstetric care. He first established, through interviews with paediatricians and internists in private practice in New Haven, Connecticut, professional views on good standards of care for 80 diseases in children and 170 diseases in adults. From this, Schonfield produced estimates of the average number of services of each kind that should be provided for the good treatment of each disease, and the average amount of time that should be given to each service unit. He then calculated the number of people that should be receiving the specified kinds and amounts of different services, based on morbidity data from the National Health Survey. Finally, the data on service needs were linked to manpower requirements through the amount of time actually spent each year by paediatricians and internists in patient care (2,227 and 2,198 hours respectively). The results showed that good primary care required the services of some 133 doctors per 100,000 population, compared with the available supply of about 65 primary-care physicians per 100,000 in 1966 and about 59 in 1970. This implied a

substantially greater gap between need and supply in America in the 1970s than Lee and Jones estimated to have existed in 1933, since the Schonfield estimate reflected only a portion of primery care, whilst the Lee - Jones estimate was for total medical care. Schonfield commented that 'with a shortage of this magnitude in the supply of physicians for primary care, not all who require such care receive it' (page 575).

Although this type of approach offers an indication of the number of doctors that might be required in an ideal world of large resources, it embodies a number of weaknesses, some of which are common to the ratio approach. As Klarman (1969) notes, each step is more complex and controversial than one might suppose, and the assumptions underlying them are not always stated and examined. One assumption is that doctors do not vary significantly in their capacity to provide the services that are needed; another is that doctors can be persuaded to enter the appropriate specialties and to distribute themselves geographically in relation to the distribution of the diseases to be treated; a third assumption is that doctors will accept and follow the standards of good practice that are specified for the diagnosis and treatment of each disease; and a fourth assumption is that morbidity data from health interviews are an acceptably reliable and valid indicator of the clinical needs of populations.

A further important difficulty in the Lee - Jones and Schonfield approach is the normative assumption that the needs of the population, identified by the procedures described above, could and should be met. Even if the supply of doctors increased to the indicated level, there is no assurance that each patient with the specified diseases would demand the precise amount of care that the experts had decided he should receive. In fee-for-service systems, the imposition of price is likely to be an effective deterrent for many patients, and even in the absence of price, other social and cultural obstacles may exist. Not only is it doubtful whether the total volume of assessed need would actually result in an effective demand for the appropriate care, it is also doubtful whether the revealed volume of need should be met. At the very least, this is a normative assertion that requires further justification. In claiming that all the benefits of care which are technically capable of achievement should actually be provided, the cost of providing them is ignored. So, too, is the possibility that, in providing these benefits, the doctor's contribution may sometimes be capable of substitution by other less costly inputs.

Perhaps because of these substantial problems, there have been no British studies comparable to those of Lee and Jones and Schonfield et al. The notion of need has entered more obliquely. It has been noted, for example, that Paige and Jones (1966) and the Todd Commission (1968) both discussed the changing

patterns of medical needs and attempted to make allowances for them in specifying future manpower requirements, but the level of generality was far higher than in the two American studies. Indeed, a close reading of both documents fails to reveal the precise ways in which the broad anticipated patterns of need influenced the ultimate judgements about manpower requirements.

A more specific use of indicators of need is represented in a recent study by Buxton and Klein (1979) which attempted to quantify the differences between the actual populations of FPC areas and the notional populations derived from applying a formula similar to that produced by the Resource Allocation Working Party (DHSS, 1976). The principal need indicator used in the study was the standardised mortality ratio, although as the authors pointed out, this may be a less appropriate indicator of the need for primary than for hospital care. The derivation of notional populations of FPC areas, based upon the weighting for need, enables notional average list sizes to be calculated that may be more sensitive to inequalities in the distribution of manpower than the conventional list sizes produced by the Medical Practices Committee. The study therefore offers a more refined index by which an equitable distribution can be assessed. It is, however, of little direct relevance to the problem of need-based judgements about a reasonable list size, for it is concerned more with the way in which an existing stock of manpower should be distributed than with the size of the stock in the first place.

Demand-based assessments of adequacy

An alternative to the need-based approach in judging the adequacy of manpower supply is that based upon patterns of demand. It would be misleading to present this as a distinctive methodology, partly because of the variety of different approaches that can loosely be categorised under the heading, and partly because of the imprecise way in which demand is often distinguished from need. However, as the preceding section of this report has shown, attempts to specify manpower requirements based upon assumptions concerning future patterns of demand have been quite common in Britain, and are for that reason worth examining. In essence, the approach postulates that a shortage of manpower exists when demands for care cannot be met. In its pure form the argument encounters the objection that, in the absence of price, the demand for care is likely to be far higher than even the most optimistic estimates of manpower supply could satisfy, and is for that reason an impracticable indicator of shortage. In practice, therefore, it has usually been restricted to estimating the number of doctors needed in the future in order to maintain the current levels at which demand is met (effective demand), or to allow for a certain degree of expansion in the level of demand that should be met. A simple

example of this is seen in the conclusion of the Willink Committee in 1957 that, in order simply to continue to meet the current level of effective demand among old people, the number of GPs would have to increase by 75 each year as the numbers of elderly people in the population grew (see page 21).

A more sophisticated approach is contained in the Government's discussion paper on medical manpower in the next twenty years (DHSS, 1978). On the assumption that 'the pattern of care will not change radically and that the role of the doctor will continue to be much as it is at present' (para. 35), the paper sets out four specific considerations that might affect future patterns of demand. First, demographic changes affecting the size and composition of the population will be reflected in a changing demand for care. The paper offers four different projections of population growth, based on varying assumptions about fertility, and calculates the number of additional doctors required under each assumption in order to maintain the same doctor/population ratio in 2001 as in 1975. The range is from 5,060 extra doctors on the assumption of a long-term fertility of 2.3 children per family to 560 extra doctors on the assumption that fertility will continue to decline until 1983. Each estimate of the future pattern of population growth is then weighted to take account of 'the differing needs for medical care at different ages' (para. 38). Although the language suggests a reversion to a need-based approach, the weights applied are in fact derived from HIPE and GHS data on age and sex variations in hospital utilisation and GP consultation rates. Other demographic factors which might affect demand (for example marital status and family structure) are mentioned in the paper, but are not incorporated in the weighting process.

The second consideration to be taken into account in estimating future patterns of demand is the scope for reducing regional disparities. The discussion paper notes that 'a high priority might be given to reducing these variations whether by redeployment of posts as they fall vacant or by encouraging differential growth' (para. 44). In the case of general practice, approximately 1,400 additional doctors would be needed in England and Wales and 180 in Scotland to reduce the maximum list size of all doctors to 2,500 without increasing existing list sizes. Unlike the first consideration, which is concerned essentially with establishing the manpower resources required to sustain the existing level of effective demand among a changing population in the future, a reduction in maximum list sizes to 2,500 would imply an acceptance of the need to increase manpower resources in order to allow a higher level of effective demand in future than at present.

The third consideration affecting future demand is that of changes in medical practice. The discussion paper sets out several such changes that might affect the numbers of general practitioners needed in the future, including a shift in emphasis from hospital to community care, changes in the size of practices or of their organisation, an increasing trend towards the centralisation of primary care in health centres, the redistribution of work between medical and non-medical staff, the effects of technological change, and the possibility that doctors will choose to work shorter hours than at present (para. 48). Some of these appear to be located more appropriately on the supply than the demand side of the equation, but it seems reasonable to suppose that such factors as the growth of community care and the development of medical technology will elevate the demand for care above existing levels independently of any increase in demand that may result from demographic changes or from a reduction in maximum list sizes. What is not clear from the discussion paper is the extent to which such elevated demand should be met, whether through an increased supply of doctors or (if appropriate substitutions could be identified) of other resources.

The fourth consideration affecting future demand that is discussed in the paper is that of improvements in the services provided. The paper notes that 'there are certain areas in which there is a clear unmet demand at present, and where there are not enough doctors available to provide a service which health authorities wish to provide' (para. 56). At first sight this appears to be no more than a reiteration of the 'pure' argument that a shortage of manpower exists whenever a potential demand can be identified that remains ineffective because of a lack of doctors. Yet the wording in paragraph 56 hints at an important alternative interpretation, namely that a shortage exists when resources are insufficient to provide the services that health authorities want to provide and for which they are prepared to pay. To the extent that such services are not currently being provided because of a lack of manpower, there will be an increased effective demand for these services in the future if the additional manpower is made available. This is a separate component of demand from that resulting from changes in the size and structure of the population, from reducing maximum list sizes, or from the introduction of new forms of treatment and care. Having made this point, however, the discussion paper fails to specify particular services which are thought to suffer a shortage of manpower in this sense of the term. It notes that 'there is obviously a degree of subjective judgement in deciding the desirable level of care' (para. 58), but the illustrations that are offered of the supposed shortages in particular specialties seem to reflect an ideal level of service provision rather than the level for which authorities would be prepared to pay if the necessary manpower

was available. No observations are made in the discussion paper about the shortage of general practitioners in this context.

The analysis of demand in the Government's discussion paper reveals a greater concern with the elaboration of comparative ratios than with evaluating the adequacy of the base ratio. The paper offers few guidelines by which to judge the reasonableness of existing lists. Projections about the future size and composition of the population, for example, enable estimates to be made of the number of doctors required in order to continue to satisfy the levels of use that are currently made by people in different age and sex groups, but they give no indication of whether the current levels of use are reasonable or not. This looks suspiciously like the application (albeit the sophisticated application) of comparative ratios across time. The data on projected demographic changes indicate how many doctors would be needed in 2001 in order to maintain the level of use that was achieved with the existing doctor/population ratio in 1975; but they do not explain why that particular level of use should be regarded as reasonable. A separate justification is needed. Likewise, the proposition that regional disparities should be reduced by lowering the maximum permitted individual list size to 2,500, closely resembles the way in which the Medical Practices Committee has made use of comparative ratios between areas. Why is the criterion of a designated (under-doctored) area selected as the 'correct' target for maximum list sizes? In fact the discussion paper appears to acknowledge the difficulties involved in this kind of argument. It notes that 'there is no simple relationship between the staffing levels in the various regions and the level of service they provide' (para. 45), and in the case of general practice, that 'a high list is not necessarily an indication of inadequate care or of over-burdened doctors' (para. 47). Yet there is no real consideration of the precise ways in which such insights might be utilised in shaping opinions about what is reasonable.

The second part of the analysis of demand runs into similar difficulties. It is no doubt true that changes in medical practice of the kind discussed in the paper, and improvements in the quality of service, will add to the future potential demand for the medical services generally and the primary care services particularly. If that additional demand is to be satisfied, additional manpower will be required; yet it is presumably unrealistic to expect that enough doctors can be supplied to meet the whole of it. At what point, and on what basis, is the limit to be set? What level of effective demand is to determine the number of patients for whom a GP can offer reasonable care? The discussion paper is right to draw attention to those factors which might elevate the potential demand for care in the future, but it appears to

offer few guidelines in deciding the point beyond which such demand cannot, or should not, be met.

Cost-benefit perspectives

The literature contains another distinctive mode of analysis, associated particularly with a microeconomic approach to supply and demand in non-market contexts, which offers the possibility of overcoming some of the difficulties raised in the Government's discussion paper. In essence, it involves the application of the principles (and, where possible, the formal methods) of cost-benefit analysis in aiding decisions about the worthwhileness of particular programmes or activities. In the simplest terms, an activity is worth doing if the benefits outweigh the costs; and the choice between two activities, in each of which the benefits outweigh the costs, will lie with the one that shows the greatest benefit per unit of cost. As Williams (1974a) puts it, 'cost-benefit studies stress the simple truth that the decision whether or not to pursue a particular course of action depends on both costs and benefits'. The trick, of course, is to quantify the benefits in terms that are commensurable with their costs; and this involves difficult questions about the ways in which different kinds of benefits are measured and valued, and the actors who are involved in the valuing process. It would be foolish to believe that the widespread application of formal cost-benefit techniques to decision-making in the health services is just around the corner, but progress is being made in the development of conceptual tools for measuring and valuing the benefits of health and welfare services (Culyer, et al, 1971, 1972; Rosser and Watts, 1972; Williams, 1974b; Davies, 1977; Wright, 1978), and the principle of thinking about costs and benefits together may still be enlightening to the decision-maker even in the absence of formal cost-benefit studies.

Central to the cost-benefit approach is the axiom that resources have alternative uses, and that the cost of deploying resources in one particular area is the benefit foregone in not applying them to other areas. Shannon (1968) has applied this argument to the definition of manpower shortage. Dismissing as 'spurious' the argument that 'there must be no shortage at any price', Shannon points out that opportunity costs are inescapably attached to any increase in the output of medical graduates. The cost of producing more doctors must be seen and evaluated in terms of other possible ways of using the resources, whether financial, human or physical. He concludes that 'shortages of manpower, defined by simple subtraction of likely supplies from likely "requirements" should not necessarily be taken seriously by the policy-maker. What he must be concerned with are the costs of producing this "shortfall" weighed up against the benefits of its production. One would

not like to pretend that this is a simple matter, but it is the proper way of analysing the situation' (page 62). This argument generates the proposition that doctors are in short supply when they are unable to provide the benefits of care which 'society' considers to be worth the cost of providing.

Culyer (1976) sets out the argument in more general terms. He first draws an important distinction between behavioural and normative notions of shortage. A behavioural manifestation of shortage is seen when potential demand exceeds supply and when, because price is either non-existent or artificially constrained, other rationing or filtering mechanisms must operate. Queues and waiting lists are the most obvious examples of such mechanisms; they are inevitable behavioural consequences of the elimination of price. For evaluative purposes, however, a definition of shortage is needed that transcends the particular methods used to allocate resources. Although a queue may exist for a particular service, it does not follow that it must necessarily be eliminated. A policy decision may have been taken that only a limited amount of that service shall be provided, and as long as sufficient resources exist to provide the specified amount of the service, there is no shortage, even though queues and waiting lists build up. Culyer notes that 'while shortages in the normative sense cannot be identified without assigning values to resources and outcomes, they can most certainly be identified without using market prices; for these can ... give misleading information about the true social cost and benefits' (page 72). A shortage, then, exists when a programme that is worth doing (that is to say, where the benefits are thought to outweigh the costs) cannot be done as planned because of insufficient resources.

A similar argument is advanced by Cooper (1975), who argues that 'whether we have sufficient doctors depends upon a listing of overall priorities for skilled manpower throughout the economy ... How many doctors there are is a decision reflecting society's current priorities within the overall constraints imposed by total resources. Clearly, such decisions will be less arbitrary if we have a clear picture of the tasks we wish doctors to perform. For example, both preventive medicine with yearly check-ups and a planned expansion of the use of medical auxiliaries with growing emphasis upon community care would profoundly influence the desired doctor-population ratio. To date society has given no clear lead as to what it expects of its doctors... Current numbers of general practitioners are perfectly consistent with maintaining that there exists either a surplus or deficit according to the view taken as to a doctor's pastoral and medical role' (page 44).

The general principle that decisions about the required number of doctors should be based upon the things that society regards it as worthwhile for them

to do, and that these things in turn should reflect as clearly as possible a conscious weighing of the costs and benefits of doing them, points to two possible ways of making decisions about target list sizes. One, which is usually associated with a highly rational model of decision-making, is to wipe the slate of current medical practice clean, and to start afresh in the task of constructing the ideal role of the GP in accordance with the cost-benefit preferences of society. Once the ideal role has been constructed, it would then remain to calculate the volume of medical manpower required to fulfil the role under conditions of optimum efficiency. This is an important innovatory approach that might yield useful lessons even in the absence of its full implementation, but there are obvious difficulties in pursuing it fully. It seems unlikely that much consensus could be forged about the roles and functions that 'society' would ideally wish its general practitioners to perform, and even less about the point at which the benefits following from the performance of such functions cease to be regarded as worth the costs of the extra doctors needed to perform them. Descriptive accounts of the work of general practitioners illustrate the multifarious, shifting and idiosyncratic goals towards which much of their work is directed (Lane, 1969; Hale and Roberts, 1974; and especially Berger and Mohr, 1976), and normative prescriptions of the GP's role are notable for the breadth of tasks which they encompass (Central Health Services Council, 1963, 1971; Royal College of General Practitioners, 1969, 1970, 1977; European Economic Community, 1974; Hunt, 1957; Noble, 1978, etc.). Hence, although the Office of Health Economics (1974) is right to observe, in relation to the multiplicity of possible roles which the GP might fill, that 'the extent to which there can be held to be too many or too few family doctors ... depends upon how much priority would be given to each of the roles in an ideal job specification for the family doctor and his team' (pages 24-25), it is difficult to see in real life how such an ideal job specification might emerge from among all those with an interest in its definition.

An alternative, more modest approach to the problem may feasibly be to hand in the philosophy of 'disjointed incrementalism' (Maddox, 1971) and the concept of marginal analysis. This approach begins not by wiping the state clean and stating an ideal set of objectives or priorities for future achievement, but by mapping out the changes resulting from marginal increases or reductions in existing resources, and then deciding on the basis of such empirical evidence whether the changes have been worth the cost or savings involved in achieving them. Further experimental increases or reductions in resources would continue to be made until the point is reached at which further changes in output are judged not to be worth the costs or savings involved in achieving them. There

seem to be several ways in which the incremental approach might be applied in a practical fashion to the task of deciding a reasonable list size.

First, it may be possible to create an experimental situation in which the effects of changes in the input of resources are assessed in a controlled manner. The case for this has been put by several writers. For example, Buchan and Richardson (1973) observed that 'the proposal to reduce list size could only be seriously considered if evidence and judgement clearly pointed to a likely improvement in the "quality of primary, personal, continuing, comprehensive, family medical care" ' (page 41); and they commented that 'it is not idle to wonder how family doctors would work, given much more time than is usually available. Suppose, for example, that patient demand per doctor were reduced by a combination of smaller list size, more nursing and ancillary staff, and greater efficiency in practice organisation? Would this increased time allocation per patient be used to the greater benefit of patients or would it simply result in a Parkinsonian situation in which history, examination, investigation and treatment were wastefully extended to fill the time available? ... These are difficult questions to which experimental answers are urgently needed' (page 40). In a similar vein, Acheson (1975) has argued that a logical first step in approaching such questions would be to find out what GPs would actually do if they had more time to spend with patients, and whether what they did with it would produce an outcome that could justify the extra cost involved.

A second way of applying an incrementalist approach to the problem of deciding a reasonable list size would be through the judgements of practitioners about the additional services they felt they would be able to provide if they had more time or other resources. Such judgements might be elicited either in general terms (for example by asking GPs to describe how often, and in what ways, they feel that constraints of time or the pressures of demand depress the range or quality of service they could give), or in relation to a series of specific consultations in which they engage. As with the first approach, separate judgements would be needed as to whether the prophesised improvements in service output would justify the additional time (or other resources) required to achieve them, but at least there would be an empirical indication of what was being traded against what.

A third approach, which unlike the previous two does not require original fieldwork, would be to scrutinise the research literature for evidence of systematic variations in the content of practice between doctors with different practice sizes. In what ways does the output of GPs with smaller lists differ from those with larger lists, and what are the different outcomes? Again,

further judgements would require to be brought to bear in evaluating the observed differences in practice output against the resource implications of reducing all list sizes to the lowest current levels, but in doing this, clear indications should be available of the anticipated benefits that would result from any such expansion of resources. As a preliminary exercise, the research literature is reviewed in this way later in this report.

Summary

A review of the literature has identified five reasonably distinctive approaches to the question of the adequacy of the supply of medical manpower and hence, indirectly, to the judgement about a reasonable list size for general practitioners.

First, in contexts where market forces are maximised, shortages are signalled when the demand for physicians' services exceeds their supply at current prices. In a free market prices will adjust to equilibrate supply and demand. In reality, there are few (if any) governments which allow the unfettered operation of market forces in health care, and imbalances between supply and demand are inadequate indicators of shortage in health care systems (such as the NHS) where direct pricing has at most a residual role.

Second, shortages have been inferred when the doctor/population ratio has been lower than that in another place or another time . This approach to the definition of shortage has been fairly widely used in the United Kingdom in judging the reasonableness of list sizes in general practice. The most fundamental drawback to this approach is the assumption it embodies that the base ratio against which existing ratios are judged is adequate and sufficient. Such an assumption, however, merely displaces rather than eliminates the need to specify a desirable doctor/population ratio, although this has not commonly been done in contexts where comparative ratios have been used. Further difficulties surround the use of comparative ratios unless the doctor component of the ratio is weighted to allow for variations in productivity and the patient component for variations in demand between different sub-groups. A further problem created by the way in which comparative ratios have sometimes been used is the inflationary thinking that flows from the belief that the best achieved local ratio should always be an overall target for the future.

A third approach to the definition of shortage is through professional estimates of need. The basic steps involved in this approach are first, to estimate the prevalence of different conditions of need in the community; second, to determine the amount of service required to treat each condition

to a satisfactory level; and third, to measure the average number of hours spent each month or year by doctors in providing the services to their patients. There are, plainly, substantial conceptual and methodological problems involved in actually producing estimates of need based upon this approach, and it involves important assumptions about the behaviour of doctors and patients that may not be justified. In particular, the assumption seems to be made that the total amount of identified need should be met through the supply of the indicated number of doctors, although this is a normative assumption that ignores the costs of increasing the supply of doctors to the level indicated.

Judgements of shortage based upon demand constitute a fourth approach that is discernible in the literature. It appears to be used less as a means of determining the adequacy of existing ratios than as an assurance of the future continuation of the patterns of medical care inherent in them. Beginning with existing levels of effective demand, forecasts are made of future levels of demand in the light of projected changes in the size and composition of the population and in the organisation of medical practice, and further adjustments may be made to allow for some improvements in the scale or quality of services and for reducing geographical imbalances. Demand-based notions of shortage, at least as they appear to have been incorporated in planning future manpower requirements, are very close to a sophisticated application of comparative ratios across time. They are concerned essentially to ensure that future demand is met at the same (or higher) levels than at present, allowing for projected changes in population structure, practice organisation, and so on. This is, plainly, a more elaborate approach to the future than the simple projection forward of existing, unweighted doctor/population ratios, but it does little to clarify the acceptability of the basic demand levels upon which the whole edifice is built. In the absence of direct price, the potential demand for GPs' services is likely to be considerably in excess of the available supply of manpower, but it is difficult to discern in the demand-based approach any indicators of how the limits are to be set.

The fifth approach to shortage is characterised particularly in the writings of British economists concerned with the application of cost-benefit analysis to the public sector. In this view, a resource is in short supply when something that is worth doing (that is, something in which the benefits are judged to outweigh the costs) cannot be done because of the non-availability of the resource. General practitioners' list sizes would thus be judged to be unreasonably large when they impede the provision of the scale and quality of services that 'society' regarded as worth the cost of procuring. This perspective emphasises the important distinction between the input of manpower as a

resource and the production of services as outputs: what is to be valued is not a resource for its own sake, but rather the services produced from that resource; and this in turn involves an estimation of the benefits to be derived from the consumption of the services. By concentrating on the service outputs of general practitioners and the way they are valued by those who consume them, the cost-benefit approach offers the possibility of breaking with the tautalogous argument about comparative ratios, but it nevertheless generates severe procedural problems in measuring and valuing benefits in ways that enable them realistically to be set against costs.

A FRAMEWORK FOR JUDGING A REASONABLE LIST SIZE

The two preceding sections have attempted to examine the foundations upon which judgements about a reasonable list size for general practitioners - or, more generally, judgements about the adequacy of the supply of medical manpower - have been, or might be, constructed. Two inter-related conclusions seem to emerge from this examination as necessary precursors of any further analysis.

The first conclusion is that there is no objectively 'correct' ratio of doctors to population. To search for a reasonable list size in the expectation that, once discovered, it will instantly and universally be recognised as such, is futile. It is clear not only that general practitioners do actually cope with lists of differing sizes, but also that they hold a wide range of views about an ideal list size. In selecting a target list size as the basis for determining the future supply of medical manpower, policy-makers are therefore exercising a judgement about the most satisfactory way forward in a given context; they are not enunciating an immutable law of medical care, the transgression of which would inevitably incur some kind of administrative retribution. The task of the research worker in this view is not to discover the Holy Grail, since it does not exist, but rather to provide the kind of information that will enable the policy-maker to maximise the confidence with which he makes and can defend his judgement.

The second conclusion is that no one methodology or set of considerations is uniquely 'correct' in making the judgement; indeed, although for analytical purposes a number of different methodologies have been described in this report, in reality they have elements in common. No realistic judgement, for example, could focus exclusively upon the needs of a population for medical care without taking account of the nature and quality of the care that is produced to meet that need, and conversely the concept of 'quality of care' in general practice must surely be related to the pattern of needs in the practice population as well as to the specific needs that individual patients bring to the doctor. Nevertheless it seems clear from the evidence presented so far in this report that some concepts and methodologies are likely to be more relevant than others, and certain kinds of information may be more relevant than others in increasing the rationality of the judgement. But before attempting to identify and assemble this information, a note must be appended about the notion of 'rationality', for although the word has a clear ring of objectivity about it, one man's rationality may sometimes be another man's myopia. The general tenor of this report has been consistent with the view that a rational

judgement about a reasonable list size is one that maximises the value obtained from health care resources, but this view may be rational only from certain perspectives. From other perspectives it may be equally rational to select the size of list that would, say, yield a maximum income to the doctor for a given volume of work or that would minimise the use that patients have to make of hospitals. In asserting that a rational judgement is one that maximises the utility of resources, it must be recognised that a particular connotation is being placed upon the notion of rationality.

Against the background of these general comments, this section aims to generate a set of questions that can be investigated empirically and that might yield information to increase the rationality (as defined) of the policy-maker's judgement. The arguments rehearsed so far in this report have highlighted two steps that need to be taken in forming this judgement. The first step is to understand the consequences of differing list sizes in varying contexts. A common thread running through many of the arguments is that, in general terms, large lists are less to be desired than small lists because of the unfavourable consequences flowing from them: they leave a larger volume of unmet need in the community, or they force the doctor to practise a lower standard of medicine, or they impose an unreasonable workload on the doctor, and so on. It has been shown, however, that such arguments have been based more upon supposition than fact, for they have made little reference to any clear evidence of the associations between list size and any of the possible dimensions of outcome that might be relevant to the judgement. One important research task, then is to document these associations.

By itself, however, an understanding of the range of consequences flowing from variations in list sizes would not be sufficient grounds for a rational judgement about a reasonable list size, for it would fail to distinguish between consequences that are regarded as worth achieving and those that are not. For example, even if the quality of care can be shown to be inversely related to list size, this fact alone would not aid the judgement about a reasonable list size for it would fail to indicate the point at which gains in quality cease to justify the costs of achieving them through reductions in list size. The second step to be taken is therefore to assign some kind of social valuation to the consequences stemming from variations in list size, and to do this in a way that enables these valuations to be weighed against the costs of procuring them. Although this step cannot properly be taken without that foundation of reliable evidence which it is the task of the researcher to provide, the step itself has to be taken by the policy-maker, not the researcher, for it is only the policy-maker who has the authority, acting on behalf of and in some sense

accountable to 'society', to make the cost-benefit choices that are implicit in the process of weighing resources against consequences.

The consequences of variations in list size

In approaching the task of documenting the consequences of different ratios of population to doctors, ways must first be found to classifying those consequences. Three distinct levels of classification can be identified. First, there will be interactive consequences for the doctors' consulting times and rates. Assuming for the moment that all other things are equal (though we shall have to return to that assumption in due course), a GP with, say 3,000 patients must differ in his pattern of work from one with 1,500 patients in one, or a combination, of the following ways: he will maintain the same consultation rate by spending less time, on average, with each patient seen; he will have a lower consultation rate through seeing the same total number of patients during the course of a standard working week or year; or he will work for longer hours. The doctor with 3,000 patients on his list cannot, in the course of a standard working week or year, maintain the same consultation rate and also spend the same average amount of time in each consultation as the doctor with only 1,500 patients.

A relevant first step in judging a reasonable list size would therefore be to document typical variations in the consulting times and rates of GPs with different list sizes. We do this later in the report, using material that has already been published. However, an understanding of such variations remains an inadequate basis for the judgement, for it gives little indication of the significance that should be attached to them. It is useful as a first step to know whether doctors with large lists typically have a lower consultation rate than those with small lists, or whether they have shorter consultations, but having gained this knowledge, a judgement must still be made about whether these differences matter. A second level, therefore, at which the consequences of different list sizes may be analysed is that of the variations in the style and content of practice which flow from them. If, for example, doctors are typically coping with larger lists by reducing the average amount of time they spend with their patients, it would be helpful in evaluating the significance of this to know how these shorter consultations differ from those of their colleagues with smaller lists. What are these doctors not doing which those with smaller lists and longer consultation times are doing? Alternatively, if GPs are typically coping with larger lists by reducing their consultation rates, it would be relevant to know whether a lower proportion of the practice population is consulting these doctors each year than in smaller practices, or whether the same proportion is consulting, but on fewer occasions each. In

the first of these two cases, questions need to be asked about the nature of the care that is given to patients in different practices. Who are the patients who are less likely to see their doctors at all in larger than in smaller practices, what are their needs, and how else are they being met? In the second of the two cases, questions need to be asked about the continuing or follow-up care received by patients in different sized practices. Do patients in larger practices have fewer consultations, on average, for each episode of illness than those in smaller practices, and if so, in what ways does their total pattern of care differ?

A relevant second step in judging a reasonable list size would therefore be to document typical variations in the content of care given by GPs with differing list sizes. Yet even if this can be done (and we attempt to do so later in the report), the strongest basis for the judgement would remain to be established, for the ultimate touchstone by which 'society' can judge the value of different doctor/population ratios is by the quality of care that is delivered from practices of different sizes. A third level, therefore, at which the consequences of differing list sizes may be analysed is that of the variations in the quality of care which flow from the different ways of coping and the different patterns of practice that characterise doctors with different practice sizes. Although attempts to define and measure the quality of care in general practice are still at an elementary level, the concept of quality is well-rooted in contemporary notions of a reasonable list size. As has been shown, for example, the BMA has consistently argued that lists of 2,500 and above do not allow GPs to spend enough time in consultation with their patients to achieve an acceptable standard of care; but this has the status only of an hypothesis, and requires careful testing. A five-minute consultation may yield a better standard of care than a one-minute consultation, but would a ten-minute consultation (the achievement of which should, in the opinion of Buchan and Richardson, 1973, be a top priority) enable the doctor to give better care than a five-minute consultation? What extra things would he do with the additional time, and would the outcome to the patient be any more favourable?

The problem of non-uniformity

The argument now encounters a substantial difficulty. So far, the explicit assumption has been made that GPs differ in nothing but the number of patients on their lists, and that all practitioners with comparable lists will provide the same pattern and standard of care. Likewise, in comparing, later in this report, the practice patterns of GPs with different list sizes, the assumption is made that by reducing list sizes to, say, 2,000 patients, all doctors would behave in the future in the way that those with 2,000 patients

behave now. This type of assumption is inevitable in any consideration of general practice as a whole. It runs through the arguments of all the sources cited in the first section of this report. Yet it is plainly an erroneous assumption, and is vitiated by at least three major considerations.

First, GPs are not confronted with a uniform amount and intensity of clinical need in their practice populations, even when those populations are of comparable size. If therefore, the notion of the 'quality of care' is applied to the practice population as a whole, and not just to those members of it who happen to cross the threshold of the GP's surgery, it is to be expected that doctors will differ in the quality of their output even when they have comparable resources and list sizes. The point has been argued most eloquently by Tudor Hart (1976) in refuting the charge that his practice in South Wales is pursuing eccentric policies of care. The practice, which participated in the Second National Morbidity Study, yielded a consultation rate during the year of the study that was 64% higher than the national average; but Tudor Hart produced a variety of data to show that the whole of the excess lay in an 'increased rate of episodes of sickness, rather than revisits and follow-up' (page 885). It would consequently be damaging and inequitable to apply the same canons of reasonableness to list sizes in this part of South Wales as to areas of the country with fitter populations and lower levels of clinical morbidity.

Second, GPs differ in their personal and professional characteristics in ways that affect their style and quality of work. This point has already been stressed, and need not be repeated at length. An apt illustration is offered by Taylor (1954) in describing two doctors in similar areas with similar list sizes. 'One doctor complained bitterly of work resulting from the National Health Service. He and his two partners looked after 9,000 patients in an unhealthy industrial area. They were good doctors, but poor organisers and their surgery conditions were appalling. Here in fact there was no discipline, but rather a diffuse and querulous kindness. Another doctor presents a very different picture. He is brusque and gruff and wastes no words on the lead-swinger. Every complaint is investigated with complete efficiency. He has a fairly high degree of contempt for the human race, but the highest standards of conduct in dealing with them. He is respected by his patients rather than popular with them' (pages 84-85). The uniqueness of each general practitioner, personally and professionally, will always confound the assumption that a predictable pattern of practice and quality of care will flow from the achievement of specified doctor/population ratios.

Third, as hinted in Taylor's description, practices differ in the efficiency with which they are run. The concept of efficiency is pertinent to the judgement of a reasonable list size because of the possibility that inefficient ways of doing things might artificially restrict the quantity and perhaps also the quality of care that can be given. The normative assumption is generally made that, provided other objectives of primary care (such as accessibility) are not infringed to an unacceptable extent, a more efficient way of doing things is to be preferred to a less efficient way. If this is indeed a widely held assumption, it would seem to follow that an increase in manpower in order to overcome deficiencies in services caused by inefficient practices would not be regarded as justified. An extreme example might illustrate the point. A GP who employed no ancillary help at all in his practice, and who carried out all his consultations as home visits, would have a low consultation rate and might, as a result, be providing a pattern of care to his practice population below the level for which 'society' would be willing to pay. Clearly, however, the initial solution would lie not in the supply of another doctor to the practice, but rather in improving its efficiency, for it would presumably be possible to increase the output of services, with no diminution in their quality, by engaging ancillary help in the practice and transferring many of the consultations from the patients' homes to the surgery. Unless, therefore, the efficiency of each practice is standardised, divergent patterns and standards of care will continue to emerge from practices even of the same size.

It is difficult to know how to tackle the problem of non-uniformity. Three possible ideal stances may be adopted. The first is to insist upon the uniqueness of each practice and practitioner, and to deny the legitimacy of any attempts to generalise. From this perspective, any generalised statement about a reasonable list size would be rejected; only those that related to specific, unique contexts would be acceptable. Whatever the rigour of this stance, it is clearly not very helpful for those whose decisions must necessarily reflect generalised views, experiences and evidence. At the other extreme, a second possible stance is simply to ignore the uniqueness of each practice. From this perspective, data are analysed and conclusions are drawn at a system level, the assumption being made that the individualistic features of practices and practitioners are randomly distributed among the groups in the analysis. For example, in comparing the pattern and quality of care between groups of doctors with different list sizes, the assumption may be made that differing categories of population needs and differing degrees of efficiency of practice organisation are randomly distributed among each group. This method of approach will tend to yield a single target list size, but it may recognise the need for

variations around the target in specific contexts, depending upon local factors. For example, in advocating a target average list of 1,700 patients (see page 9), the BMA is presumably not claiming that all doctors should ideally have this number of patients, but rather that this should be a norm around which individual variations may well occur in response to local needs and conditions. A third stance, whilst accepting the uniqueness of each practice, professes the legitimacy of classification systems that permit the comparison of 'similar' with 'similar', if not actually 'like' with 'like'. The aim here is to derive valid criteria for grouping similar kinds of practices together, with variations in the quantity and quality of service output being compared between practices of different sizes within the same groupings. For example, the problem of allowing for variations in the levels of clinical need among practice populations might be eased by concentrating the analysis within a defined area that is believed to reflect a reasonably homogeneous epidemiological make-up. Likewise, variations in the efficiency with which practices are organised and operated might be accommodated by comparing different-sized practices in similar premises and with a similar availability of primary care team members. It would presumably be very difficult to make a comparable allowance for the variability in the personal characteristics and professional abilities of GPs, but it seems reasonable to assume that, if satisfactory account could be taken of the variability in population needs and practice organisation, the significance of this latter factor might diminish. One important consequence of this third approach to the problem of non-uniformity is that it would generate a range of reasonable list sizes for different contexts, rather than one single figure for the country as a whole.

Of these three approaches to the problem of non-uniformity, the first seems to be unrealistically rigorous and the second unrealistically assumptive. The third approach seems to offer the best compromise between on the one hand the need to draw generalised conclusions about general practice if manpower policy is to advance, and on the other hand the need to acknowledge the variability that exists in primary care. However, it will become plain in the next section, when we review the research literature for answers to the questions that logically emerge from this framework, that existing data rarely permit the philosophy of this third approach to be applied in their analysis. It will be seen, therefore, that notwithstanding all the problems arising from the assumptions it embodies, it is the second approach that has had perforce to be adopted in much of the next section of the report. However, this brief discussion of the possible options that are available for coping with the problem of non-uniformity is pertinent to the concerns of the concluding section of the report, namely the areas in greatest need of further investigation.

Questions to be addressed to the research literature

The framework that has been developed in this section yields up a set of fairly specific questions, the answers to which might provide a basis for judging a reasonable number of patients for GPs to have on their lists. Of course, to the extent that the relevance of the framework is discounted, so also will the pertinence of the questions. In particular, it may be objected that in focusing upon those aspects of practice that can most easily be measured, the opportunity is missed of reviewing the full range of possible tasks in general practice. There is really no adequate defence to this objection; all that can be said is that this report is proving difficult to write even within the limited framework it presents. To add to its complexity would defeat the writer, and probably the reader also.

1. How much time do general practitioners spend in patient care, and how does this vary with variations in list size?

2. What is the average amount of time that general practitioners spend in consultations with their patients, and how does this vary with variations in list size?

3. What is the average number of consultations per patient per year, and how does this vary with variations in list size?

4. What proportion of the population consults a general practitioner each year, and how does this vary with variations in list size?

5. What is the ratio of doctor-initiated to patient-initiated consultations, and how does this vary with variations in list size?

6. What is the content of care in general practice, and how does this vary with variations in list size?

7. What is the quality of care in general practice, and how does this vary with variations in list size?

8. (Ideally) how do all of these variations themselves differ between different practice locations and contexts?

In the next section of the report we review the extent to which answers to these questions can be unearthed from published research material.

THE TIME SPENT IN PATIENT CARE

The previous section attempted to construct a coherent framework within which to locate the meaning of the question about a reasonable list size, and it concluded by posing a set of questions derived directly from the framework. The purpose of the remainder of this report is to review the extent to which answers to these questions are already available in the research literature. This is done in the belief that it is sensible to look carefully at what is already known before embarking upon further fieldwork studies, but it will become apparent that, in spite of the large number of studies of general practice that have been carried out, there are serious problems in synthesising their results. Many studies have been conducted by individual GPs on their own practices, and there are consequently serious inconsistencies in definitions, time periods, practice contexts, and so on. An attempt is made to evaluate the reliability and validity of the data that are summarised in this report, but readers of the report cannot escape the responsibility of forming their own judgements about the extent to which the available information offers a solid foundation for the policy judgements implicit in the framework.

The first two questions summarised at the end of the preceding section concerned the time spent by general practitioners in various aspects of patient care. It was argued that GPs with larger lists could maintain the same consultations rates as those with smaller lists only by spending more total time in the different activities of patient care, or by reducing the average amount of time they allocated to each consultation, or both. This section tries to assess the extent to which the consequences of different list sizes are reflected in these dimensions of time.

Total time spent in patient care

Early studies cited by the College of General Practitioners (1965) showed that 'a doctor is in contact with his patients for approximately 35-42 hours per week' (page 23), although it was noted that the range around this average was wide. However, not all the sources quoted in the College's report did in fact include a reference to the total number of hours spent in patient care. Of those that did, Mair and Mair (1959) reported an average of just over 43 hours spent in surgery and home consultations during a six-day week by each of two partners in 1957-8, and a further 9 hours spent in administration; and Crombie and Cross (1964) reported an average of just over 20 hours per week spent on contact with patients by a GP in suburban Birmingham in 1952, with an additional 5 hours each on travelling, administration and 'just talking', and 14 hours on medical activities outside the NHS.

The College also reported in 1965 the results of a survey of 370 GPs, carried out in August 1964, which showed an average of 35½ hours per week spent in the consulting room and on home visiting, with a further 3½ hours spent on hospital and other work (College of General Practitioners, 1965). Because the results of the survey were criticised as unrepresentative of the year as a whole, a subsequent study was made of the hours worked during a busy time of the year (Eimerl and Pearson, 1966). The study was conducted by the Merseyside and North Wales Faculty of the College, and 134 faculty members kept records of their work during a one-week period in February 1965. The results showed that, in the course of a full week, the respondents spent an average of 36½ hours in consultations in the surgery and the home, and a further 7½ hours on hospital and other work. The August 1964 survey gave no indication of the range of responses around the mean values, but the February 1965 survey did. Just over a tenth of the 134 doctors spent fewer than 30 hours per week on all activities, a similar proportion spent between 30 and 35 hours, a quarter spent between 35 and 40 hours, a fifth spent between 40-45 and 45-50 hours, and the remainder (about 12 per cent of the total) spent more than 50 hours.

Further information about the time spent in patient care is available from studies contained in the Report on the Practice Nurse (Royal College of General Practitioners, 1968). Hodgkin and Gillie reported the total number of hours worked by two doctors in an urban housing estate practice between 1965 and 1968. Recording was done at three different points in time: at the outset of the study period, after the employment of a practice nurse, and after the introduction of planned changes into the practice routine to save time. At the beginning of the period, the two doctors each worked, on average, for about 13 hours per week on visits, 17 hours on surgery consultations, 3½ hours on ante-natal and post-natal clinics, 4 hours on administration and almost 4½ hours on driving. By the end of the study period, these figures had changed to 8½ hours per week on visits, 17½ hours on surgery consultations, 2½ hours on ante- and post-natal clinics, 3 hours on administration and 3½ hours on driving.

In the other three practices included in the Practice Nurse Study, the average number of hours worked per week by the GPs in each practice were 33, 50 and 57 before the introduction of the nurse, and 34, 41 and 53 after the introduction (these figures are calculated from data on page 17). The breakdown of these times into different activities indicated that between about one-third and two-fifths of the total time was spent on new consultations and visits, and a similar proportion on 'all other medical work'.

Of slightly more recent origin than the practice nurse studies is a study of the work of all the general practitioners serving a whole community,

consisting of a small Scottish town and its rural hinterland (MacDonald and McLean, 1971). Eleven practitioners participated in the study by keeping workload records during each of three separate weeks in November 1967 and February and July 1968. Summing the data across the three weeks, MacDonald and McLean found that each doctor worked for an average of just under 39 hours each week, with a further 76 hours 'on call'. Of this total, an average of 17 hours were spent each week on surgery consultations, 8 on home and hospital visits, 7 on travelling, 3 on administration and 3 on other work. The individual variations around these averages were, however, quite large. For example, the range in total hours worked was from just under 26 to 53½; in hours on call from 38 to 103; and in hours spent on surgery consultations from 12½ to 24.

Two other recent studies offer further information. Garraway (1973) found that a single-handed doctor in North-west England worked for an average of 37 hours per week on surgery and home consultations, but no information is given about the time spent on other activities. Hughes et al (1976) reported the average number of hours worked per week by each doctor in a five-man practice in Derbyshire between 1971 and 1974. The average amount of 'contact time' per doctor (including face-to-face contact time as well as movements in and out of surgery, note-writing, and so on) ranged between 37⅞ and 40⅜ hours per week over the four years, and the average time 'on call' ranged between 28 and 32 hours per week.

Three conclusions may reasonably be drawn from this mixed bag of information. First, the average number of hours spent each week by general practitioners in consultation with their patients and in immediately associated activities such as administration and travelling to home visits is between about 35 and 42. The majority of studies, whether of single practices or of groups of practitioners, lie within this range, irrespective of year, place or season. Second, however, the distribution of hours worked by individual doctors covers a wide range, from perhaps as few as 25 to as many as 55. Third (and of most significance for the argument in hand) there is some indication of a positive association between the number of hours worked and the size of the practice. Of the studies quoted above (which claim to be the principal studies in the literature in which total time is reported), no information about list size is given by the College of General Practitioners (1965), Crombie and Cross (1964) and Eimerl and Pearson (1966). However, among the four practices in the Practice Nurse Study (Royal College of General Practitioners, 1968) a fairly regular trend is discernible in which, both before and after the introduction of a practice nurse, the reported average number of hours worked by the GPs in each

practice increased from about 33-34 per week in the practice with an average list of 2,490 patients to about 53-57 per week in the single-handed practice with some 4,200 patients. This latter practice stands out as having had by far the most hours worked by the practitioner.

A similar story is revealed in the study by MacDonald and McLean (1971). Among the six practices represented by the eleven GPs in the study, the two with the largest average list sizes also contained the GPs working the greatest average number of hours per week. Hughes et al (1976) likewise report that, when the average list size in the Derbyshire practice rose in the period 1971-74, the average number of contact hours per GP also rose, and when the average list size fell, so too did the number of hours. Indeed, the average number of contact hours spent by each GP on his registered patients each year remained virtually constant throughout the four years for which the data are reported. It is this kind of evidence which supports the suggestion of a positive relationship between list size and hours worked. However, there are some inconsistencies in these studies in the range of tasks included, and other investigations have reported conflicting results. For example, Hodgkin and Gillie in the Practice Nurse Study reported a decline in the average number of hours worked during a three-year period in which list size increased; Fry (1972) reported a decline in the average number of hours per week in contact with patients from 43 to 27 during a period (1954-72) in which the average list size in the practice increased from 2,700 to 4,500; and at the other extreme Mair and Mair (1959) worked some 52 hours each per week in 1957-8 (including administration) with an average of only about 2,000 patients.

A different kind of evidence about the relationship between list size and working time is seen in the results of surveys such as that undertaken by Mechanic (1974) in which the perceptions of doctors are sampled. Among the 807 doctors in Mechanic's survey, the proportion who reported themselves to be very dissatisfied with the amount of time and effort they had to devote to their practices increased from 3 per cent of those with lists under 1500, to 16 per cent of those with lists between 2,000 and 2,499, and to 26 per cent of those with lists above 3,500. The proportion of doctors reporting themselves to be very dissatisfied with the amount of time given to patients likewise increased regularly with increasing list size, as also did the proportion who reported that the pressure of time affected their behaviour. These results, from a large number of doctors, strongly indicate the greater dissatisfaction and pressure felt by GPs with larger than with smaller lists, but they do not directly reveal whether doctors respond to that pressure by spending more time on the job. It seems more reasonable to conclude from the meagre information that is available that the total amount of time spent in the practice is as

likely to reflect a range of personal and idiosyncratic considerations as it is to reflect any systematic variations in list size.

Average time spent in consultations

The average amount of time that GPs spend in each consultation is not necessarily related either to list size or to the total amount of time spent in caring for patients. Doctors with larger lists may spend just as much or even more time, on average, in their consultations as those with smaller lists; but unless they also spend a greater amount of total time in contact with patients they will see relatively fewer of their patients during a specified time period. The previous sub-section was unable to find any conclusive evidence that list size and hours worked are positively associated, and it therefore seems that doctors with larger lists are coping either by spending less time, on average, with each patient or by reducing their consultation rate. The objective of this sub-section is to review the evidence in the literature about the average time spent in consultations and its relationship to list size. The literature contains rather better evidence about the average time of consultations than about the total number of hours worked, although the observation by Lees and Cooper (1963) that 'regrettably little work has been done on the use of a doctor's time' (page 431) remains almost as true now as when it was uttered.

From the reported studies it appears that the mean length of a surgery consultation is between about 5.0 and 6.5 minutes, and the mean length of a home consultation (excluding travel time) is between about 10 and 15 minutes. Reported times of surgery consultations, in order of increasing average time, include the following: 5.0 minutes (Fry, 1952; Floyd and Livesey, 1975); 5.1 minutes (Crombie and Cross, 1964); 5.2 minutes for return consultations (Morrell, 1971); 5.5 minutes in February (Eimerl and Pearson, 1966); 5.7 minutes (MacDonald and McLean, 1971); 5.8 minutes for new consultations (Morrell, 1971); 6.0 minutes (Hughes et al, 1976); 6.0 minutes in total time per patient and 5.0 minutes in face-to-face time (Buchan and Richardson, 1973); 6.4 minutes (Garraway, 1973); 6.6 minutes (College of General Practitioners, 1965); 6.7 minutes in August (Eimerl and Pearson, 1966); 7.2 minutes (Watts, 1952); 8.3 minutes (Wood, 1962); 8.3 minutes in 1958 (Mair and Mair, 1959); a mean of 8.7 minutes with a median of 6.0 minutes (Westcott, 1977); and 8.9 minutes in 1957 (Mair and Mair, 1959). Some of the extreme times can probably be discounted for various reasons as atypical of the 'average' general practitioner. Westcott's figures, for example, were based upon two weeks' self-recording as a vocational trainee, and the other high values all derived from studies carried out before the implementation of the Family Doctor Charter in 1966 and the subsequent expansion in the use of ancillary staff.

As with the total time spent in patient care, the individual variations around the average length of consultation are quite large. The notion of variation is used in two quite separate ways. First, the variation may be defined as the range between the longest and shortest individual consultations in the series under study. In the studies reported above, for example, the ranges in individual consultation times were 2-32 minutes (Westcott, 1977), 0.4 - 31.1 minutes (Buchan and Richardson, 1973) and 2-35 minutes (Wood, 1962). Second, in studies which report the average consultation time for a group of doctors, the variation may be defined as the range between the shortest and the longest average consultation time among all the doctors in the study. For example, Floyd and Livesey (1975) found a range in the average time per patient among five doctors from 4.0 to 6.0 minutes (with a mean of 5.0 minutes), and MacDonald and McLean (1971), in their study of eleven doctors, found a range of between 4.0 and 10.0 minutes (with a mean of 5.7 minutes). The importance of this distinction becomes apparent when the correlates of consultation time come to be considered.

The reported times of home consultations appear to be less reliable than those of surgery consultations because of inconsistencies in the use of travelling time. Reported times, in order of increasing average time, include the following: 7.7 minutes (MacDonald and McLean, 1971); 11 minutes (Watts, 1952); 'between ten and fifteen minutes' (Fry, 1952); 10.0 minutes (Hughes et al, 1976); 12.1 minutes in total time per patient and 5.6 minutes in face-to-face time (Buchan and Richardson, 1973); 12.3 minutes (Wood, 1962); 15.3 minutes (Eimerl and Pearson, 1966); 17.7 minutes (College of General Practitioners, 1965); 23.9 minutes (Garraway, 1973). The best indications of the probable range of individual home consultation times are found in Buchan and Richardson (1973), who reported face-to-face times in individual home consultations ranging from 0.2 to 46.0 minutes, and in Wood (1962) who reported a range of 2 to 60 minutes. The best evidence about the variations in average home consultation times between individual GPs is found in MacDonald and McLean's (1971) study of eleven Scottish doctors, among whom the range was from 6.3 to 10.9 minutes.

Three conclusions may be drawn from this mixture of data on consultation times. First, the average time of a surgery consultation is between about 5 and 6½ minutes. Reported averages in excess of about 6½ minutes may not be representative of contemporary general practice. Second, the average time of a home consultation is between about 10 and 15 minutes, although reported averages falling substantially outside this range cannot easily be discounted. Third, there are considerable variations, in the case of both surgery and home consultations, in the time taken for individual consultations and in the average

times taken by individual GPs. In the latter case, for example, there are reports of differences of up to 150% in the average time taken by GPs in surgery consultations, and of up to 75% in the average time taken in home consultations.

How, then, do these differences relate to list size? In addressing this question, the distinction must be preserved between the two different notions of variation discussed above. The reasons why individual doctors spend more time with some of their patients than with others have nothing to do with list size. The most important influence appears to be the diagnosis (Buchan and Richardson, 1973; Westcott, 1977; Morrell, 1971): mental, psychoneurotic and behaviour disorders tend to occupy the most amount of time, and infective diseases the least. The second type of variation, between the average consultation times of different doctors, is of much greater interest to the argument developed in this report, for it is possible that doctors may cope with larger lists by reducing the average amount of time they give to each patient. In fact there is no clear evidence that this occurs. Of the studies of individual practices cited above, insufficient information is given about list size to make meaningful comparisons, but two of the studies of groups of practitioners do enable consultation times to be related to list size.

First, Buchan and Richardson's (1973) study, which may be the most rigorous investigation yet conducted of how general practitioners use their time, included 1,636 surgery and 477 home consultations carried out by 22 doctors in and around Aberdeen. The timing of the different activities performed during the consultation was probably controlled more meticulously than in any other study. Buchan and Richardson concluded that 'the number of patients per doctor showed no correlation with face-to-face consulting time, but it should be noted that the average list size both in this study and in the North-East region of Scotland is not high. ... It is possible that in large practices the time spent per patient is reduced and that in very small practices the opposite is true. All we can say is that in this sample there was little evidence that list size is a major factor governing the length of time spent with patients' (para. 3.14). Buchan and Richardson also failed to find any association between the average face-to-face time and the number of patients seen per surgery session. 'In other words, these doctors worked at a consistent rate which appeared to be independent of the number of patients attending the surgery' (para. 3.18). This tendency on the part of GPs towards a rigid style and tempo of work irrespective of the pressure of demand has also been noted by Eimerl and Pearson (1966) and Hodgkin and Gillie (Royal College of General Practitioners, 1968), and it is, of course, consistent with the observation that the total time spent in patient care and the average time spent per

consultation are largely unrelated to list size. It is, however, not easy to reconcile with the findings of Mechanic's (1974) survey that doctors with larger lists were more likely to feel dissatisfied with the amount of time they could give to each patient. If doctors do indeed allocate what they consider to be the 'right' amount of time to each patient, irrespective of the numbers of patients on their lists, then frustrations might be expected to be directed more towards the limited number of patients to whom the 'right' amount of care could be given than towards any limitation in the time available to those patients who are seen and treated.

The second study that enables consultation times to be related to list size is that of MacDonald and McLean (1971). It is possible from their data to relate only the average list for each of the six practices with the average consultation times of the doctors in each practice, and this form of analysis may conceal important relationships. However, when handled in this way, the data confirm the conclusion of Buchan and Richardson that the two variables are unrelated. No systematic relationship was found between the rank order of the practices in terms of their average list sizes and their rank in terms of the average duration of either surgery or home consultations. Indeed, the two smallest practices actually contained the highest and the lowest mean surgery consultation times. Unlike Buchan and Richardson, however, MacDonald and McLean did find a strong association between the average time of surgery consultations and the weekly average number of patients seen in the surgery. As the authors put it, 'the more patients a doctor had to see in the surgery the less time he was apt to devote to each' (page 685). This finding did not conflict with the absence of any systematic relationship between list size and consultation time because in this particular study the number of patients seen each week bore no relationship to the average list of each practice. It does, however, lie uneasily against the impression created by other studies of a constant work routine on the part of doctors that is more or less unaffected by the pressure of demand.

Perhaps the most pertinent conclusion to be drawn from the material presented in this sub-section is that much still remains opaque about the extent to which doctors respond to larger list sizes by reducing the time they spend, on average, in consultation with each patient. The evidence, such as it is, suggests that list size is not usually an important determinant of the length of consultations, although extremely large or small lists might have some effect. There is conflicting evidence about the extent to which the length of a consultation is responsive to the pressure of demand: some writers have remarked upon the rigidity with which doctors construct their routines and

their imperviousness to the pressures of demand, but one study, at least, has demonstrated an association between weekly demand and consultation length that is not likely to be the result of accident or chance. In the context of the framework developed in this report, it seems unlikely that doctors with small lists are spending more time on each consultation than those with larger lists (although it is arguable as a separate issue that they should be spending more time in order to give a better standard of care). However, if it is the case that GPs usually cope with larger lists neither by increasing the total hours spent in the practice nor by reducing the average time spent with each patient, it follows that they must have a lower consultation rate. It is to this, then, that we next turn.

CONSULTATIONS IN GENERAL PRACTICE

Much more information has been gathered about the number of consultations that take place in general practice than about their length, but precisely because there is so much information in the literature about consultation rates, it must be treated with considerable care. The objective of this section is to review the evidence in the literature about the realtionship between list size and consultation rates to assess the extent to which doctors may be coping with larger lists by having fewer consultations.

Three principal measures of consultations are to be found in the research literature. The most common measure is usually called the consultation rate, and is derived by expressing the total number of consultations made by a doctor in the course of a year as a rate per 1,000 patients at risk of consulting. Exactly the same information is conveyed by shifting the decimal point three places to the left to show the average number of consultations made each year by each patient at risk. Most studies of consultations have incorporated this measure, and many have derived separate consultation rates for men and women, surgery and home consultations, and so on.

The second measure is usually called the patient consulting rate, and is the proportion of patients registered in the practice who consult on different numbers of occasions during the course of a year. The patient consulting rate cannot be inferred from the consultation rate, but the consultation rate can be calculated from data about the patient consulting rate. Of the two rates, the patient consulting rate requires the more elaborate data-collection system, and it may be for this reason that relatively few studies of the pattern of consultations in general practice have incorporated it.

A third measure of consultation concerns the place of consultation. Most (though not all) studies have made the basic distinction between consultations that take place in the surgery and in the patient's home, and some have further distinguished telephone consultations and consultations or contacts with other members of the primary care team. Information about the place of consultation is commonly expressed either as the proportion of consultations that occurs at home, or as a classifying variable for consultation and patient consulting rates. It is used in the latter way in the analysis that follows.

In addition to these three principal measures of the pattern of consultations, a small number of studies have collected information on the average number of consultations per patient consulting during the year, the average number of consultations per episode of care, and the ratio of doctor-initiated

to patient-initiated consultations. Such measures offer potentially important clues to the consequences of variations in list size, but there is little evidence to relate these measures of consultation to list size.

Methodological problems

Before reviewing the evidence itself, attention must be drawn to problems inherent in the data and in any attempts to compare the results of different studies. Many of these problems are common also to the data summarised in the preceding section. First, reservations must be lodged about the quality and the comparability of data on consultations. Most of the studies cited below are the work of self-selected general practitioners looking at their own practices, and it may reasonably be assumed that they differ in the assiduity with which they record their work as well as in the consistency with which they define the basic measurements. As Howie (1977) has noted, 'rates depend on definitions and are influenced by the accuracy of recording visits and appointments; the accuracy of the estimate of practice size; the inclusion or exclusion of figures for immunizations, child care, ante-natal care and other special sessions; the quantity and type of work done by the nurse or health visitor; and on the number of services to temporary residents and other local factors' (page 23). Some studies make clear distinctions between different types of consultations, but many do not (Knox and Morrell, 1974). Second, most studies have lasted for less than one year's duration, and annual consultation rates have therefore had to be constructed from much shorter recording periods. Allowance is rarely made for possible seasonal atypicalities in the recording periods, and there is no standard definition of the number of weeks in a working year. Third, different kinds of studies present different kinds of data. At one extreme are the studies of individual practitioners, in which list size and consultation rates can be related with a good deal of precision; at the other extreme are studies typified by the National Morbidity Survey, in which average list sizes and average consultation rates can be related only at such high levels of generality as the standard regions, thereby concealing a good deal of the 'real' relationship between them. In between these extremes are studies of practices which yield average measurements for each doctor in them, and studies of groups of doctors which show, for example, the average consultation rates for all doctors with list sizes within a specified range. Fourth, the accuracy of the data on list size may be variable, especially in practices where no up-to-date register exists or where there is a large turnover of patients in the course of a year. Fifth, an association between list size and consultation rates, even if it is found to exist, may be difficult to interpret. Many factors in addition to list size may influence the pattern of consultations, and there is no guarantee, in contrasting

the consultation rates of doctors with different list sizes, that like is being compared with like - or even similar with similar. Most of the study reports include at least a brief description of the salient features of the practice under investigation, but in reality it is almost impossible to use such information as classifying variables.

Consultation rates

These problesm can to some extent be eased by looking at the evidance in different ways. We do this for each of the principal measures of consultation described above, beginning with the consultation rate. The first approach is simply to accept all the evidence at face value and see whether, across a large number of observations, there are any associative trends between list size and consultation rates. If a sufficiently large number of observations is available, any underlying association might be expected to show through, with other factors being more or less randomly distributed. Twenty-four studies have been identified in the literature in which the annual surgery and home consultation rates of single-handed or partnership doctors can be related to list size. The studies yield a total of 80 separate observations, the greater number of observations than of studies being explained by the fact that several studies report a series of consultation rates over several years. The 24 studies cover a a wide span of dates (1959 to 1976), of locations (Northern Scotland to South-West England), of list sizes (1,473 to 5,411), of annual surgery consultation rates (1.3 to 7.1) and of annual home consultation rates (0.1 to 3.6). The 24 studies are: Baldwin (1959), Barber (1971), Bolden and Morgan (1975), Cobb and Baldwin (1976), Dawes and Cottrell (1964), Duncan and Orcharton (1964), Fry (1972), Garraway (1973), Grene and Johnson (1971), Hardman (1965), Hughes et al (1976), MacDonald and McLean (1971), McGregor (1973), Mair and Mair (1959), Marsh and McNay (1974a), Morrell, Gage and Robinson (1970), Noble (1973), Posner (1965), Royal College of General Practitioners (1968), Scott and McVie (1962), Steen (1967), Stevenson (1964), Weston Smith and O'Donovan (1970) and Wood (1962).

The 80 observations relating list sizes to annual surgery and home consultation rates are set out in the scatter diagrams (figures 1 and 2). The mean list size is 2,749 (S.D. 923), the mean surgery consultation rate is 3.2 (S.D. 0.9), and the mean home consultation rate 1.3 (S.D. 1.0). The much higher coefficient of variation among the home than among the surgery consultations is consistent with the observations of several investigators about the large differences between GPs in the proportions of home visits they make (Lees and Cooper, 1963; Williams, 1970; Marsh, McNay and Whewell, 1972; Cobb and Baldwin, 1976). Indeed, the literature contains reports of variations in home visits as a

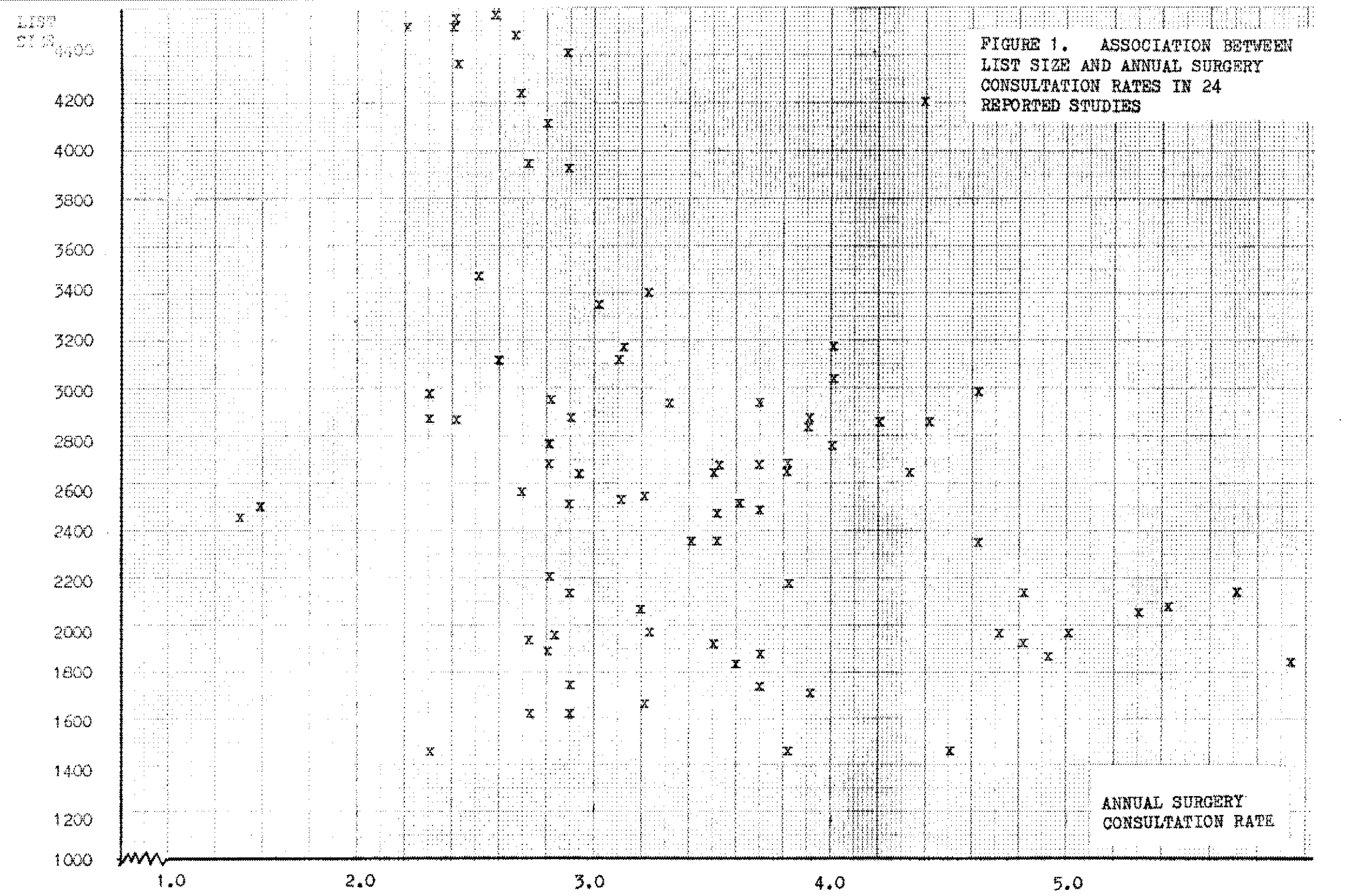

FIGURE 1. ASSOCIATION BETWEEN LIST SIZE AND ANNUAL SURGERY CONSULTATION RATES IN 24 REPORTED STUDIES

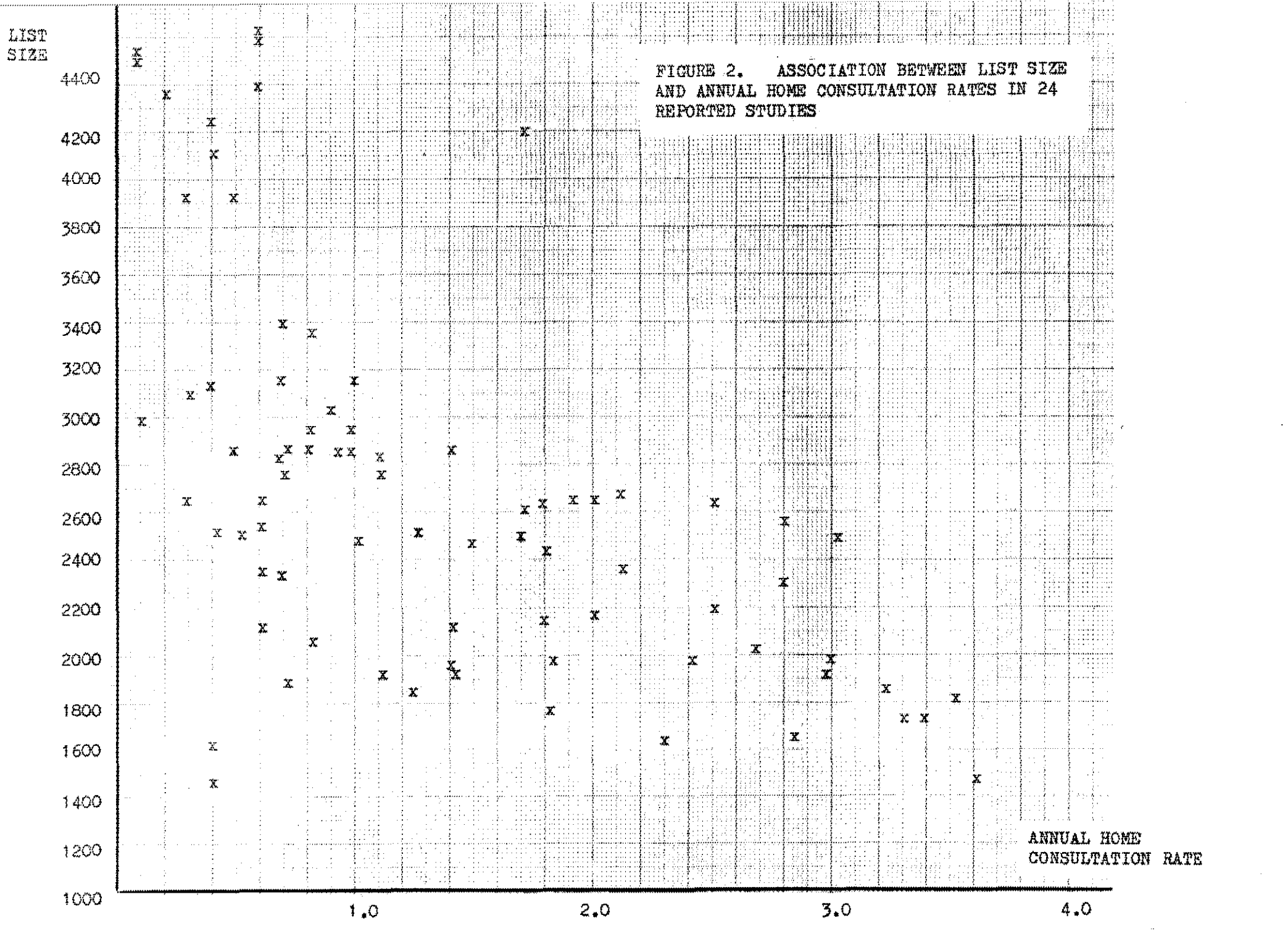

FIGURE 2. ASSOCIATION BETWEEN LIST SIZE AND ANNUAL HOME CONSULTATION RATES IN 24 REPORTED STUDIES

proportion of total consultations ranging from as low as 4 per cent (Weston Smith and O'Donovan, 1970) to as high as 60 per cent (Mair and Mair, 1959; McGregor, 1973).

Taking first the surgery consultation rate, simple observation of the data indicates a negative relationship between list size and consultation rates: with one exception, for example, all the practices with list sizes above 3,000 had surgery consultation rates of 3.0 or less, whilst the majority of practices with list sizes below 2,500 had consultation rates of 3.0 or more. Regression analysi s confirmed the existence of a negative correlation ($r = -0.35$), but the association is fairly weak, with only 12 per cent of variance among surgery consultation rates being explained by list size. It seems reasonable to conclude that, accepting the evidence at face value, there is a clear tendency for doctors with larger lists to have lower annual consultation rates than those with smaller lists, although doctors with medium-sized lists (between about 2,500 and 3,000) display a range of rates that bears little obvious relationship to the numbers of patients on their lists.

Considerable caution is needed in accepting the evidence at face value, however. For example, most of the observations in which list size exceeded 3,500 were drawn from different years in one practice (Fry), and the possibility exists that an individualistic style of practice might exert a disproportionate effect upon the overall relationship. On the other hand, it is interesting that at least some of the observations in which the consultation rate is substantially above or below the usual rate for that list size can be explained in terms of unusual features in the practices concerned. McGregor (1973), over a twenty-year period, reported a very low surgery consultation rate indeed (1.3), but this occurred in a rural area, with a widely scattered population, a large proportion of hospital work, and a compensatingly high proportion of home visits (59 per cent over the twenty years). Mair and Mair (1959) likewise reported some low surgery consultation rates for the number of patients in the practice, but as with McGregor, these appeared to be offset to a large extent by an uncommonly large proportion of home visits. The combined surgery and home consultation rates reported in each of five years by Mair and Mair were not significantly out of line with those from other practices with similar list sizes. The single most 'deviant' observation, in which a list size of 4,200 was related to a surgery consultation rate of 4.2, derived from the Practice Nurse Study (Royal College of General Practitioners, 1968). There is no obvious reported explanation for this, but it is believed that the doctor in question had at least the part-time services of an assistant, and in any case he had a working week of 56¾ hours - well above the average for general practitioners (see page 67).

This finding is consistent with the general conclusion emerging from this review of the literature, that doctors tend to work for a similar amount of time whatever the size of their lists, and more patients therefore tend to be accommodated by reducing the consultation rate. If, however, (as in this case), a doctor works for substantially more hours than average, he can sustain a high consultation rate among an unusually large number of patients.

Turning now to the home consultation rate, simple observation of the scatter diagram (figure 2) again confirms a negative association between this variable and list size. With few exceptions, practices with list sizes above 2,800 had home consultation rates of 1.0 or less, whilst most of those with lists below 2,500 had consultation rates in excess of 1.0. The negative association is confirmed by regression analysis, and is stronger than the negative association between list size and surgery consultation rates ($r = -0.54$). This suggests that although there is a tendency for doctors to cope with larger lists by reducing the average number of surgery consultations per patient each year, there is an even stronger tendency for them to cope by cutting back the average number of home visits they make to their patients each year.

As with the data on surgery consultations, much caution must be exercised in accepting the data on home visits at their face value. Almost all the observations in which the list size exceeded 3,500 were drawn from Fry's practice, and there is the possibility that an idiosyncratic style of practice may distort the perception of a general relationship between list size and home visiting. Indeed, some of Fry's home visiting rates are among the lowest reported anywhere in the literature, irrespective of list size. Moreover, it is apparent in examining Fry's data that a regular decline in visiting rates has occurred since the mid-1950s, and hence the low proportion of visits may reflect a more widespread trend in general practice during the last 20 years as much as the increasing list size in this particular practice. Against this, it may be observed that the association between list size and home consultation rates appear to depend rather less heavily upon Fry's data than does the association with surgery consultation rates, for even among practices with average lists below 3,000 there remains a clearly discernible tendency for home consultation rates to rise as list size falls. Finally, as with the data on surgery consultations, special circumstances may account for some of the practices in which home consultation rates were particularly out of line with the average for any given list size. For example, Barber (1971) and Morrell et al (1970) both reported lower home consultation rates than might be expected from the average list sizes of the doctors concerned, but in both cases the doctors had substantial commitments outside the practice (in a hospital and medical school respectively).

A second way of approaching the evidence on list size and consultation rates is through the time trends within the same practices. The literature identifies several practices in which consultation records have been maintained over a period of time (Royal College of General Practitioners, 1973), but only a small number of these contain the information necessary to relate changes in consultation patterns to changes in list size. Where this can be done, however, the approach has the considerable advantage of controlling for some of the extraneous factors which might influence inter-practice variations in rates, although it must be noted that in some cases the records have been kept over a period of time precisely in order to evaluate a change or innovation in the practice. In such cases it is often impossible to tell how much of the change in consultation patterns has resulted from the innovation and how much from any change in list size.

From the 24 studies listed above, six present surgery and home consultation data for two years or more: Dawes and Cottrell (1964), Fry (1972), Hughes et al (1976), Mair and Mair (1959), Steen (1967) and Stevenson (1964). In addition, two studies (Hodgkin, 1973; Craig, 1974) present data on trends in home and surgery consultations combined, and Marsh (1968) has published information on changes in the number of home visits, related to list size, over an eight-year period. Most of these sources are at least consistent with the hypothesis that list size and consultation rates are inversely associated, and some of them support the more refined hypothesis that home consultation rates are more sensitive to changes in list size than are surgery consultation rates.

Stevenson's (1964) data from a three-man practice in Ayrshire showed that whilst the practice average list size increased regularly between 1957 and 1963 from 1,719 to 2,024, the surgery consultation rate fell regularly from 3.7 to 3.0, and the home consultation rate fell from 3.4 to 2.7. Expressed in percentage terms, the list size increased by 17.7 per cent whilst the surgery consultation rate over the same period decreased by 18.9 per cent and the home consultation rate by 20.6 per cent. Consequently the average number of consultations given by each doctor remained almost constant from one year to the next, differing by only 5 per cent between the beginning and the end of the period. A similar story is reported by Hodgkin (1973), although his data do not distinguish between surgery and home consultations. The routine collection of workload data in Hodgkin's practice revealed that, whilst the practice average list size increased from 2,012 in 1960 to 2,934 in 1969 (an increase of 45.8 per cent), the consultation rate over the same period declined from 5.1 to 3.5 (a decrease of 31.4 per cent). This period coincided with the introduction of a practice nurse and of new administrative routines (see (see page 65), both of which may have had an effect upon the rate of

consultations; but Hodgkin noted that the average number of consultations given by each doctor in the practice remained virtually constant, regardless of other changes. To the extent that Stevenson's and Hodgkin's findings are representative, they reinforce the view that doctors tend to follow personal routines of practice which are more or less impervious to outside influences and which account, of course, for the inverse relationship between list size and consultation rates.

The data recorded by Dawes and Cottrell (1964) are interesting because of their reference to a period (1958 to 1963) in which the average list size in the practice remained absolutely constant. The hypothesis of an inverse relationship between list size and consultation rates would suggest that, in the absence of any change in list size over time, there should be no marked changes in the consultation rates; and this is more or less what Dawes and Cottrell found. The total consultation rate varied in a non-systematic way between 5.2 and 5.6; the surgery consultation rate between 3.3 and 3.6; and the home consultation rate between 1.7 and 2.1. These figures are the highest and lowest rates in any of the six years, and they reveal variations in consultation rates that are not only of a non-linear kind, but are also considerably smaller than those noted over a similar time-period by Stevenson and by Hodgkin.

Fry's (1972) data cover a much longer span than any other, and for the reasons noted above they must be treated cautiously. The relationship between list size and consultation rates is less apparent in Fry's data than in Stevenson's or Hodgkin's, but it is discernible. In the case of surgery consultations, the highest rates were generally recorded in the years when the practice average list size was lowest, and conversely the low consultation rates were usually associated with high average list sizes in the practice. A similar tendency is to be found in the data on home consultations, although the point has already been made that the steady decline in home visiting is almost certainly the result of a conscious change in the style of practice as much as the response to an increasing list size. Fry reported a number of changes occurring in the practice over the 21 years which might have affected the rate of consultations, but the durability of the association between list size and consultation rates is further evidence of its validity.

Two smaller studies are not inconsistent with the hypothesis. Steen (1967) recorded surgery and home consultation rates in a group practice in 1963, 1965 and 1966 when the practice average list size increased from 2,500 to 3,100, whilst the surgery consultation rate fell from 2.7 to 2.4 and the home consultation rate fell from 0.5 to 0.3. Craig (1974) reported an increase in the overall consultation rate in a West Midland's practice from 2.2 to 2.3

between 1971 and 1973 when the list size fell from 3,950 to 3,460. In the latter year, however, more surgery hours were provided, and this (rather than the decline in list size) may have accounted for the rise in the consultation rate.

Two further studies show a contrary trend to the general pattern. The study of a group practice in Derbyshire (Hughes, et al, 1976), showed a constant surgery consultation rate and only a slightly changing home consultation rate over a four-year period (1971-4) during which the practice average list size rose from 3,020 to 3,183, fell to 2,783 and then rose again to 2,841. It would be expected that, with little significant change in the consultation rates, the increases in the practice average list size would produce a concomitant increase in the total number of hours spent in contact with patients, and conversely that decreases in the list size would result in a decrease in hours worked; and this was found to be the case (see page 68). Thus, although the doctors in this practice appeared to be unusual in not responding to a growth in the number of patients by reducing the consultation rate, they did compensate by increasing the total time they spent in contact with their patients. The other study, by Mair and Mair (1959), is less readily explained. During the period 1954-58, when the practice average list size increased regularly from 1,473 to 2,115, the surgery consultation rate also increased steadily from 2.1 to 4.6. It is true that the home consultation rate fell in linear fashion from 3.6 to 1.4 during the same period, but Mair and Mair offered three explanations for this, none of which made any reference to the expanding number of patients on the list. There is no obvious way of accounting for this apparently deviant set of data, although it may be noted that even the largest reported list size (2,115 in 1958) was quite small by general standards, and the overall consultation rate associated with it (6.0) was by no means unusual for lists of about that size.

A third way of approaching the relationship between list size and consultation rates is through the conclusions drawn by other writers who have examined it. Cartwright (1967) found from her national study of 422 GPs that the estimated overall annual consultation rate (based upon reported consultations in a two-week period) declined from 5.6 among doctors with lists below 1,500 to 4.2 among those with lists above 3,000. The rates for doctors with intermediate list sizes were not linearly distributed, but Cartwright took the relationship between list size and consultation patterns seriously by discussing possible explanations. 'Such a trend might arise because patients of doctors with large lists are discouraged from consulting their doctor when he is busy and has many other patients waiting, because doctors with small lists encourage their patients to come back to see them rather more frequently, because

chronically ill patients seek out doctors with small lists, or because doctors whose patients do not consult frequently on the average are able to take on a larger number of patients' (pages 163-4). After reviewing the evidence to hand, Cartwright concluded that the latter hypothesis was the most plausible, namely that doctors whose patients consult relatively infrequently are thereby enabled to take on larger numbers. We return to this important conclusion later.

Wright's (1968) account of a survey of 77 GPs in the South-West Faculty of the Royal College of General Practitioners included information on the association between list size and both surgery and home consultations. The study involved detailed workload recording for four specified weeks each quarter in 1964-5, and a total of some 51,000 consultations were included. A direct consultation was defined as 'a conversation with medical content, conducted directly between patient and doctor'; an indirect consultation was 'one involving the intervention of telephone, message, or third party' (page 5). The results showed that among GPs with fewer than 2,000 patients the overall mean direct consultation rate was 4.8; among those with 2,000-2,999 patients it was 4.0; and among doctors with 3,000 or more patients the rate fell to 3.4. Wright commented that 'consultation rates ... appear significantly higher in doctors with small lists (less than 2,000), suggesting that work may expand to fill the time available for it ... Conversely, those doctors with large lists (over 3,000) show the lowest mean consultation rate, suggesting that workload does not increase proportionately to list size'(page 7). With regard to home visiting, Wright found that the proportion of home visits declined very slightly from an average of 31% among GPs with fewer than 2,000 patients, to 28% among those with 2,000-2,999 patients, and to 27% among those with 3,000 or more patients. Wright concluded that 'the mean home visiting ratios for doctors with small lists and those with large lists does (sic) not show any major difference. However, although there were only 18 doctors who had small lists (less than 2,000), they provided seven of the 12 doctors with high visiting ratios. There is some evidence therefore, that doctors with small lists not only have higher consultation rates, but also higher home visiting ratios, than do doctors with large lists' (page 9).

Richardson et al (1973), in their study of workload recording by 142 GPs in the North-East Scotland Faculty of the College, found like Wright that a relatively small average list size was associated with a relatively high consultation rate, although the source of this finding (Table 8, page 140) is confusing, to say the least. It appears that, among the 23 GPs with average consultation rates in excess of 18.6 per 1,000 per day, the mean list size was 1,380, whilst among the 26 GPs with daily consultation rates below 11.5,

the mean list size was 1,895. Data are also presented on the mean percentage of home visits carried out by doctors with lists above and below 2,000. With the location of the practice (city/country) and the age of the doctor (under 45/over 45) controlled, the results indicated a somewhat lower visiting rate among GPs with the larger than with the smaller lists. Richardson et al concluded, however, that 'analysis of variance showed a slight but not significant association (of home visiting) with list size' (page 138).

It seems reasonable to conclude from these different ways of looking at list size and consultation rates that the two variables are inversely associated, and that it is typical for general practitioners to cope with larger lists by having lower surgery and (especially) home consultation rates. This conclusion is consistent with the evidence presented in the previous section on the time spent in general practice and the average time of consultations. If, as appears to be the case from the fragmentary evidence available, GPs with larger lists do not usually work for longer hours or spend less time with patients than those with smaller lists, it follows that they must have fewer consultations for each thousand patients on their lists. It is gratifying that the research evidence supports this, and refines it by showing that the reduction in consultation rates is probably more marked among home than among surgery consultations.

However, the picture is not yet complete, for a reduction in the consultation rate may affect the practice population in different ways. It may, for example, mean that proportionately fewer patients are seeing the doctor at all in the course of a year, or it may mean that the same proportion of patients are seeing the doctor on fewer occasions each. The effect, in terms of the pattern of consultation, may be very different in each case and may lead to different evaluations of the consequences of rationing. We turn next, therefore, to the evidence on patient consulting rates - that is, the proportion of registered patients consulting their doctor on different numbers of occasions during the year.

Patient consulting rates

Whereas it is relatively simple for GPs to keep records of their consultation rates, it is rather more difficult for them to keep track of their patient consulting rates, and it is presumably for this reason that few such studies appear in the literature.

The best data are probably those collected by Cartwright (1967) in her

national study of 422 GPs and their patients. The information on patient consulting rates is drawn from people's own accounts of the number of times they consulted their doctor in the year preceding the interview, and for this reason it is open to possible errors of recall. Taking the replies at face value, however, they show a modest tendency for patient consulting rates to decrease as their doctors' lists increase in size. The proportion of patients reporting at least one consultation with their doctor in the preceding 12 months was 71% of those on lists of less than 1,500; 69% of those on lists of 2,000 - 2,499; and 65% of those on lists of 3,000 or more. Conversely, the proportions of patients reporting five or more consultations in the previous year were 31%, 29% and 22% respectively among those on lists of less than 1,500, 2,000 - 2,499 and 3,000 or more. The differences are fairly small, and the rates for doctors with intermediate list sizes were not linearly distributed; but they offer some indication that at least part of the lower consultation rates experienced by doctors with larger lists may be the result of a lower proportion of patients seeing their doctor in the course of a year.

Other studies of single practices are consistent with this conclusion. Marsh and McNay (1974a; 1974b) reported that only 53 per cent of male patients and 63 per cent of female patients consulted during the course of a year in which the GP's list size exceeded 3,100, whilst at the other extreme Morrell (1971), Scott and McVie (1962) and Barber (1971) reported patient consulting rates of 78 per cent, 73 per cent and 83 per cent respectively with practice average list sizes of 1,485, 1,993 and 1,612. The National Morbidity Surveys (General Register Office, 1958; Office of Population Censuses and Surveys, 1974) reported an identical patient consulting rate in each of the two years (67 per cent), but the data are presented in a way that precludes detailed analysis of the relationship between this variable and list size. The best that can be done is to relate the average list size in each region to the regional patient consulting rate: the result is not inconsistent with a negative association between list size and patient consulting rates, but not surprisingly the regional differences on each variable are quite small.

The initiation of consultations

Another effect of the reduced consultation rate consequent upon an increase in list size may be a shift in the ratio of doctor-initiated to patient-initiated consultations. The evidence reviewed so far in this section points towards conflicting expectations in this matter. On the one hand, the reduction in overall consultation rates with increasing list sizes (and particularly the reduction in home consultations) suggests that doctors with larger lists may be initiating fewer follow-up consultations for each episode of care than those

with smaller lists. If this is so, it would be reflected in a diminishing ratio of doctor- to patient-initiated consultations as list sizes increase (Last, 1965). On the other hand, the fact that a lower proportion of patients appear to see their doctor at all in the course of a year in larger than in smaller practices suggests that part of the difference in overall consultation rates is explained by a reduced patient-initiated demand, for whether or not a patient sees his doctor at all in the course of a year is largely up to him. To this extent, there is no necessary reduction in the ratio of doctor-initiated to patient-initiated consultations as list sizes grow.

The published evidence on this matter is extremely sketchy, and probably suffers from a lack of direct comparability between different investigations in the definition of terms. It also fails to point to any clear-cut conclusions. Cartwright (1967) concluded from her survey of 1,397 people that no association existed between list size and the proportion of consultations initiated by doctors themselves. 'The proportion of consultations said by the patients to be for the first time for that episode, or at the suggestion of the doctor, did not vary with the number of patients the doctor looked after' (page 165). Wright's (1968) survey of the workload of 77 GPs in the South-West of England in 1964-5 reached a similar conclusion. The proportion of follow-up (doctor-initiated) to total consultations was 59% among doctors with fewer than 2,000 patients, 58 per cent among those with 2,000-2,999 patients, and 56% among those with 3,000 or more patients. These differences are consistent with the hypothesis that the proportion of doctor-initiated consultations would diminish with increasing list size, but they are plainly insignificant. Wright concluded that, although doctors with smaller lists had significantly higher consultation rates than those with larger lists (see page 82),'this expansion does not result from an increased proportion of follow-up consultations; that is, it is not a load imposed directly by the doctor upon himself' (page 7). Likewise, the lower consultation rates of GPs with large lists (over 3,000) could not be explained by a reduced follow-up load.

Against the evidence of Cartwright and Wright, Richardson et al (1973) offer circumstantial evidence of a negative association between list size and the proportion of doctor-initiated consultations. This study of the workload of 142 GPs in North-East Scotland found a wide variation between the doctors in their follow-up policies, but noted nevertheless that doctors with high overall consultation rates also had high proportions of return visits. Since the study also found that high overall consultation rates were associated with small average list sizes, it is possible that doctors with smaller lists made more use of follow-up consultations than those with larger lists. However, no direct evidence is given on this.

The small amount of evidence that is available from other studies offers no further clarification. The literature contains reports of variations in the ratio of doctor-initiated to patient-initiated consultations ranging from 0.6:1 (Marsh and Kaim-Caudle, 1976) to 1.1:1 (Rider, et al, 1969), but there does not appear to be any associative pattern between these variations and the variations in the list sizes of the practitioners concerned.

Summary

It is now possible to summarise this section on consultation patterns. The firmest conclusion is that, notwithstanding the relatively large amount of published research evidence, there remains much that is obscure about the way in which consultation patterns vary in practices of different size. Most (though not quite all) of the reported studies are of self-selected practitioners, and differences in the definitions and methodologies used, together with the impossibility of controlling for other factors that might affect the pattern of consultation, render hazardous any attempt to relate list size and consultation rates. A second conclusion is that, ignoring such hazards and accepting the reported evidence at face value, there appears to be a broad negative association between list size and both surgery and (especially) home consultation rates. Doctors with larger lists generally have lower surgery and home consultation rates than those with smaller lists. There is fairly clear evidence that, as list sizes increase, home consultations are cut back more markedly than surgery consultation rates. This second conclusion would be predicted from the findings in the previous section that GPs with larger lists do not appear to spend either more total time on patient care or less average time with each patient than do GPs with smaller lists. A third conclusion is that part of the lower consultation rates that occur among doctors with larger lists is probably caused by a lower patient consulting rate. It appears that relatively fewer patients consult their doctors at all in the course of a year in larger than in smaller practices, and correspondingly more consult on at least one occasion. A very crude calculation based on data reported by Cartwright (1967, Table 54) indicates that about a quarter of the difference between the consultation rates of doctors with fewer than 1,500 patients and those with 3,000 or more patients was due to the lower patient consulting rates among the latter. The remainder of the difference was due to the smaller average number of consultations per patient consulting. A fourth conclusion is that no substantial evidence exists that the control exerted by GPs over their own workload through the use of doctor-initiated follow-up consultations is related to the number of patients for whom they care. It might be expected that doctors with larger lists would achieve a lower consultation rate by

holding down the element of demand over which they have some control; but there is no clear evidence in the literature that they are more likely to do this than are doctors with smaller lists.

If this is a correct interpretation of the evidence, it leads to the rather startling conclusion that the lower levels of demand that are evident in larger than in smaller practices reflect the behaviour of the patients rather than the doctor. It is interesting that a similar conclusion was reached by Cartwright (1967) and Wright (1968). Cartwright, after reviewing a number of possible explanations for the association between list size and consultation rates, concluded that the most plausible explanation is either that 'doctors whose patients consult relatively infrequently take on larger numbers, or that those whose patients consult frequently recruit more doctors into their partnership' (page 165). Wright noted with respect to his data that 'the differences in consultation rate between doctors with small lists and those with large lists thus appear to be dependent on some subtle difference in patient-doctor relationship. The patient creates more contact with the small-list and less with the large-list, doctor. We can only speculate on the mechanism of this difference' (page 7, emphasis added). It may, of course, be the case that the lower apparent level of patient-initiated demand in larger practices is strongly influenced by the expectations of the doctor, mediated to his patients ov r a period of time. We know of no investigation into this delicate aspect of practice.

THE CONTENT OF CARE IN GENERAL PRACTICE

The two preceding sections presented and evaluated the evidence in the research literature pertaining to the first five questions listed on page 63. The data indicated ways in which time and consultation patterns in general practice may typically be related to variations in list size. However, by concentrating on the fact rather than the content of interaction between doctor and patient, the salience of the analysis is weakened. A consultation is not an end in itself, it is a means to the more distant end of enhancing the patient's welfare; and any analysis of the consequences of differential doctor/patient ratios must therefore take account of what the doctor does, and the effects of what he does, as well as whom he sees. The aim in this section is therefore to review the evidence about the sixth question posed on page 63, namely the ways in which the content of care in general practice varies with list size.

Methodological problems

In setting out the available evidence, care is required in selecting the base upon which to calculate rates. The evidence reviewed above suggests that variations in practice size are associated more intimately with the rate at which patients are seen than with either the number of patients seen or the average amount of time allocated to them. Doctors with larger lists appear, typically, to cope with them not by spending more time in patient care or in conducting shorter consultations, but by seeing a smaller proportion of their patients and seeing them less often. If this is a correct interpretation of the evidence, it is possible that the content of practice may show less variation between practices of different sizes when expressed as rates per thousand consultations than when presented as rates per thousand registered patients. If, in other words, the most important effects of variations in practice size are to be found in the pattern of care to the practice population as a whole rather than to those members of the population who happen to cross the doctor's threshold, it seems important to relate the things that GPs do to the total number of patients for whom they are responsible as well as to the number of patients whom they happen to treat in the course of a specified period of time.

This consideration points to one of the difficulties in using published research material for this purpose, that not all studies of the content of care yield data that can be related both to consultations and to the population at risk. There are other difficulties, some of which have been reviewed above in connection with the material on consultation rates (see page 75). For example, there is the obvious fact of the paucity and selectivity of data.

Most studies have been located in self-selected or volunteer practices, and doctors observing their own practices have found it rather easier to count numbers of patients (and even the amount of time they spend with them) than to list the transactions that take place between doctor and patient. Items which are noted routinely for administrative reasons are most likely to be known, but although prototype systems exist for recording other items of care given by GPs (Eimerl and Laidlaw, 1969), it is instructive that the periodical reviews of trends in general practice published by the College have contained much less information under the heading 'techniques used' than the headings 'consultation rates' or 'time spent'. Second, a large part of the information that is available about what GPs actually do cannot be related to list size. Information is collated by the DHSS about the work of GPs in prescribing, in certifying injury and sickness benefit claims, in performing cervical cytology examinations, in requesting pathological and radiological investigations, in referring patients for specialist care, and so on; but none of this can be sub-divided by the practice size of the doctors concerned. Similarly, the Second National Morbidity Study (OPCS, 1974) collected quite detailed information on the rates of different types of referrals (Table 19), but the data can be related to list size only through the indirect technique of contrasting referral rates for regions with different mean list sizes. A third difficulty concerns the inconsistent use of definitions. This is well illustrated by Carstairs and Skrimshire's (1968) review of published studies of outpatient referrals, which showed considerable inconsistency in distinguishing between persons and referrals. In other cases it is not always clear whether referrals for radiological investigation are classified as 'outpatient referrals' or as 'diagnostic investigations'.

Referrals for specialist care or opinion

In an attempt to impose some sort of order on the published material, this section proceeds by reviewing the evidence about referrals for specialist care or opinion, requests for investigative tests, and treatments used by GPs themselves. Looking first at referrals for specialist care or opinion, the second National Morbidity Study found that, in England and Wales as a whole, the annual referral rates per 1,000 population were 18.3 to inpatient care (direct admission to hospital) and 86.0 to outpatient care (including domiciliary consultations). Regional variations were marked, ranging between 14.5 and 38.9 for inpatient referrals and between 67.2 and 160.7 for outpatient referrals. The research literature contains a considerable number of studies of the correlates of GPs' referral patterns, but few have explicitly examined the effect of practice size. Those that have are unanimous in concluding that list size and the rate of referral per thousand patients at risk are not associated.

In their study of the use of hospitals in Barrow in 1957-8, Forsyth and Logan (1960) noted the 'striking absence of any apparent relationship' between size of list and the rate of in-patient referrals among the 16 practices in the area; and in their later survey of the use of the out-patient departments of some 80 hospitals serving about 2 million people, Forsyth and Logan (1968) reported that 'size of practice list had no effect at all. Practitioners with comparable practice sizes had widely differing rates of referral to out-patients... Some with less than 1,500 patients to look after had low rates while others with lists exceeding 3000 were found among the highest users of the facilities' (page 41). Scott and Gilmore (1966), in introducing their study of the use of out-patient departments of Edinburgh hospitals, observed that 'studies have shown that general practitioners vary in the number of patients they refer in relation to the number of patient consultations they have or to the number of patients at risk' (page 5), but the references that are cited in support of this assertion do not demonstrate any systematic relationship between the variables. Moreover, Scott and Gilmore's own study of the out-patient referral of Edinburgh residents by 30 GPs in ten practices in May and June 1962 led them to refute that conclusion. 'We found a wide range in referral rates from individual doctors, ranging from 0.6 to 25.8 per 100 patients at risk. However, we were unable to establish any correlation between referral rates and ... size and type of practice ...' (page 12). Wright's (1968) survey of 77 GPs in the South-West England Faculty of the Royal College of General Practitioners yielded out-patient referral rates of 31 and 39 per thousand consultations among, respectively, 18 doctors with fewer than 2,000 patients and 21 doctors with more than 3,000 patients; and rates of 148 and 135 referrals per thousand patients at risk among the two groups of doctors. (These figures included NHS and private referrals, but excluded domiciliary consultations.) Wright did not attach significance to these differences, commenting that 'doctors with smaller lists tend to ... seek consultant advice as often as their colleagues with larger lists' (page 25). Data collected by the Birmingham Research Unit of the Royal College of General Practitioners, derived from the first 100 proformas received from volunteer practitioners, and referring to a total of 65,000 consultations carried out in four weeks in October and November 1977, showed that the referral rate per thousand consultations to hospital OP and IP departments was significantly greater among doctors with fewer than 500 consultations in the four-week recording period than among those with more than 500 consultations (Journal of the Royal College of General Practitioners, 1978b). However, no information was sought about list size, and as the report pointed out, 'the doctors participated voluntarily, and cannot be considered as a representative sample' (page 521).

Of the area studies, the data collected in the Second National Morbidity Study (OPCS, 1974) showed a negative rank correlation at the regional level between list size and out-patient referrals: that is, regions with the larger regional average list sizes tended to display the lower regional average referral rates, and vice versa. The explanation for this finding may lie in the relationship between the supply of and the demand for resources. Regions with a relatively low provision of GP manpower tend to have relatively low provisions of hospital services (Cooper and Culyer, 1967), and as Cooper (1975) has noted, patients cannot be referred to hospital specialists who do not exist. At a lower level of territorial aggregation, a statistical analysis of GP list sizes and hospital in-patient discharge rates among the districts of the South-East Thames Region concluded that there was no evidence of any relationship between them (SETRHA, 1974).

The general conclusion emerging from these studies of groups of doctors, that list size and hospital referrals are not systematically associated, is generally confirmed by studies of individual doctors. An early, but widely-quoted, study of the referral habits of a group of general practitioners is that reported by Starey (1961). The study involved 30 GPs, practising in Berkshire, Buckinghamshire and Oxfordshire, who kept records on each patient referred to a hospital out-patient clinic over a three-month period in 1960. The data are presented in a way that enables the list size of each doctor to be related to his annual number of referrals (calculated as four times the number of referrals in the three month recording period) expressed as rates per thousand patients on the list and per thousand consultations each year (calculated as 52 times the average weekly number of consultations). The results show that, however the data are manipulated, no clear associations are to be found between list size and either of the two referral rates. An equally wide range in referral rates was found among doctors with large as with small lists. For example, among the 8 doctors with lists of 3,000 or above, the range in the annual number of referrals was from 21 to 107 per thousand patients at risk, whilst among the 8 doctors with lists of 2,000 or below, the corresponding range was from 46 to 122. The only observation offered by Starey about the influence of list size on referral habits (anticipating Wright's finding) was that 'the average referral rate (per thousand consultations was slightly higher in practices numbering over 3,000 patients than in those with 2,000 or less' (page 221), although the reverse was true for the average referral rate per thousand patients at risk. Starey concluded that 'there were... wide differences between one practice and another of the same type ... (which) can only be explained by the differing methods and ideas of individual practitioners' (page 221).

Other studies of individual practices, when put together, further confirm the conclusion that list size has very little to do with a patient's chances of being referred for specialist care or opinion. Studies that relate list size to the annual number of referrals per thousand consultations, or per thousand patients at risk, or both, have been reported by Barber (1971), Berkeley (1976), Bolden and Morgan (1975), Fraser et al (1974), Fry (1972), Marsh and McNay (1974b), Morrell et al (1971), Price (1973), Hopkins (1956), Carmichael and Stevenson (1963) and Evans and McBride (1968). They paint a consistent picture of very wide differences in out-patient referral rates that bear no constant relationship to practice size. Referrals per thousand patients in these studies ranged from 8 per annum in a practice of 2,100 patients to 205 per annum in a practice of ,1370 patients; and referrals per thousand consultations ranged from 14 per annum in a practice of 1,862 patients to 56 in a practice of 4,504 patients. Notwithstanding all the inadequacies of making comparisons between practices which are anything but comparable, the weight of available evidence strongly supports the conclusion that large variations exist between individual GPs in the rate at which they refer patients for specialist care or opinion, but that, whatever factors may be associated with such variations, practice size is not among them.

Requests for diagnostic investigations

Turning next to the request for diagnostic tests, data collated by the Department of Health and Social Security show an increase in the number of GP-initiated pathology requests in England from 1.0 million in 1959 to 7.8 million in 1976, and an increase in the number of GP-initiated referrals for radiological investigation from 1.8 million in 1959 to 24.6 million in 1976 (DHSS, 1977). The 1976 figures yield rates per thousand population of 169 for pathology requests and 529 for radiological investigations. (It would appear that the method of counting these units is not compatible with the way in which general practitioners themselves usually record their requests, for there is a substantial discrepancy between these rates - especially the rate for radiological investigations - and the rates that typically emerge from studies of GPs' behaviour.) The Second National Morbidity Study (OPCS, 1974) reported an overall rate of referral for investigation of 110 per thousand population, with regional variations from 77 in Yorkshire and Humberside to 164 in the East Midlands. (These figures appear to include radiological as well as pathological investigations.) Rose and Abel-Smith's (1972) study of 132 doctors in three areas of one county showed that the estimated annual numbers of requests for pathological investigations, expressed as rates per thousand patients on the doctors' lists, were 130, 130 and 100 in the three areas, in

spite of substantial variations between the areas in the proportions of doctors using the laboratory facilities during the three-month survey period. No information was collected about list size, however.

Studies that have included, or made reference to, list size are divided in their estimation of the significance of list size in explaining the use of open-access diagnostic facilities. Some writers discount any relationship between the two variables. Forsyth and Logan (1960, 1968) explicitly noted the absence of any association between list size and the rate of use of direct-access facilities in their two studies; and Logan (1964) reached a similar conclusion from his earlier work on the use of direct-access facilities across 12 towns in England. Data are given, for example, on 34 practices in Bolton, in which referral rates for X-ray and pathology are categorised by type of practice (solo vs. group) and list size ($\leq$ 2,500 vs. $>$ 2,500); the results show that although the rates appear to be very low by the standards of subsequent studies, they were not related to list size. Logan commented that 'this suggests that the decision for the GP to refer a patient for ... direct-access investigation is intrinsic and within himself rather than in the external circumstances of the practice' (page 19). More recently, Taylor et al (1975) recorded the numbers of vaginal swabs, faecal specimens, throat swabs and urine specimens submitted by 104 general practitioners in Aberdeen during 1973-4, and noted the wide variation between individual practitioners in their use of these laboratory facilities. However, in discussing the possible reasons for such variation, Taylor et al discounted the structural features of the practices and their populations. 'It seems most unlikely that such a large variation as that between, for example, the five doctors who submitted 62% of the total throat swabs and the 40 who submitted none could be wholly or even largely explained by such factors. Insofar as list size can be taken to indicate differences in workload, our own calculations showed that the use of results based on rates per 1,000 patients per year made little difference to the positions of individual doctors in the rank orders' (page 636). Two pieces of circumstantial evidence further supporting the view that practice size is unrelated to the use of open-access diagnostic facilities are provided by Green (1973) and Backett et al (1966). Green's extensive review of the literature on the use of open-access pathology services by general practitioners led him to construct an 'identikit' picture of the typical high pathology user (page 323); but practice size did not appear as one of the distinguishing features. Backett et al noted from their study of the use of hospital services by GPs in North-East Scotland that doctors who were high users of open-access facilities were also high users of outpatient clinics. Although Backett et al give no information about the list sizes of these 'high use' doctors, the fact that referral rates to outpatient

clinics appear to be influenced by list size suggests that rates of use of diagnostic facilities may likewise be unaffected by list size.

Against this weight of evidence, however, are ranged the conclusions of other studies claiming to have found an association (usually inverse) between practice size and the use of open-access facilities. Scott and Gilmore (1966) analysed the use made of laboratory and X-ray services by all GPs in Edinburgh during a three-month period in 1962. The table showing the relationship between the use of these services and practice size is ambiguous because it does not delineate a time-span and it does not clearly distinguish between laboratory and X-ray services (Table 7, page 25); but a general trend is evident. In single-handed practices, the rates of use per thousand patients at risk were 3.8, 10.8 and 11.5 respectively among doctors with 3,000 or more patients, 2,000-2,999 patients, and less than 2,000 patients. In partnerships the rates were 12.7, 20.8 and 20.3 respectively. Scott and Gilmore commented that 'in both single-handed and partnership practices the use of open-access facilities varies inversely with the size of the practice, but the effect of practice size is greater in single-handed than in partnership practices' (page 25). A possible clue to the explanation of this association is found in an earlier paper by Scott (1964) describing the work of the Family Doctor Diagnostic Centre in Edinburgh. 'The busy, frustrated or overworked doctor of necessity develops a number of protective mechanisms to avoid taking decisions. Among these mechanisms, which can become almost reflex reactions, is the blurring of the difference between diagnosis and treatment. The student is taught that in the classic sense diagnosis must precede treatment. The GP knows only too well that in absolute terms it is comparatively seldom that he has a clear-cut diagnosis in respect of each consultation' (page 129).

Other studies offer some support for the findings of Scott and Gilmore. In Wright's (1968) survey of 77 practitioners in the South-West England Faculty of the Royal College of General Practitioners, GPs with fewer than 2,000 patients were found to request routine pathological investigations at an annual rate of 85 per thousand population and 18 per thousand consultations, compared with rates of 51 and 15 respectively for GPs with more than 3,000 patients. Wright noted that 'doctors with smaller lists tend to investigate their patients more fully than their colleagues with larger lists' (page 25). However, the rates for radiological investigations were reversed, being higher among doctors with more than 3,000 patients than among those with fewer than 2,000 patients. Price (1973) recorded the use of pathological and radiological investigations by a suburban Manchester practice over a three-year period (1968-1971) in which the average list size of each partner declined from 2,750 to 2,467, and found that the estimated annual use of X-ray and pathology

services, expressed as rates per thousand patients, increased as the average list size in the practice fell. Annual X-ray requests rose over the three-year period from 42 to 51 per thousand patients, and pathology requests rose from 122 to 158 per thousand patients. Evans and McBride's (1968) study of a group practice in Stratford-upon-Avon assembled data on all X-ray requests and haematology investigations initiated by each of the four partners in the practice between March and September 1966. The list sizes of the four partners ranged from 1,500 to 2,400. After adjusting the data to a yearly basis, the results of the study showed an almost consistent positive relationship between list size and the use of these facilities. For example, annual X-ray requests per thousand patients increased in an almost linear fashion from 22 for the partner with 1,500 patients to 45 for the partner with 2,400 patients; and requests for haematology investigations increased linearly from 3 per thousand patients for the partner with 1,500 patients to 15 for the partner with 2,400 patients. Lastly, the report from the Birmingham Research Unit of the Royal College of General Practitioners, derived from the first 100 proformas received from volunteer practitioners, and referring to a total of 32,000 consultations carried out in a two-week period in August and September 1977, noted that requests for investigations declined as the numbers of consultations undertaken by the participating doctors increased in the two-week period (Journal of the Royal College of General Practitioners, 1978a). However, as with the comparable report from the Unit on outpatient referral rates (see page 90), the participating doctors cannot be regarded as a representative group, and as the report pointed out, 'information about list size was not sought and therefore we cannot relate these results to it' (page 62).

To summarise this sub-section on the use of diagnostic tests, no clear conclusions can be drawn about the way in which this aspect of the content of care is associated with list size. Studies have been cited, covering a wide range of dates, which discount any such association; others claim to have found a consistently greater use of diagnostic tests among doctors with smaller than with larger lists. In this respect, at least, the picture is more clouded than in the case of outpatient referrals. It is further complicated by the possibility that GPs with low rates of use of open-access diagnostic facilities are doing their own tests. The report from the Birmingham Research Unit (see above) found that, across a range of pathological and radiological investigations, specimens were collected and analysed in the practice in 25 out of each 1,000 consultations; in comparison, specimens were collected in the practice and analysed elsewhere in 56 out of each 1,000 consultations, and they were collected and analysed elsewhere at a rate of 31 per 1,000 consultations. However, no information is available on any variations in these rates between practices of different sizes. A further important deficiency in much of the literature, to

which Green (1973) has drawn attention, is the absence of any attempts at multi-variate analysis of the factors underlying the use of open-access facilities. Several studies have included sufficient numbers of doctors to make this feasible, and the results of such analysis would highlight which among a number of inter-related variables are most clearly associated with differences in the use of these facilities.

Treatments used by GPs

Prescribing

Lastly in this section we turn to the meagre evidence on the procedures used by general practitioners themselves. As noted above, the fullest information is available for the two procedures which are required to be recorded for administrative or executive reasons: prescribing and certification. The 1977 Report of the Review Body on Doctors' and Dentists' Remuneration contained data showing an increase in the number of prescriptions written by general medical practitioners from 296 million in 1970 to 335 million in 1975 (Review Body, 1977). The 1975 figures yield average annual prescribing rates of 13,700 per GP, 6,254 per thousand registered patients, and 2,085 per thousand consultations (assuming an average consultation rate of 3.0 per patient per year). No information was contained in the Review Body Report about variations in prescribing rates between doctors with different list sizes, but several post-war studies have examined the relationship between the two variables. The results are inconclusive.

An early study by Dunlop et al (1953), based upon a 1% sample of all prescriptions issued in Scotland during October 1951, related prescribing rates to the factors thought to influence them. A correlation of +0.7 was found between the number of ECIO forms issued per unit of population in each area and the mean list size of the areas. Dunlop et al commented that 'it will be readily understood that script rate may be related to list size simply because the latter depends on population density' (page 696), although Martin (1957) observed that general social conditions affecting both morbidity levels and the attractiveness of areas to doctors may have been a more plausible explanation for the correlation than population density per se. Martin's own study was also an ecological one, involving an analysis of the relationship between the prescribing behaviour of GPs in 67 medium-sized county boroughs in 1951 and all other available information about the boroughs that 'had the remotest likelihood of being related to prescribing' (page 68). A variety of statistical techniques were employed in analysing the data. Correlation analysis failed to confirm the results of Dunlop et al: no significant correlation was found between the mean

list size of the 67 boroughs and the average number of prescriptions issued per patient in 1951. Various measures of morbidity, and historical patterns of prescribing, were correlated much more closely with prescribing frequency in 1951 than was list size. However, factor analysis indicated that 'frequency of prescribing is fairly closely associated with ... large lists' (pages 92-3); from this perspective, therefore, the results of Dunlop et al were confirmed by Martin.

A similar exercise by Gray and Cartwright (1953) yielded more direct confirmation, although the results of this study are not strictly comparable with those of either Dunlop's or of Martin's. Using interview data from some 2,000 adults living in county boroughs who were interviewed in the Survey of Sickness in February and March 1952, Gray and Cartwright compared the weekly consumption rates of all prescribed medicine, and the proportions of consultations resulting in a prescription, between county boroughs with differing ratios of population to doctors. The weekly consumption rate of all prescribed medicine per 100 adults was positively associated with the population/doctor ratio, declining regularly from 15.6 in boroughs with more than 2,900 people per doctor to 10.7 in boroughs with fewer than 2,300 people per doctor. There was also some evidence of a positive association between the population/doctor ratio and the proportion of consultations resulting in a prescription. In boroughs with more than 2,900 people per doctor, 83% of consultations yielded a prescription, whilst in boroughs with fewer than 2,300 people per doctor, the proportion dropped to 66%. Gray and Cartwright concluded that 'there is a suggestion that prescribing increases ... with the size of doctor's list' (page 28).

Yet another ecological study supporting this conclusion is that of Joyce et al (1967). The study assembled information on all prescriptions issued in one month in 1962 in three industrial towns, and this was subsequently related to the personal characteristics of a sample of the prescribing doctors, derived from personal interviews with them. The results showed that the mean number of prescriptions issued in the study month by each doctor interviewed was 1,169 in town 1 with average list of 3,038 patients; 1,174 in town 2 with an average list of 3,349 patients; but only 718 in town 3 with an average of 2,624 patients per doctor. Joyce et al commented that 'the mean frequency of prescribing in town 3 was consistently lower, in most instances significantly so' (page 175). The study also confirmed the conclusion of several other investigators that a lower age, higher educational qualifications and an orientation towards 'whole person medicine' were associated with lower prescribing of drugs of all kinds.

Against this evidence, a few other studies have reported no observed association, or a negative association, between list size and prescribing rates.

In a study of the drugs prescribed by a small sub-sample of the 142 doctors participating in the 1969/70 North-East Scotland workload study (Richardson et al, 1973, see page 82), Berkeley and Richardson (1973) related variations in annual prescribing rates to a number of indicators of practice structure (including list size) and the doctors' personal characteristics, 'but no significant correlations were found' (page 160). (No data are given in the paper in support of this conclusion.) The only study that has been discovered to report a negative association between list size and doctor's perceptions of their prescribing is that of Dunnell and Cartwright (1972). Data from interviews with a random sample of 1,412 adults in Great Britain showed that about two-thirds of the consultations they had had with their GPs during the two weeks prior to the interviews had resulted in a prescription being written. In a subsequent postal survey of the GPs concerned (to which 326 doctors, or 56%, responded), each doctor was asked whether he thought he gave prescriptions more or less frequently than this. The results showed that 29% of doctors with fewer than 2,000 patients felt that their prescribing was above the average, compared with 22% of doctors with 2,000-2,999 patients, and 13% of those with 3,000 or more patients. Dunnell and Cartwright commented that 'this variation in prescribing patterns with list size seems somewhat surprising in the light of (another) finding that many doctors felt they would write fewer prescriptions if they had more time' (page 76). However, further analysis of the replies showed that doctors with small lists were actually less likely than those with large lists to think that they would write fewer prescriptions if they had more time. 'Those findings suggest that if doctors with larger lists had more time/they might see their patients more often and would therefore not cut down their prescribing in the way they predict' (page 76).

Data from studies of individual practices that relate prescribing rates to list size have been reported by Bain and Haines (1975); Weston Smith and O'Donovan (1970); Hughes et al (1976; repeat prescribing only); and Barber (1971). They add very little to the studies summarised above: in most cases list sizes in these practices were larger than the national average, but the prescribing rates (whether based upon the number of consultations or the number of patients at risk) ranged around the national average rates in quite unpredictable ways.

It seems, therefore, that the weight of evidence points towards a positive association between list size and prescribing rates. The most substantial investigation refuting this conclusion (by Dunnell and Cartwright) employed a somewhat indirect measure of the rate of prescribing, and cannot be compared directly with the earlier work of Dunlop, Martin and Grey. However, much

caution is required in interpreting this conclusion. First, prescribing rates have no intimate connection with the quality of care. Doctors with larger lists may be prescribing more frequently than those with smaller lists because the pressures under which they perceive themselves to be working may prohibit the use of other more satisfactory (but more time-consuming) patterns of management. A second problem arises from the failure in several of these studies to distinguish adequately between the number of scripts issued and the total number of different preparations prescribed. Textual interpretation often implies a concern with numbers of preparations, whilst the data seem to relate to numbers of scripts. Third, it is often impossible on the basis of the published results to distinguish between repeat prescriptions (with or without a consultation) and new prescriptions. Finally, the base that is used in the calculation rates, will affect the nature of the conclusions that are drawn.

Certifying

Information about certificates issued by GPs in connection with new claims for injury and sickness benefits was given in the 1977 Report of the Review Body on Doctors' and Dentists' Remuneration. The total number of these certificates issued by GPs in 1975 was 10.2 million, yielding average annual rates of 419 per GP, 191 per thousand registered patients, and 64 per thousand consultations. Much of the published research and commentary on sickness absence has concentrated on the characteristics of workers who claim sickness absence benefits and on the social and economic circumstances that are associated with variations in rates of absence (Office of Health Economics, 1965). Less has been written about the contribution of the GP to certification rates, although the arguments developed in this report suggest the plausibility of an hypothesis linking the size of a doctor's list with his frequency of sickness absence certification. Information about the frequency of certification in individual practices has been published by Handfield-Jones (1964), Ashworth (1965), Grossmark and Sharer (1967), Carne (1969) and Morrell (1971). These studies confirm that, as with most activities in general practice, there are wide variations between individual practitioners in their certification rates, and for most GPs, the issuing of certificates forms a not inconsiderable part of their workload. As Taylor (1974) has put it, 'the time has come to admit quite openly that medical certificates are now, for all practical purposes, issued on demand' (page 330). However, the studies cited above show no apparent relationship between certification rates and list size, and in any case, the data they present are rarely comparable, dealing with varying time periods and age groups, and not always permitting the calculation of rates on comparable bases.

Other treatments used by GPs

Apart from referring and issuing prescriptions and sickness absence certificates, surprisingly little information has been published about what GPs actually do, and almost nothing that permits even tentative statements to be made about typical differences in the content of care between those with large and small lists. Information about the proportion of consultations that involve a physical examination of the patient has been published by Morrell (1971) and Hull (1968). Morrell's data showed that, over a period of one year in a teaching practice with an average list of 1,485, physical examination of the patient was carried out at a rate of 850 per thousand new consultations and 524 per thousand follow-up consultations. Hull's data are not directly comparable with Morrell's, being confined to 1,000 consecutive new cases presenting over a three-month period in a rural practice with an average list of 2,625 patients per doctor. The results showed local examination of the patient was performed at a rate of 480 per thousand consultations, system examination at a rate of 300, and general examination at a rate of 180 per thousand consultations. Barber's (1971) report of one year's activity in a practice of 1,612 patients in a Scottish new town contained information about the use of non-drug treatments. For example, 'advice' (with or without other treatments) was used at a rate of 130 per thousand consultations during the year, and the next most common treatments were 'dressings' and 'diet', both of which were used at a rate of 11 per thousand consultations. By contrast, drugs were prescribed, on average, at 866 out of each thousand consultations. Barber is at pains to point out probable inaccuracies in the data on non-drug treatments, but his study appears to be one of the very few that has collected any information at all about the content of practice other than prescribing, certifying and referring.

Only two reports have come to hand that point, however cautiously, to possible variations in the content of care between doctors with larger and smaller list sizes. Wright's (1968) study of 77 practitioners in South-West England elicited information about the number of 'practical manouevres' usually carried out by them. Nine such procedures were listed, including the stitching of cuts, the injecting of piles and varicose veins, the manipulation of fractures and joints, and the cauterization of cervices. On average, each doctor usually performed 3.1 of the nine procedures, but the average was higher among the 21 doctors with more than 3,000 patients (3.5) than among the 18 doctors with fewer than 2,000 patients (2.6). Other factors associated with a high use of these procedures included rural practice, partnership practice, and being over 50 years of age.

Cartwright (1967) used similar methods and reached similar conclusions. Questions were included in her postal survey of a national sample of GPs asking

respondents to judge the frequency with which certain procedures were carried out in their practices. Seven procedures were listed, including the stitching of cuts, the strapping of sprains, the estimation of haemoglobin and the use of a laryngoscope. The replies were scored in a way that favoured doctors carrying out procedures which the majority did not (appendix 6, page 276). The average score for all doctors was 3.7, but it was somewhat higher among those with 2,500 or more patients (3.9) than among those with smaller lists (3.5). Cartwright also reported some data on other aspects of care. 'There was no variation with list size in the proportion who thought it appropriate for general practitioners to be consulted about such things as children getting into trouble or family discord, nor was there any difference in their views on cervical smear tests. But the proportion who thought that ideally general practitioners should carry out some (other) regular check-ups on middle-aged people increased from 41 per cent of those with lists of under 1,500 to 58 per cent of those with lists of 2,500 or more. Possibly those with larger lists are more likely to have come across conditions which might have been picked up by such check-ups' (page 163). Cartwright's final conclusion, however, was that 'the size of the doctor's list seems to make comparatively little difference to the doctor's own perception of his role' (page 163).

THE QUALITY OF CARE IN GENERAL PRACTICE

The seventh question identified on page 63 concerns the quality of care that is delivered from practices of different sizes. It is an important consideration in the judgement of a reasonable list size, for it has been argued that society's decision about the number of doctors that it wishes to have must reflect the value it places upon any variations in the standards of care that may flow from an increased input of manpower resources. Variations in consultation times and rates, and in the content of care, that are associated with changes in the availability of manpower are significant intermediate data in reaching an appropriate evaluation, but they fall short of the firmest touchstone. For example, the apparent tendency for doctors to reduce their consultation rates in response to increases in practice size might be regarded as an indicator of an undesirable decline in the overall quality of care: fewer patients may be receiving the amount of care that 'society' wishes them to have, and some patients may not be receiving the amount of care for specific problems that 'society' regards as appropriate to their needs. In either case, however, an understanding of the variations in consultation rates that are associated with differences in list size offers, by itself, an imperfect basis for evaluation. Further questions need to be asked about whether these variations matter, and such questions lead inexorably into the treacherous swamps of quality.

Before setting a tentative foot into the mire, two self-protective comments must be lodged. First, it is not merely the eccentric view of a lone academic that questions about the quality of care are of central relevance to the judgement of a reasonable list size. Many of the expressions of opinion reviewed in the first two sections of this report (especially opinions originating from within the medical profession) reflect the view that list sizes are too large to permit the practice of an acceptable standard of care, and should <u>on these</u> grounds be reduced. The analysis rehearsed in this report, if it is correct, identified some critical questions to be asked about this view. What dimensions of quality are sensitive to variations in the input of manpower resources, and can they be arrayed in a way that enables policy-makers to choose the point at which further gains in quality cease to justify the cost of achieving them? The second protective comment is that, for present purposes, the consideration of quality can be confined to those aspects that are plausibly related to the number of patients for whom a GP provides care. Much of the voluminous literature on the quality of care can conveniently be sidestepped on these grounds.

The concept of quality, especially as it is applied to medical care, has been described as 'a portmanteau word that can be stretched almost without limit to have packed into it whatever one chooses' (Gillie, 1966, pages 1-2). The Oxford English Dictionary offers a lead by defining the quality of something as being its 'nature, kind or character', and it suggests that the word is most used in a comparative sense to assess the degree or grade of excellence possessed by something. The notion of quality as a relative rather than an absolute attribute is particularly helpful for the present purpose of contrasting the quality of care between different practice contexts, not of contrasting achieved quality with some ideal notion of absolute quality.

The structure of care

Much of the literature about the quality of medical care acknowledges a debt to the pioneer writings of Sheps (1955) and, particularly, Donabedian (1966), whose tripartite division of care into the component parts of structure, process and outcome has influenced many subsequent writers. The structure of care concerns the 'settings in which it takes place and the instrumentalities of which it is the product. ... It is concerned with such things as the adequacy of facilities and equipment; the qualifications of medical staff and their organisation; the administrative structure and operations of programs and institutions providing care; fiscal organisation and the like' (page 170). The assumption is made that, given the proper settings and instrumentalities, good medical care will follow. The classic studies by Peterson et al (1956) and Clute (1963) in North America included measurements of the facilities, equipment and training that were assumed to be the requisites of 'good' general practice; and in this country the critical descriptive accounts of general practice by Collings (1950), Hadfield (1953) and Taylor (1954), and studies such as that of Eimerl and Pearson (1968) into the equipment available to general practitioners in their work, are representative of this approach. There is some evidence that these structural aspects of the quality of care are not only differentially related to list size, but may actually be more favourably represented among practices with larger lists. Support for this conclusion at an ecological level is offered by the study of Butler et al (1973), based upon a postal survey of 1,721 GPs in England. They concluded that 'conditions of general practice in designated areas are somewhat more aligned to contemporary notions of good medical care than those in restricted areas. To the extent that partnerships, based on health centres with a full range of ancillary help, and with adequate free time for the GP to study and relax, are accepted as valid signs of good general practice, then the greatest room for improvement is seen in those places with the best doctor/patient ratios' (pages 109-110).

A similar conclusion was reached by Bridgstock (1976), based upon data from the Swansea national cohort study. Cartwright's (1967) national survey of 422 GPs in England and Wales found that those with larger lists were more likely to attend continuing education courses, and more likely to employ ancillary help in their practices, than those with smaller lists. Great care is required in interpreting results of this kind, for list size per se may be a less salient independent variable than others (such as urban or rural location) which interact with list size in quite complex ways.

The process of care

Donabedian's second component category of care is that of process. 'Judgements (about quality) are based upon considerations such as the appropriateness, completeness and redundancy of information obtained through clinical history, physical examination and diagnostic tests; justification of diagnosis and therapy; technical competence in the performance of diagnostic and therapeutic procedures; evidence of preventive management in health and illness; co-ordination and continuity of care; and so on' (page 169). The rationale underlying process studies is that, if the proper things are done, the outcome will be good. The designation of things, or actions, or modes of acting as 'proper' may be explicit or implicit. Explicit designations require the prior agreement of a group of clinicians about the criteria that may validly be used to indicate good quality in the process of care: such predetermined criteria are then applied to the work of the doctors under review. The Professional Standards Review Organisations in the USA rely heavily upon the application of explicit criteria in assessing the quality of hospital care. Implicit designations involve the judgement of each case on its own merits, with each judge or auditor making reference to his own internalised notions of what is proper in that particular case. Self-audit and case conferences are familiar examples of this.

Although the language and concepts of audit and quality assessment in medical care have become increasingly familiar in this country during the last five years, they remain alien to many doctors and to most investigations in general practice. The studies reviewed in the preceding section, for example, illustrate the kind of information that has typically been collected about the process of care in general practice, and they indicate ways in which this process might be influenced by the numbers of patients for whom GPs have to care; but they lack the evaluative dimension that would categorise certain modes of care as being 'better' than others. The only valid conclusions to be drawn from such studies is that the total pattern of care (including consultation rates as well as the content of care) appears to differ in fairly predictable

ways between larger and smaller practices. Because judgements have rarely been made, either explicitly or implicitly, about the 'desirability' or 'correctness' of certain patterns of care, no conclusions can be drawn about the relationship between list size and the quality of the caring process. Any statement about quality requires an evaluative judgement about the relative merits of different ways of doing things.

In spite of this general limitation in the research literature, several studies have been published that do incorporate an evaluative dimension, and that indicate ways in which future research might progress, even though they yield few substantive clues about the significance of list size. The most useful indicators in this respect are the reported self-judgements of GPs that the workload pressures arising from large lists prevent them from giving the quality of care to their patient that they would wish. Mechanic's (1974) investigation has already been referred to in this connection (see pages 14 and 29). Across a variety of dimensions of care, the respondents in this study expressed increasing dissatisfaction with the quality of work as their list sizes increased; but, as noted above, the data are not sufficiently detailed to permit inferences about the point at which further gains in quality cease to justify the cost of the additional manpower. One way of building upon Mechanic's results would be through the collection of sensitive data about the amount and intensity of the workload in practices of different size, and the coping mechanisms adopted by the practitioners, in order to provide substance to what are at present the somewhat disembodied expressions of opinion by the participating doctors.

Another category of evaluative studies in general practice embraces the attempts that GPs have made at self-audit. Several examples are to be found in the literature. Doney (1976) described an internal audit of the care of diabetics in a group practice of 20,175 patients. After noting the pattern of the disease in the practice, the way in which the criteria of diabetic control were recorded in the case notes, and the nature of the care received by diabetic patients, Doney formed the judgement that 'the strict recording of the criteria of diabetic control and the regular follow-up of diabetic patients were poor in this practice in the period under study' (page 741). Gruer *et al* (1977) reported a collaborative audit involving GPs, surgeons and community physicians, aimed at improving the diagnosis and management of acute abdominal pain. Reviews were made of the diagnosis and management of 407 patients referred to a Scottish hospital over a six-month period with symptoms of acute abdominal pain, and criteria were drawn up by the hospital staff and GPs together that might improve the diagnostic accuracy of the GPs and reduce the number of 'unnecessary' referrals. Subsequent studies were carried out to test whether

the application of the criteria did in fact lead to the better management of acute abdominal pain both by hospital doctors and general practitioners.

Reilly and Patten (1978) reported an audit of the prescribing behaviour of GPs in a teaching practice in Belfast. Records were kept of all the prescriptions written by the doctors during a one-week period, and the details of each prescription were married to other information about the patients, drawn from the record cards. The results were then presented at a meeting of the doctors, and their reactions to them were recorded. Predictions were made by the investigators about the degree and direction of future changes in the doctors' prescribing behaviour, and these were tested against their actual prescribing behaviour during a second survey week. It was found that, in comparison with the initial survey week, there was a slight reduction in the number of items per prescription, in the number of indirect prescriptions, and in the average cost of each item of medication. These changes were consistent with the views expressed earlier by the doctors about 'desirable' standards of prescribing. In an earlier study in the same Belfast practice, detailed records were kept of all the patient contacts by the primary care team in a one-week period, and in each case the team member was requested to make evaluative judgements about the continuing care that would be required in order to meet the patient's needs satisfactorily (McCready and Reilly, 1977). The authors concluded that 'discussion by the workers of findings such as these could result in a setting of objectives that would result in the more efficient use of primary care resources' (page 530).

Yet another approach to assessing the quality of the process of care is represented by Hodgkin's (1973) work on 'delay pattern analysis'. Starting from the premise that 'delay is often but by no means always an indication of slack or inefficient attitudes and skills in the doctor' (page 761), Hodgkin devised a method of recording and combining information about the delays displayed by patients in reporting their symptoms and the delays displayed by GPs in suspecting a correct diagnosis and in starting treatment. By applying delay pattern analysis to different doctors in different situations, Hodgkin believed that 'it is possible to produce a consensus picture that will allow doctors to evaluate their own performance' (page 762).

Studies such as these are illustrative of the pioneering methods being used to evaluate the quality of the caring process in general practice; but as noted above, the literature offers very few clues about the salience of list size in determining the quality of the caring process. Indeed, very specific hypotheses would be required to implicate list size as an independent variable. One such hypothesis might use the average length of consultations as an

intervening variable by positing that shorter consultations would occur in larger practices, and would be associated with more hurried and less satisfactory styles of practice. This appears to be one way of interpreting the results of Mechanic's (1974) survey, but the difficulty has already been noted of reconciling this interpretation with the absence among the research literature of any clear evidence that doctors with larger lists do actually have shorter consultations than those with smaller lists (see page 73). An alternative hypothesis, building upon the conclusion reached earlier in this report about the negative association that typically seems to exist between list size and consultation rates, might focus upon the quality of care that is given to a practice population as a whole, rather than to those individual members of it who happen to cross the surgery threshold. This approach would appear, on the basis of the conclusions reached elsewhere in this report, to offer an interesting area for further investigation. The notion of the quality of care to communities is fairly well established in the general practice literature (see, for example, Tudor Hart, 1971; Stevens, 1977), but as the studies reviewed above have demonstrated, most empirical studies of the quality of care have focused exclusively upon events occurring within the domain of the surgery, and have ignored the ways in which practice size determines who does (and who does not) enter the domain in the first place.

The outcome of care

Donabedian's third component category of care is that of outcome. Outcome studies involve an assessment of the end result of care, and they require some way of measuring what has actually happened to the patient as a result of his encounter with the medical care system. Donabedian noted that the validity of outcome as a dimension of quality has seldom been questioned, and the same sentiment is conveyed by Christoffel and Lowenthal's (1977) observation that 'outcome is regarded by a growing number of researchers as the most accurate and important index of the quality of health care' (page 888), and by Mansfield's (1973) criticism of much current work into the doctor-patient relationship that it 'pays too little regard to the fact that it is a means, not an end' (page 892). The primacy of outcome measures as an indicator of quality is perhaps self-evident, for evaluations of the quality of the structure and process of care rest ultimately upon the effects they produce. It is difficult to see in what sense one kind of structure or process is 'better' than another unless it is more likely to produce the change or improvement in the patient's condition for which the clinician is aiming, and which, on the basis of contemporary medical knowledge, he may reasonably expect to achieve.

Nevertheless, difficult questions surround both the choice of appropriate

measures of the outcome of care and the links between structure, process and outcome. It is platitudinous to observe that changes in health status may be influenced by many factors in addition to medical care, but that merely intensifies the problem of disentangling the specific effects of medical care from the residual complex bundle of inter-related influences. The problem of attributing cause to effect is heightened in the case of general practice, where many of the conditions that are treated will improve spontaneously whatever remedial measures are applied. The Royal College of General Practitioners (1973) has estimated that some two-thirds of all conditions seen in general practice are minor and self-limiting, and as Ginzberg (1975) has pointed out, it is unproductive to use outcome measures to evaluate quality for self-limiting illnesses or illnesses where no effective intervention is possible.

In spite of these difficulties, attempts have been made to specify appropriate indicators of the outcome of primary medical care. Irvine (1976) listed eight indicators that were drawn up by an RCGP meeting in 1974. These are: (1) prevention of disease or control of the disease process; (2) improvement or preservation of the patient's level of function in his family, at work and in his social activities; (3) relief of the patient's symptoms, distress, and anxiety, and avoidance of iatrogenic symptoms; (4) prevention of premature death; (5) minimising the cost of illness to the patient and his family; (6) giving the patient satisfaction with his care; (7) relieving or at least clarifying the patient's interpersonal problems; and (8) preserving the human integrity of the patient from an ethical point of view. However pertinent these indicators might be as an agenda for future research, it is apparent from a careful review of the literature that they have rarely been applied in a systematic way to the evaluation of primary medical care. Moreover, the few studies that have attempted to measure outcome have usually confined their attention to the care provided to individual patients, not to a practice population. Mourin's (1976) work on thyroid dysfunction is illustrative of this type of study. He defined the criterion of outcome as the maintenance of patients in a euthyroid state, and with this as the yardstick, Mourin was able to evaluate the quality of care given in his practice to 35 patients with present, previous or potential thyroid dysfunction. Valuable though this kind of study undoubtedly is, it is perhaps of somewhat limited use in the present context, for there is no way of linking these results to those of other studies for the purpose of assessing the influence of list size on the quality of care.

Two other studies offer interesting alternative approaches. Thomas (1978) reported briefly on an inquiry conducted in his own practice to test the hypothesis that longer consultation times would generate more favourable clinical outcomes for patients than shorter consultations. Two hundred patients in whom

no diagnosis could be made were randomly allocated for one of four treatment styles: short consultations ($\leq$ 5 minutes) at which no treatment was given; short consultations at which some treatment was given; long consultations ($>$10 minutes) at which no treatment was given; and long consultations at which some treatment was given. The outcome criterion used was simply that of whether or not patients had returned to see any of the doctors in the practice with the same or a different complaint within one month of the intial consultation. All patients were invited to return within a week if they were no better. The results (by which Thomas declared himself to be 'surprised') showed no significant difference in outcome between the four groups: in each case, about three-quarters of the patients did not return at all, and about one-in-ten returned with a different complaint. The adequacy of Thomas's outcome criterion can easily be criticised, but the study is unusual in focusing directly upon one of the mechanisms that have been claimed to mediate between list size and the outcome of care, namely the length of time that doctors are able to spend with their patients. Taken at face value, the results seem to cast further doubt on the validity of this claim. Not only is there no clear evidence that doctors with smaller lists do actually spend more time, on average in each consultation, it also appears that longer consultations may not generate a more favourable outcome even when they do occur. However, there is much that can be criticised about Thomas's study, even though it marks an initial attempt to meet Buchan and Richardson's (1973) plea for a systematic exploration of the consequences flowing from a longer consultation time (see page 52).

The second approach that suggests a way of relating practice size to the outcome of care is represented by the 'iceberg' concept developed by Last (1966, 1971). By concentrating attention upon the amount of <u>untreated</u> morbidity in the community, rather than upon the effects of treatment of <u>known</u> morbid episodes, an insight is available into the quality of care to a whole population, not just to those members who come within the medical care system. Last commented, with respect to his study of 171 general practitioners, that 'this was a sensitive measure of the quality of care. (The doctors') own testimony of what they had seen in their practices during a full year was compared with the the numbers of certain conditions and pathological processes which had been revealed in surveys of representative samples of total populations. The comparison indicated a considerable component of undetected disease, some of it seriously and potentially lethal, in English general practice. The method ... is available to anyone who cares to use it in evaluating the quality of medical care given to a defined population' (1971, page 6. 10). The importance of this approach lies in the mechanism it suggests for linking practice size to the outcome of care. Whereas Thomas's study can be applied to the hypothesis that

larger lists might lead to a poorer outcome for individual patients by virtue of the lesser amount of time that doctors can spend with each patient, Last's methodology suggests the alternative hypothesis that larger lists might lead to a poorer outcome for practice populations as a whole by virtue of the reduced number of patients whom GPs can see, diagnose and treat during a specified length of time. This latter hypothesis is more in tune with the conclusions reached earlier in this report. Last found that, over the course of a year in the 'average' practice, only about one-third of patients with serious acute or chronic conditions, including unrecognised and pre-symptomatic cases, would actually be recognised by the doctor; but no information is given that permits this result to be related to list size.

One of the eight measures of outcome described by Irvine (1976) that has perhaps been investigated more fully than any other is that of patient satisfaction. One such study has been reported fully by Marsh and Kaim-Caudle (1976). It is an important study because it was designed specifically to test the reactions of consumers to the style and organisation of a practice of over 3,000 patients (see above, page 15). Marsh had claimed that a practice of this size was perfectly compatible with the maintenance of proper standards of care, and whilst he found that the absence of defined and acceptable standards of the quality of clinical care made it impossible to substantiate the claim, he and Kaim-Caudle nevertheless regarded the degree of satisfaction among the practice population as a relevant clue to the standard of care that was being delivered. The study was designed to test several hypotheses, among them that patients were satisfied with the overall health care service they received, and that they were satisfied with the care given by the individual members of the primary health care team. The survey was restricted to adult members of the practice population, and it utilised a stratified sample design with differential sampling fractions. The five strata covered chronic housebound patients; recipients of intensive care; patients whose first contact during an episode of illness had been with nurse; minimal users; and normal users. A total of 417 patients were sampled, and they were interviewed in their own homes, early in 1973, by a team of university-based interviewers. A final interview response rate of 82 per cent was achieved, and the respondents were judged by the investigators to be representative of the full sample.

Marsh and Kaim-Caudle are appropriately cautious about the significance of their results. They point out, for example, that 'the replies reflect the memory, perceptions and opinions of patients; they are not necessarily correct in any other respect. ... Thus, an affirmative reply to the question, "Do the receptionists rush you?" is not evidence that the patients were rushed, but a valid statement that they felt they were rushed' (page 129). Nevertheless,

Marsh and Kaim-Caudle expressed no doubts about their principal conclusions. 'The overwhelming majority of respondents (95 per cent or more), irrespective of age and social class, who expressed an opinion, assessed the practice premises as good or excellent, considered that the receptionists gave an average or good service, were satisfied or very satisfied with the overall treatment they received from the doctor, considered him as concerned or very concerned with the patient's problem, approved both treatment by a nurse at the surgery after they had seen the doctor and also follow-up visits by her after an initial visit by the doctor, and assessed the health visitor's well-baby clinic as well organised and her immunisation clinics as helpful or very helpful' (page 119). Later, they commented that 'these findings thus prove conclusively that the new style of health care ... with its very low doctor workload can give a high level of satisfaction to some 90 per cent of patients' (page 134).

Nevertheless, a close reading of the survey results indicates pockets of expressed dissatisfaction that might have stemmed directly from the method of working adopted by the GP to enable him to cope with such a large list. For example, a quarter of all the respondents felt that the doctor was always or sometimes reluctant to visit them at home, and one in five of these respondents thought that their health had suffered as a result. One in four of all respondents, excluding the chronic housebound, had struggled to the surgery at some time when they would have preferred a home visit by the doctor. More than a quarter reported trouble getting an appointment to see the doctor. With respect to the role of the nurse, the survey found that about one in ten of all patients had been visited by the nurse when they had requested the doctor, and of these, more than half would have preferred to see the doctor. One in six of them felt that their health had suffered because the nurse called rather than the doctor.

It is difficult to evaluate these results. The existence of a measure of dissatisfaction among Marsh's practice population is not necessarily evidence of a fault that ought to be corrected. No form of personal service is likely to be wholly satisfactory to all recipients all the time, and a style of practice in which the doctor always responds to the request for a home visit would be regarded by many as the sign of inefficient or even bad practice. The problem here, as in other dimensions of quality, is that of deciding the point at which expressions of consumer dissatisfaction matter and should be corrected. The decision is not aided by the lack of comparable data from practices of differing sizes and organisations. If it could be demonstrated that a significantly higher proportion of Marsh's patients than of patients in smaller practices are dissatisfied with the doctor's reluctance to visit them at home, and if more of them feel that their health suffers as a result, there may be some basis for

evaluation; but directly comparable data do not exist. For example, Cartwright's (1967) multi-purpose study included information about the views of her sample of 1,397 respondents on various aspects of the care they received from their GPs, but the replies were not analysed according to the doctors' list sizes. With respect to the accessibility of doctors, 84 per cent of the sample thought they would be able to get hold of their gen eral practitioner, or someone acting for him, if they needed him on a Sunday afternoon or in the middle of the night; and just over half of those who had in fact tried to get hold of him in a hurry at some time during the previous 12 months had been able to do so within half an hour. (In evaluating the significance of these findings, it must be remembered that the survey was conducted before the introduction of the 'Charter' and before the widespread use of deputising services.) With respect to the doctor's approachability, 88 per cent of the sample thought their doctor was good about taking his time and not hurrying them, 93 per cent that he was good about listening to what they had to say, and 75 per cent that he was good about explaining things to them fully. These findings contrast interestingly with those of Mechanic's (1974) study of the doctors' perceptions (see above, page 29). At face value, they suggest that doctors may be more critical than patients about the hurried and sketchy nature of a good deal of care in general practice. The overwhelming impression created by these results, however, (and confirmed by the other studies summarised above) is one of substantial satisfaction expressed by most patients in formal interviews about their primary medical care. There is certainly insufficient evidence to draw any conclusions about the way in which patient satisfaction is responsive to practice size. More effort may need to be expended in refining the methodologies by which the attitudes of patients are assessed before any such conclusions can be reached.

THE RESEARCH EVIDENCE SUMMARISED

The previous four sections of this report have attempted to abstract and classify the evidence in the research literature about the consequences of variations in the list sizes of general practitioners. The use of published research data in this way is fraught with difficulty. A pervasive threat to the validity of the entire exercise lies in the assumption that statistical associations between list size and various measures of the pattern and content of care are causal associations, and that changes in list size will therefore produce concomitant changes in the indicators of practice style. Although such an assumption is justified in the analysis of consultation times and rates (where variations in list size must be reflected either in changes in consultation times or rates, or in the total number of hours worked by practitioners), it is less obviously justified in the analysis of the content and quality of care, where a diverse array of factors may confound it. In spite of the confident assertions of the medical profession that large lists generate a poor standard of practice, there are few specific hypotheses identifying the mechanisms by which this link might operate, and they are not clearly supported by the available data. Further problems surround the intrinsic quality of the data used in this analysis. There are substantial doubts about the representativeness of many of the studies used in the analysis, and questions have been raised in the text about the lack of comparability in the definition and measurement of concepts, about the variable duration of studies and the consequent difficulties in deriving comparable annual rates, about the accuracy both of reported list sizes and of the data collected in fieldwork studies, and about the different ways in which similar kinds of data have been analysed and presented. Above all, there are very large gaps in the information that would ideally be needed to examine each question properly.

The first question was whether the total amount of time spent by general practitioners in patient care differs according to the size of their practice lists. Although the literature contains reports of a wide variation in the typical working week of individual practitioners, most studies have located the average number of weekly hours within a surprisingly narrow range, between about 35 and 42 hours excluding time on call. There is some evidence from studies in Scotland, North-East England and Derbyshire that doctors with larger lists may work longer hours than those with smaller lists, but this evidence is judged to be fairly insubstantial, and is not supported by other studies (also from Scotland and North-East England). There does, however, appear to be good evidence from a national study of a large number of GPs that, as list size increases, proportionately more doctors express dissatisfaction with the amount of time they have to spend on their practices.

The second question was whether the average amount of time spent by GPs in consultations with their patients differs according to the size of their practice lists. In fact there is no substantial evidence to justify a positive answer. Although there is a wide range in the reported times taken for individual consultations, and in the average time taken by individual GPs, most studies locate the mean time of surgery consultations between about 5 and 6½ minutes, and of home consultations between about 10 and 15 minutes. Two studies, both in Scotland, that permitted the consultation times of groups of doctors to be related to the size of their lists, each failed to find any association between the two variables. These findings are consistent with the opinion, frequently expressed in the literature, that GPs tend to develop a consistent method of working that is fairly impervious to external pressures.

The third question was whether the average number of consultations made by patients each year differs according to the list sizes of their doctors. Three principal measures of consultation rates were identified in the literature, and several different techniques were used to analyse the large amount of reported information. Accepting these data at face value, the analyses pointed consistently towards a broad negative association between list size and both surgery and (particularly) home consultation rates. Although there are many reported cases which do not easily fit this association, it appears to be the conclusion, not only of this analysis but also of other investigators, that doctors with larger lists typically cope by having fewer surgery and home consultations per thousand patients on their lists than doctors with smaller lists. This conclusion is also consistent with the opinion, noted above, that GPs tend to have fixed methods of working, irrespective of the numbers of patients on their lists, which they can either impose by virtue of the rationing mechanisms which they control or which they are permitted to adopt by virtue of the demand patterns existing in their practices.

The fourth question was whether the proportion of patients consulting their GPs each year differs according to the size of their doctors' lists. The meagre evidence that is available, drawn mainly from a fairly dated national survey, suggests that patient consulting rates are inversely related to list size: that is, the more patients a doctor has on his list, the lower will be the proportion of patients whom he sees each year. A very crude calculation from the national survey indicates that about a quarter of the difference between the consultation rates of doctors with fewer than 1,500 patients and those with 3,000 or more patients may be due to the lower patient consulting rates among the latter, with the residue of the difference stemming from the lower average number of consultations per patient consulting.

The fifth question was whether the ratio of doctor-initiated to patient-initiated differs with list size. As with information about patient consulting rates, the evidence on this is sparse, but data from national and regional studies, as well as from individual practices in different parts of the country, fail to indicate any clear association between the two variables. If it is indeed the case that doctors with larger lists are just as likely to initiate follow-up consultations as those with smaller lists, it follows that the lower levels of demand that are clearly evident in larger than in smaller practices reflect the behaviour of patients rather than doctors. Other studies support this conclusion.

The sixth question was whether the content of care differs according to the number of patients for whom GPs are responsible. The importance was noted of distinguishing between rates based on the number of patients at risk and those based on the number of patients consulting. The analysis is necessarily confined to those aspects of care for which data are normally collected; it is thus a very selected analysis. With respect to the referral of patients for specialist care or opinion, the evidence from a large number of studies leads consistently to the conclusion that, whatever factors may be associated with the very wide range in the referral rates of individual practitioners, list size is not among them. Opinion is divided about the effect of list size on the use of hospital diagnostic facilities. Two studies, separated in time and space, discounted any such effect, and circumstantial evidence from other investigations supports this conclusion. Against this, a larger number of studies have provided evidence of a negative association between the two variables.

Several studies enable prescribing rates to be linked to list size in various ways. Investigations carried out in the 1950s, and one in the 1960s, pointed to a positive association between the mean list sizes and the mean prescribing rates in different geographical areas, but more recent surveys, including a national survey of people in Great Britain, have reported either no association, or a negative association, between the two variables. There are, however, difficulties in interpreting the evidence that are not resolved by incompatibilities in the definition of terms and in the base upon which prescribing rates are calculated. A small amount of information has been published about the issuing of sickness certificates and the use of non-drug treatments in general practice, but insufficient to draw any firm conclusions about the effect of list size on these components of care. There is a suggestion from one national and one regional study that the use of non-drug treatments may be positively associated with size of list.

The seventh question was whether the quality of care differs according to the number of patients for whom GPs are responsible. A firm answer cannot be given until a degree of uniformity has been reached in the definition and measurement of the quality of care. Although some interesting attempts have been made to do this, there does not yet exist an adequate consensus of opinion, nor fund of data, to permit more than the sketchiest of impressions. The division of care into component parts of structure, process and outcome has influenced much of the literature on the quality of care, and a similar framework is adopted here.

The structure of care comprises what Donabedian calls 'the settings in which it takes place and the instrumentalities of which it is the product.' The qualitative evaluation of different caring structures is itself problematic, for it cannot be done rationally without reference to the processes and outcomes of care that flow from different structures. The tendency in the literature for qualitative judgements about caring structures to be made without reference to processes and outcomes invalidates any firm conclusions about the influence of list size on this aspect of quality. However, there is evidence from three national studies of general practitioners that structural features which are generally regarded as 'good' or 'desirable' are more commonly associated with larger than with smaller lists.

Studies of the process of care likewise suffer from the absence of qualitative evaluations without which the 'good' processes cannot be distinguished from the 'bad'. The literature contains evidence that GPs with larger lists subjectively feel the quality of their care to be lower than do those with smaller lists, but such studies fail to specify the ways in which the quality of the caring process is impaired, nor do they indicate the mechanisms that link list size with the process of care. Various attempts have been made by GPs to audit (that is, to make critical evaluations of) their own methods of care. Though inherently valuable, such attempts are of little help to the problem in hand because they have concentrated on quite limited aspects of the process of care; they have been concerned almost exclusively with the care given to individual patients rather than to practice populations; they have not utilised comparable definitions of good quality; and they have not been sufficiently numerous to make sensible inferences about the influence of list size on the quality of the caring process.

The outcome of care, though widely regarded as the most valid indicator of the quality of care, presents equally complex problems in its measurement. The elusiveness of valid outcome measures is emphasised particularly in general practice, where many of the conditions that are seen will improve spontaneously

whatever remedial measures are applied. Nevertheless, attempts have been made to define the outcome variables of general practice, and some studies have endeavoured to apply them to real practices. Two types of studies that are of particular interest by virtue of their focus upon the links that might mediate between list size and outcome are those that have examined the outcome of consultations of different lengths, and those that have assessed the outcome of care for practice populations (for example, by estimating the amount of untreated morbidity existing in different sized practices).

The dimension of outcome which has been investigated most intensively is that of patients' evaluations of the care they receive. One such study, mounted deliberately in a practice that consciously uses the primary health care team to enable the doctor to accept an uncommonly large list, found a high overall level of satisfaction, although pockets of dissatisfaction were felt about certain aspects of care directly attributable to the chosen style of practice. However, the value of these results is diminished by the universally high esteem in which patients hold most aspects of the care they receive from their GPs. More sensitive methodologies for assessing patients' opinions may need to be developed before the significance of list size in shaping the outcome of care can fully be understood.

SUMMARY AND CONCLUSIONS

This report has been prepared in response to the question posed by the Department of Health and Social Security: what is a reasonable list size for a general medical practitioner to provide an adequate level of care? It was shown in the introduction to the report that the question carries important implications for various aspects of health service policy, and that it has been posed (in one form or another) by many prestigious committees, commissions, groups and individuals. There appears to be a widespread and serious interest in exploring the question. However, it was also argued early in this report that, in the form in which it is currently phrased, the question is problematic because it fails to specify a context in which a valid answer might be given, and therefore does little to clarify the kind of information that might be relevant in formulating an answer. The Royal Commission on the National Health Service, for example, recommended that before a maximum or minimum list size for GPs is adopted, considerable further research should be undertaken; but it is not unequivocally clear what kind of research would contribute in a relevant way to the elucidation of the question. Much would appear to depend upon the way in which the sense or meaning of the question is interpreted.

The report tackled the problem of the meaning of 'a reasonable list size' by examining the grounds on which the advocacy of a target GP/patient ratio has been based in a variety of post-war reports, documents and academic commentaries. The actual ratios that have been regarded as 'reasonable' or 'desirable' have (at least in the case of general practice) tended to cluster within a fairly narrow range around the actual national ratio, but the bases of these views have been varied. References are documented in the report to the salience of such factors as the needs of the population for medical care, the standard or quality of care that is delivered, the costs of care, the workload placed upon the GP, and so on. The report drew two general conclusions from this profusion of opinion. First, it was concluded that no single list size could be identified as inherently more 'correct' or 'reasonable' than any other. To search for a reasonable list size in the expectation that, once discovered, it will instantly and universally be recognised as such, is futile. Second, it was concluded that no rational case can be made for viewing any one approach to the definition of a reasonable list size as inherently superior to any other. It is as valid for one group to emphasise the heavy demands that are made upon doctors with lists of 2,500 patients or more as it is for another group to draw attention to the costs of producing and servicing enough new doctors to secure significant reductions in average list sizes.

If these two conclusions are valid, it follows that the choice about a 'reasonable' or a 'maximum' or a 'minimum' list for general practitioners must

represent a judgement about the most appropriate course of action in any given context. The choice necessarily involves the exercise of discretion; it cannot be eliminated or made redundant by the result of a definitive investigation purporting to demonstrate a uniquely 'correct' list size. The choice may, however, be aided, informed, and perhaps made more rational through the results of research, and it is here that the concept of 'reasonableness' must be articulated in order to identify the kind of information that will be relevant in aiding, and perhaps enhancing the rationality of, the judgement.

The evidence reviewed in the report indicates that a core theme in the arguments of those who have tackled the issue is that of the consequences of variations in list size. Those who support a reduction in list sizes have tended to emphasise the undesirable or unacceptable consequences of large lists, and have claimed that such consequences would be eliminated, or at least ameliorated, by improving the doctor/patient ratio. Conversely, those (much fewer in number) who argue the case for larger lists have drawn attention to the limitations of smaller lists and have pointed to the favourable consequences that would flow from increasing the number of patients on a doctor's list. It was noted in the report, however, that claims about the anticipated consequences of increasing or reducing the list sizes of GPs have been based more on supposition than on verifiable evidence, and that one way in which research might contribute to the choice of a reasonable list size is by documenting more thoroughly than has yet been done the range of consequences that are sensitive to variations in doctor/patient ratios.

The logic of the argument, however, suggests that this alone is not sufficient, for even if it could be demonstrated that, say, the quality of care in general practice bore an inverse relationship to list size (as claimed by the British Medical Association), a choice would still be required about the point at which further gains in quality are considered not to be worth the further reductions in list size that would be necessary to achieve them. A key concept here is that of cost, for reductions in the national average list size incur costs that must be weighed against the value or benefit attached to the consequences flowing from such reductions. Ultimately, therefore, the choice about a reasonable list size must be a policy judgement weighing the anticipated costs or savings from changes in list sizes against the value that is placed upon the consequences of such changes. The choice inevitably involves the exercise of value judgements, but it can be informed, and perhaps made more rational, by the availability of good information about the likely costs and consequences of variations in doctor/patient ratios.

The report has tried to assemble information relevant to the choice by reviewing the evidence in the research literature about variations in practice

style, content and outcome that typically occur between GPs with lists of different sizes. There are many pitfalls in this kind of exercise. Apart from obvious problems surrounding the quality, typicality and comparability of data assembled from many separate studies, deeper structural problems also exist. For example, the ultimate uniqueness of each GP and each practice context raises serious doubts about the validity of generalised statements, and the complex inter-relationships of the distinctive features of different practices render hazardous any conclusions about the causal influences of list size on each measure of 'consequence'. Moreover, the review is necessarily restricted to those aspects of practice that have been studied and reported, and it is therefore weighted quite heavily towards items that are easily quantified at the expense of other (perhaps more important) items that are much less easily encapsulated in statistical containers. It is a wholly valid criticism of this section of the report that it projects a rather narrow, mechanistic view of general practice.

The principal dimensions of 'consequence' reviewed from the research literature were: the total time spent by general practitioners in patient care and the average time spent on each consultation, the pattern of consultations, the content of care, and the quality of care. The time variables do not appear to be influenced to any significant extent by list sizes: doctors with larger lists might be expected to spend more time on the job or to spend less time on average with each patient than those with smaller lists, but this is not generally the case. General practitioners seem to develop a distinctive tempo of work regardless of the number of patients for whom they are responsible. There are, however, fairly clear indications that both surgery and home consultation rates are inversely associated with list size; that is, the more patients a doctor has on his list, the lower is his consultation rate per thousand patients. Most of the difference in consultation rates between GPs with large and small lists is due to the lower average number of consultations per patient consulting, not to the lower proportion of patients consulting, and it appears that these variations in consultation patterns reflect the behaviour of the patients rather than the doctors. Referrals for specialist care or opinion are not related to list size, but there is some evidence that doctors with larger lists are less likely than those with smaller lists to use hospital diagnostic facilities and more likely to carry out various medical and surgical procedures. The data in the literature about the quality of care in general practice are insufficient to permit many sensible conclusions to be drawn about its relationship to list size, except to note that, whilst GPs with larger lists express more frustration and dissatisfaction with the quality of their work than those with smaller lists, patients in large practices do not appear on the

available evidence to be any more or less dissatisfied with the care they receive than those in smaller practices.

The broad impression created by this review of the research literature is that the consequences of variations in list size are less evident in the things that GPs actually do than in the number and selection of patients to whom they do them. It appears that general practitioners tend to develop individual practice styles, including the length of time they typically spend in consultation with different categories of patients, that remain fairly impervious to the differing pressures of demand from different sizes of lists. Much more sensitive to variations in list size is the rate at which doctors see their patients, for the strongest single finding emerging from the review of the research literature was the inverse relationship between list size and consultation rates. It follows, then, that the dimension of quality that is threatened by large lists may be less the quality of care to individual patients and more the quality of care to practice populations by virtue of the reduced access that patients in larger practices have to their doctors. Some work has been carried out on this aspect of quality, especially epidemiological studies of the amount of morbidity in practice populations that is known and unknown to the doctors concerned, but most attempts to examine the quality and outcome of care have focused upon events occurring in the GP's surgery, not in the community. This should not surprise us, for it is consistent with the traditional concern of the doctor with the specific patient who seeks his care, rather than the collectivity of patients who may potentially need care but who have not yet claimed the doctor's attention. One danger of this is that by basing judgements about a reasonable list size upon evidence about the consequences flowing from a change in doctor/population ratios, conservative practices will be perpetuated and the scope for radically creative thinking about new roles for general practitioners will be diminished.

Future research

By setting up a systematic framework for reviewing the research literature in general practice, this report has identified a host of gaps and deficiencies in the current state of knowledge, some of which may be worth plugging with further investigations. The aim in this concluding section is simply to summarise the variety of research tasks that emerge from the analysis. No attempt is made to present them as formal research proposals, and no assumptions are made about their feasibility. The extent to which they can be implemented will depend in large part upon the willingness of GPs to participate in collaborative research.

First, the review of the research literature in this report emphasises how

relatively little is known about the work of the average general practitioner. The great majority of studies included in the review are of self-selected practices, and whilst it is not known how representative these practices are of all practices, there are obvious grounds for doubting their typicality. Provided they collect the right sort of data, studies making much more extensive use of randomly selected practices would be of great value in indicating the degree of generalisability of existing data, and in isolating the effects of list size from the effects of other factors which influence the output of services. Multicentre studies may be necessary to generate a sufficient number of subjects.

Second, there is a glaring lack of experimental and quasi-experimental studies in general practice. Such studies are, of course, difficult to set up, but more effort might usefully be expended in this direction. It would, for example, be profitable to study the effects, through the use of appropriate control s, of planned changes in list size, extensions in consultation lengths and the introduction of teams of variable composition.

Turning now to the substance of future investigations, it would be interesting to cull the views of representative groups of patients and GPs about the role which they feel general practice should be fulfilling in the latter quarter of the twentieth century. We may have been too pessimistic in our opinion that insufficient consensus exists at this level of strategic thinking to inform the choices of policy-makers. It would, for example, be interesting to know the reactions of representative groups of doctors to the vision set out in the report of the Northumberland LMC (1978), and whether they concur with the resource implications contained therein.

Considerable gaps still remain in our knowledge about the effects of reducing the size of GPs' lists. Studies aimed at filling some of these gaps would have to go beyond the conventional measures of consultation rates and patterns to the development of sensitive indicators of changes in the amount and nature of demand, in the content and style of the doctor's work, and in the effectiveness of it to the patient. The object of such studies would be to substantiate and fill out the claim of many doctors that smaller lists are a necessary prerequisite to better standards of care. An experimental design for such studies would be an excellent way of setting about them. An alternative approach, suggested earlier in the report, would be through the judgements of doctors about the additional services and benefits they feel they could provide if they had fewer patients and more time. Studies of this kind might take the form of a conventional survey of a representative group of practitioners in order to elicit their generalised judgements on such matters, or they may be combined with conventional workload studies and invite the participating doctors to make such judgements about each one of a series of patients under review.

An important conclusion emerging from the review of published research data was that the major threats to the quality of care arising from over-large lists may be those directed at whole practice populations rather than at individual patients. Some work has been done of relevance to this, but not much. Important questions remain unanswered about the characteristics of those people in larger practices who 'miss out' on access to their doctors relative to those in smaller practices. What kinds of health care needs are left untreated, what other sources of help are elicited, and what are the consequences to the sufferer? Epidemiological studies and surveys of patient satisfaction may be appropriate investigative tools here.

Finally, questions about the efficiency of the structure and organisation of general practice have been raised at several points in the research. The general point has been made that judgements about the balance between costs and benefits should ideally be made in contexts of optimal efficiency; if they are not, the possibility exists of increasing the benefits with no corresponding increase in costs by improving efficiency. Questions of efficiency are relevant to many aspects of general practice, not merely the immediate issue in hand. One focus around which such questions have clustered in this report is that of the practice team. It is clear that the sharing of tasks between different team members is both technically feasible and acceptable to patients, but the evidence points to considerable variety in the size and compositions of teams, and in the way they are used. In some cases, practice and community nurses accept work which, in their absence, would be done by the doctor; in other cases they seem to uncover new work that would not have surfaced but for their presence. In the former context, the efficiency of the doctor is enhanced by transferring tasks that can be performed adequately by less highly skilled staff, and if the time released in this way is used to extend the doctor's care to more patient, there may well be a reduction in the notional average cost of care per patient per year. In the latter context, the doctor's efficiency remains unchanged, and the introduction of the nurse results in additional costs for additional work done. Hence the notional average cost of care per patient per year will increase, but the standard of care will also rise. Questions such as these about the consequence for efficiency and standards of different team structures do not appear to have been tackled very extensively in the literature, and they may constitute another fruitful area for further investigation.

REFERENCES

Acheson, H.W.K. (1975). 'Medical audit and general practice', The Lancet, 1, 511 - 513

Ashworth, H.W. (1965). 'Sickness absence and return to work', Proceedings of the Royal Society of Medicine, 58, 825 - 826

Backett, E.M., Summer, G., Kilpatrick, J. and Dingwall-Fordyce, I. (1966). 'Hospitals in the North-East Scotland region', in Problems and progress in medical care, second series (ed. G. McLachlan), London: Nuffield Provincial Hospitals Trust

Bain, D.G.J. and Haines, A.J. (1974). 'Workload in general practice', British Medical Journal, 1, 641

Bain, D.J.G. and Haines, A.J. (1975). 'A year's study of drug prescribing in general practice using computer-assisted records', Journal of the Royal College of General Practitioners, 25, 41 - 48

Baldwin, J.T. (1959). 'Items of service in general practice', British Medical Journal, Supplement, 7 March, 77

Ball, J.G. (1978). 'Workload in general practice'. British Medical Journal, 1, 868 - 870

Barber, J.H. (1971). 'Computer assisted recording in general practice', Journal of the Royal College of General Practitioners, 21, 726 - 736

Berger, J. and Mohr, J. (1967). A fortunate man, London: Allen Lane The Penguin Press

Berkeley, J.S. (1967). 'Reasons for referral to hospital', Journal of the Royal College of General Practitioners, 26, 293 - 296

Berkeley, J.S. and Richardson, I.M. (1973). 'Drug usage in general practice', Journal of the Royal College of General Practitioners, 23, 155 - 161

Bolden, K.J. and Morgan, D.C. (1975). 'Moving to a health centre - the effect on workload and patients', Journal of the Royal College of General Practitioners, 25, 527 - 531

Bridgstock, M. (1976). 'General practitioners' organisation and estimates of their workload', Journal of the Royal College of General Practitioners, 26, Supplement 1, 16 - 24

British Medical Association (1965). 'A charter for the family doctor service', British Medical Journal, Supplement, 13 March, 89

British Medical Journal (1978a). 'Right to maintain professional standards', Supplement, 16 September, 841

British Medical Journal (1978b). 'Maintaining professional standards', Supplement, 5 August, 449

British Medical Journal (1979). 'GP charter working group report', British Medical Journal, 1, 564 - 572

Buchan, I.C. and Richardson, I.M. (1973). Time study of consultations in general practice, Scottish Health Service Studies, 27, Edinburgh: Scottish Home and Health Department.

Butler, J.R., Bevan, J.M. and Taylor, R.C. (1973). Family doctors and public policy, London: Routledge and Kegan Paul

Buxton, M.J. and Klein, R.E. (1979). 'Population characteristics and the distribution of general medical practitioners', British Medical Journal, 1, 463 - 466

Carmichael, L. and Stevenson, J.S.K. (1963). 'Why are they waiting?' British Medical Journal, 1, 736 - 738

Carne, S. (1969). 'Sickness absence certification. Analysis of one group practice in 1967', British Medical Journal, 1, 147 - 149

Carstairs, V. and Skrimshire, A. (1968). 'The provision of outpatient care at health centres: a review of data available for planning', Health Bulletin, 26, 12-22

Cartwright, A. (1967). Patients and their doctors, London: Routledge and Kegan Paul

Cartwright, A. and Anderson, R. (1979). Patients and their doctors 1977, Occasional Papers no. 8, London: Royal College of General Practitioners

Cartwright, A. and Scott, R. (1961). 'The Work of a nurse employed in a general practice', British Medical Journal, 1, 807 - 810

Central Health Services Council (1954). Report of the Committee on General Practice Within the National Health Service (Cohen Report), London: HMSO

Central Health Services Council (1957). Report of the Committee to consider the future numbers of medical practitioners and the appropriate intake of medical students (Willink Report), London: HMSO

Central Health Services Council (1963). Report of the Sub-Committee on the field of work of the family doctor (Gillie Report), London: HMSO

Central Health Services Council (1971). Report of a Sub-Committee on the organisation of group practice (Harvard Davis Report), London: HMSO

Christoffel, T. and Loewenthal, M. (1977). 'Evaluating the quality of ambulatory care', Medical Care, 15, 877 - 897

Clute, K.F. (1963). A study of medical education and practice in Ontario and Nova Scotia, Toronto: University of Toronto Press

Cobb, J.S. and Baldwin, J.A. (1976). 'Consultation patterns in a general practice', Journal of the Royal College of General Practitioners, 26, 599 - 609

College of General Practitioners (1965). Present state and future needs, Reports from General Practice No.2, London: College of General Practitioners

Collings, J.S. (1950). 'General practice in England today: a reconnaissance', Lancet, 1, 555 - 585

Cooper, M.H. (1975). Rationing health care, London: Croom Helm

Crombie, D.L. and Cross, K.W. (1964). 'The workload in general practice', The Lancet, 2, 354 - 356

Crombie, D.L., Pinsent, R.J.F.H., Lambert, P.M. and Birch, D. (1975). 'Comparison of the first and second National Morbidity Surveys', Journal of the Royal College of General Practitioners, 25, 874 - 878

Culyer, A.J. (1976). Need and the National Health Service, London: Martin Robinson

Culyer, A.J., Lavers, R.J. and Williams, A. (1971). 'Social indicators - health', in Social Trends No.2, 1971, Central Statistical Office, London: HMSO

Culyer, A.J., Lavers, R.J. and Williams, A. (1972). Health indicators', in Social indicators and social policy, (eds. A. Shonfield and S. Shaw), London: Heinemann Educational Books

Davies, B.P. (1977). 'Needs and outputs', in Foundations of Social Administration, (ed. H. Heisler), London: The Macmillan Press Ltd.

Dawes, K.S. and Cottrell, G.R. (1964). 'General practice in Scotland', British Medical Journal, 2, 314

Department of Health and Social Security (1976). Sharing resources for health in England, London: HMSO

Department of Health and Social Security (1976). Priorities for Health and Personal Social Services in England, London: HMSO

Department of Health and Social Security (1977). Health and personal social services statistics for England, 1977, London: HMSO

Department of Health and Social Security (1978). Medical manpower - the next twenty years, mimeographed

Donabedian, A. (1966) 'Evaluating the quality of medical care', Milbank Memorial Fund Quarterly, 3, part 2, 166 - 203

Doney, B.J. (1976). 'An audit of the care of diabetics in a group practice, Journal of the Royal College of General Practitioners, 26, 734 - 742

Duncan, R.M. and Orcharton, A.M. (1964). Quoted in Stevenson, J.S.K. (1964) op.cit.

Dunlop, D.M., Inch, R.S., and Paul, J. (1953). 'A survey of prescribing in Scotland in 1951', British Medical Journal, 1, 694 - 697

Dunnell, K. and Cartwright, A. (1972). Medicine takers, prescribers and hoarders, London: Routledge and Kegan Paul

Eimerl, T.S. and Laidlaw, A.J. (1969). A handbook for research in general practice, Edinburgh: E. and S. Livingstone

Eimerl, T.S. and Pearson, R.J.C.(1966). 'Working time in general practice: how general practitioners use their time', British Medical Journal, 2, 1549 - 1554

Eimerl, T.S. and Pearson, R.J.C. (1968), 'Tools for the job', Journal of the Royal College of General Practitioners, 15, 447 - 452

European Economic Community (1974). The General Practitioner in Europe, Report of the working party appointed by the Second European Conference on the Teaching of General Practice, The Netherlands: Leeuwenhorst

Evans, E.O. and McBride, K. (1968). 'Hospital usage by a group practice', Journal of the Royal College of General Practitioners, 16, 294 - 306

Fein, R. (1967). The doctor shortage: an economic diagnosis, Washington DC: The Brookings Institution

Floyd, C.B. and Livesey, A. (1975). 'Self-observation in general practice - the bleep method', Journal of the Royal College of General Practitioners, 25, 425 - 431

Forsyth, G. and Logan, R.F.L. (1960). The demand for medical care, London: Nuffield Provincial Hospitals Trust

Forsyth, G. and Logan, R.F.L. (1968). Gateway or dividing line? London: Nuffield Provincial Hospitals Trust

Fraser, R.C., Patterson, H.R. and Peacock, E. (1974). 'Referrals to hospital in an East Midlands city', Journal of the Royal College of General Practitioners, 24, 304 - 319

Fry, J. (1952). 'A year of general practice: a study in morbidity', British Medical Journal, 2, 249 - 252

Fry, J. (1972). 'Twenty-one years of general practice: changing patterns', Journal of the Royal College of General Practitioners, 22, 521 - 528

Garraway, W.M. (1973). 'The community health team in general practice', Update, 6, 121 - 132

General Register Office (1958). Morbidity statistics from general practice, Studies on Medical and Population Subjects No. 14, London: HMSO

Gillie, A. (1966). 'Quality in general practice', Journal of the Royal College of General Practitioners, 11, Supplement 3, 1 - 8

Ginzberg, E. (1975). 'Notes on evaluating the quality of medical care', The New England Journal of Medicine, 292, 366 - 368

Gray, P.G. and Cartwright, A. (1954). 'Who gets the medicine?', Applied Statistics, 3, 19 - 28

Green, R.H. (1973). 'General practitioners and open-access pathology services', Journal of the Royal College of General Practitioners, 23, 316 - 325

Grene, J.D. and Johnson, R.D. (1971). 'General practice morbidity monitoring', Journal of the Royal College of General Practitioners, 21, 417 - 421

Grossmark, F.B. and Sharer, P. (1967). 'A study of certification in general practice', The Practitioner, 199, 354 - 355

Gruer, R., Gunn, A.A. and Ruxton, A.M. (1977). 'Medical audit in practice', British Medical Journal, 1, 957 - 958

Hadfield, S.J. (1953). 'A field survey of general practice 1951-2', British Medical Journal, 2, 683 - 706

Hale, G. and Roberts, N. (1974). A doctor in practice, London: Routledge and Kegan Paul

Handfield-Jones, R.P.C. (1964). 'Who shall help the doctor?, Lancet, 2, 1173 - 1174

Hardman, R.H. (1965). 'Morbidity in two areas', Journal of the Royal College of General Practitioners, 9, 226 - 240

Hodgkin, G.K. (1973). 'Evaluating the doctor's work', Journal of the Royal College of General Practitioners, 23, 759 - 767

Hopkins, P. (1956). 'Referrals in general practice', British Medical Journal, 2, 873 - 877

Howie, J.R.G. (1977). 'Patterns of work', in Trends in general practice 1977 (ed. J. Fry), London: Royal College of General Practitioners

Hughes, A.O., Skinner, J.L., Miller, D.S., Skinner, E.M., Enoch, P.J., Varnam, M.A. and Shevlin, A.B. (1976). 'Estimates and projections of workload in general practice', British Journal of Hospital Medicine, 7, 63 - 69

Hull, F.M. (1968). 'Diagnostic pathways in rural general practice', Journal of the Royal College of General Practitioners, 18, 148 - 155

Hunt, J.H. (1972). 'The renaissance of general practice', Journal of the Royal College of General Practice, 22, Supplement 4, 5 -

Irvine, D. (1976). 'Contemporary professional practice', in A question of quality, (ed. G. McLachlan), London: Nuffield Provincial Hospitals Trust

Journal of the Royal College of General Practitioners (1972). Editorial: how many patients?, 22, 491 - 493

Journal of the Royal College of General Practitioners (1977). Editorial: how many patients?, 27, 3 - 7

Journal of the Royal College of General Practitioners (1978a). 'Practice activity analysis 3, investigation', 28, 60 - 62

Journal of the Royal College of General Practitioners (1978b). 'Practice activity analysis 5, referrals to specialists', 28, 252 - 253

Joyce, C.R.B., Last, J.M. and Weatherall, D.M. (1967). 'Personal factors as a cause of differences in prescribing by general practitioners', British Journal of Preventive and Social Medicine, 21, 170 - 177

Judek, S. (1964). Medical manpower in Canada, Ottawa: Queen's Printer

Kennedy, T.M. (1973). 'Annual reports from general practice, 1971, Langholm', Update, 6, 1285 - 1300

Klarman, H.E. (1969). 'Economic aspects of projecting requirements for health manpower', Journal of Human Resources, 4, 360 - 376

Knox, J.D.E. and Morrell, D.C. (1974). 'Studies of general practice: demand, need and quality', British Medical Bulletin, 30, 209 - 213

Lane, K. (1969). The longest art, London: George Allen and Unwin Ltd.

Last, J.M. (1965). 'Objective measurement of quality in general practice', mimeographed, quoted in Cartwright, A. (1967) op.cit.

Last, J.M. (1966). 'The Iceberg: completing the clinical picture in general practice', Lancet, 2, 28 -

Last, J.M. (1971). 'Measurement of quality of care and medical audits', in Methods of health care evaluation, (eds. D.L. Sackett and M.S. Baskind), Hamilton, Ontario: MacMaster University (mimeographed)

Lave, J.R., Lave, L.B. and Leinhardt, S. (1975). 'Medical manpower models: need, demand and supply', Inquiry, 12, 97 - 125

Lavers, R.J. (1978). 'The morale of NHS stsff', in Economic aspects of health services, (eds. A.J. Culyer and K.G. Wright), London: Martin Robertson

Lee, R.I. and Jones, L.W. (1933). The fundamentals of good medical care, Chicago: University of Chicago Press

Lees, D.S. and Cooper, M.H. (1963). 'The work of the general practitioner', Journal of the College of General Practitioners, 6, 408 - 435

Logan, R.F.L. (1964). 'Studies in the spectrum of medical care', in Problems and progress in medical care (ed. G. McLachlan), London: Nuffield Provincial Hospitals Trust

McCready, P.E. and Reilly, P.M. (1977). 'Audit of a general practice', Journal of the Royal College of General Practitioners, 27, 527 - 530

MacDonald, A. and McLean, I.G. (1971). 'Study of the work of general practitioners', The Practitioner, 207, 680 - 688

McGregor, R.M. (1973). 'The work of a family doctor', Journal of the Royal College of General Practitioners, 23, 787 - 791

Maddox, G.L. (1971). 'Muddling through: planning for health care in England', Medical Care, 9, 439 - 448

Mair, A. and Mair, G.B. (1959). 'Five-year study of a general practice', British Medical Journal, Supplement, 20 June, 281 - 284

Mansfield, P. (1973). 'The study and evaluation of general practice', Journal of the Royal College of General Practitioners, 23, 887 - 894.

Marsh, G.N. (1968). 'Visiting - falling workload in general practice' British Medical Journal, 1, 633 - 635

Marsh, G.N. and Kaim-Caudle, P. (1976). Team care in general practice, London: Croom Helm

Marsh, G.N. and McNay, R.A. (1974a). 'Team workload in an English general practice', British Medical Journal, 1, 315 - 318

Marsh, G.N. and McNay, R.A. (1974b). 'Factors affecting workload in general practice', British Medical Journal, 1, 319 - 321

Marson, W.S., Morrell, D.C., Watkins, C.J., and Zander, L.I. (1973). 'Measuring the quality of general practice', Journal of the Royal College of General Practitioners 23, 23 - 31

Martin, J.P. (1957). Social aspects of prescribing, London: William Heinemann Ltd.

Maynard, A. and Walker, A. (1978). 'Medical manpower planning in Britain: a critical appraisal', in Economic aspects of health services, (eds. A.J. Culyer and K.G. Wright), London: Martin Robertson

Mechanic, D. (1974). Politics, medicine and social science, New York: John Wiley and Sons

Mechanic, D. (1977). 'The growth of medical technology and bureaucracy: implications for medical care', Milbank Memorial Fund Quarterly, Health and Society, 55, 61-77

Medical Practices Committee (1949). First Report, duplicated

Medical Practices Committee (1951). Third Report, duplicated

Medical Practices Committee (1954). Sixth Report, duplicated

Migue, J.L. and Belanger, G. (1974). The price of health, Toronto: The Macmillan Company of Canada

Morrell, D.C. (1971). 'Expressions of morbidity in general practice', British Medical Journal, 2, 454 - 458

Morrell, D.C., Gage, H.G. and Robinson, N.A. (1970). 'Patterns of demand in general practice', Journal of the Royal College of General Practitioners, 19, 331 - 341

Morrell, D.C. and Nicholson, S. (1974). 'Measuring the results of changes in the method of delivering primary care - a cautionary tale', Journal of the Royal College of General Practitioners, 24, 111 - 118

Mourin, K. (1976). 'Auditing and evaluation in general practice', Journal of the Royal College of General Practitioners, 26, 726 - 733

Noble, E.J. (1973). 'Practice workload', Update, 7, 989 - 991

Noble, J.S. (1978). 'An expanding remit for general practice', British Medical Journal, Supplement, 10 June, 1966

Northumberland Local Medical Committee (1978). Charter 2: a new general practitioner service, mimeographed

Office of Health Economics (1965). Work lost through sickness, London: OHE

Office of Health Economics (1974). The work of primary medical care, London: OHE

Office of Population Censuses and Surveys and the Royal College of General Practitioners (1974). Morbidity statistics from general practice, Studies on Medical and Population Subjects No.26, London: HMSO

Office of Population Censuses and Surveys (1978). The General Household Survey 1976, London: HMSO

Paige, D. and Jones, K. (1966). Health and Welfare Services in Britain in 1975, Cambridge: The University Press

Peterson, O.L., Andrews, L.P., Spain, R.S. and Greenberg, B.G. (1956). 'An analytical study of North Carolina general practice', Journal of Medical Education, 31, 1-165

Posner, B.H. (1965). 'The pattern of work in an industrial practice', The Practitioner, 195, 799 - 803

Price, D.T. (1973). 'A survey of the use made of hospitals by a provincial suburban practice', Update, September, 617 - 625

Price, D.T. (1974). 'Teamwork in general practice', British Medical Journal, 1, 520

Rafferty, J. (ed.) (1974). Health manpower and productivity, Lexington, Mass.: Lexington Books

Reilly, P.M. and Patten, M.P. (1978). 'An audit of prescribing by peer review', Journal of the Royal College of General Practitioners, 28, 525 - 530

Reinhardt, U.E. (1975). Physician productivity and the demand for health manpower, Cambridge, Mass.: Ballinger Publishing Company

Review Body on Doctors' and Dentists' Remuneration (1966). Seventh Report, Cmnd.2992, London: HMSO

Review Body on Doctors' and Dentists' Remuneration (1968). Ninth Report, Cmnd.3600, London: HMSO

Review Body on Doctors' and Dentists' Remuneration (1970). Twelfth Report, Cmnd. 4352, London: HMSO

Review Body on Doctors' and Dentists' Remuneration (1972). 1972 Report, Cmnd.5010, London: HMSO

Review Body on Doctors' and Dentists' Remuneration (1975). Fifth Report, Cmnd.6032, London:HMSO

Review Body on Doctors' and Dentists' Remuneration (1977). Seventh Report, Cmnd.6800, London: HMSO

Review Body on Doctors' and Dentists' Remuneration (1978). Eighth Report, Cmnd.7176, London: HMSO

Richardson, I.M., Howie, J.G.R., Durno, D., Gill, G. and Dingwall-Fordyce, I. (1973). 'A study of general practitioner consultations in North-East Scotland, Journal of the Royal College of General Practitioners, 23, 132-142

Rider, J.G., Waters, W.H.R., Charnock, R.B. and Lunn, J.E. (1969). 'A study of workload in a mining area practice', Journal of the Royal College of General Practitioners, 17, 361 - 369

Rose, H. and Abel-Smith, B. (1972). Doctors, patients and pathology. Occasional Papers on Social Administration No.49, London: G. Bell and Sons

Rosser, R.M. and Watts, V.C. (1972). 'The measurement of hospital output', International Journal of Epidemiology, 4, 361 - 368

Royal College of General Practitioners (1968). The practice nurse, Reports from General Practice No.10, London: Royal College of General Practitioners

Royal College of General Practitioners (1969). 'The educational needs of the future general practitioner', Journal of the Royal College of General Practitioners, 18, 358 - 360

Royal College of General Practitioners (1970). Present state and future needs of general practice, Reports from General Practice No.13, London: Journal of the Royal College of General Practitioners

Royal College of General Practitioners (1973). Present state and future needs of general practice, Reports from General Practice No.16, London: Journal of the Royal College of General Practitioners

Royal College of General Practitioners (1977). Trends in general practice 1977 (ed. J. Fry), London: Royal College of General Practitioners

Royal Commission on Medical Education (1968). Report (Todd Report), Cmnd.3569, London: HMSO

Royal Commission on the National Health Service (1979). Report (Merrison Report), Cmnd.7615, London: HMSO

Schonfield, H.K., Heston, J.F. and Falk, I.S. (1972). 'Numbers of Physicians required for primary medical care', New England Journal of Medicine, 286, 571 - 576

Scott, R. (1964). 'The family doctor diagnostic centre', in Problems and progress in medical care (ed. G. McLachlan), London: Nuffield Provincial Hospitals Trust

Scott, R. and Gilmore, M. (1966). 'The Edinburgh hospitals', in Problems and progress in medical care, second series (ed. G. McLachlan), London: Nuffield Provincial Hospitals Trust

Scott, R. and McVie, D.H. (1962). 'Doctor in the house', Journal of the College of General Practitioners, 5, 72 - 85

Shannon, R. (1968). 'Manpower planning in the National Health Service', in Problems and progress in medical care, third series, London: Oxford University Press for the Nuffield Provincial Hospitals Trust

Sheps, M.C. (1955). 'Approaches to the quality of hospital care', Public Health Reports, 70, 877 - 886

South-East Thames Regional Health Authority (1974). 'Rate of hospital discharges, average length of stay, and GP list sizes and hospital activity in the S.E. Thames RHA', unpublished report, Management Services Division

Starey, C.J.H. (1961). 'A hospital outpatient referral survey', Journal of the College of General Practitioners, 4, 214 - 222

Steen, P. (1967). 'Some aspects of group practice', The Practitioner, 199, 810 - 813

Stevens, J.L. (1977). 'Quality of care in general practice: can it be assessed?', Journal of the Royal College of General Practitioners, 27, 455 - 466

Stevens, R. (1971). American medicine and the public interest, New Haven: Yale University Press

Stevenson, J.S.K. (1964). 'General practice in Scotland - why the difference?', British Medical Journal, 1, 1370 - 1373

Taylor, P. (1974). 'Sickness absence: facts and misconceptions, Journal of the Royal College of Physicians, 8, 315 - 334

Taylor, R.J., Howie, J.G.R., Brodie, J. and Porter, I.A. (1975). Use of bacteriological investigations by general practitioners', British Medical Journal, 3, 635 - 636

Taylor, S. (1954). Good general practice, London: Oxford University Press

Thomas, K.B. (1978). 'Time and the consultation in general practice', British Medical Journal, 2, 1000

Tudor Hart, J. (1971). 'The inverse care law', Lancet, 1, 405 - 412

Tudor Hart, J. (1976). 'General practice workload, needs and resources in the National Health Service', Journal of the Royal College of General Practitioners, 26, 885 - 892

Watts, C.A.H. (1952). 'A year of general practice', British Medical Journal, Supplement, 30th August, 115

Westcott, R. (1977). 'The length of consultations in general practice', Journal of the Royal College of General Practitioners, 27, 552 - 555

Weston Smith, J. and O'Donovan, J.B. (1970). 'A survey of the workload in a general practice', Journal of the Royal College of General Practitioners, 20, 202 211

Whitfield, M.J. (1973). 'A study of prescribing in general practice 1969-70', Journal of the Royal College of General Practitioners, 23, 168-182

Williams, A. (1974a). 'The cost-benefit approach', British Medical Bulletin, 30, 252 - 256

Williams, A. (1974b). 'Measuring the effectiveness of health care systems', British Journal of Preventive and Social Medicine, 28, 196 - 202