

CANCER REGISTRATION IN THE STUDY OF BREAST AND LUNG CANCER^a

J. Clemmesen

INTRODUCTION

Hospital-based cancer registration has in the past contributed to collection of treatment results, follow-up, and other fundamental clinical activities. Since the Second World War adequate diagnostic and therapeutic facilities have become available to all social groups in many countries, enabling population-based registers to be established with the intention of covering *all* cases (and only those) arising in a particular region. Ideally, cancer registration should be nationwide but experience has shown that it is difficult to establish linkage in populations greater than 5–10 million.

Uniformity in certain respects eases the problem of linkage and hence Denmark, with its relatively uniform population of 5 million living in a limited, and geographically uniform, area with good medical facilities for all social classes, is particularly fortunate. There, it is possible to organize and follow up screening campaigns and surveillance of environmental risks on a nationwide basis and the results of registration can be translated into changes in patient care and contribute to future planning.

BREAST CANCER

Methods and Results

Breast cancer is probably the best example of a naturally occurring malignant neoplasm to illustrate how registration and relatively unsophisticated statistical methods contribute to the efforts against cancer. It was useful findings from genetic studies that were the main means of justifying, within a limited period, the setting up of the Danish Cancer Registry.

^a Example based on Clemmesen (1, 2).

Genetics. In collaboration with the University of Copenhagen Institute of Human Genetics Jacobsen (3) collected family histories for 197 female and 3 male breast cancer patients known from the registry, as well as corresponding information for control families. The study resulted in a clear confirmation of the few older demonstrations of a hereditary tendency for the development of breast cancer.

Age distribution. Furthermore, Jacobsen's data showed also a slight bimodal trend in the age distribution of breast cancer, which was confirmed by analysing data from the register which at that time was still limited. Clemmesen (4) analysed data for a longer period and found a small peak in the age curve for incidence of breast cancer at 47 years with a small depression between 47 and 52 years. This "hook" distribution coincides with the culmination of cervical cancer incidence and with the average age of menopause, which at that time in Denmark was 48 years. Hence, the irregularity in the age distribution of breast cancer could be ascribed to hormonal changes at the time of menopause.

The irregularity became obvious also in data from other countries and regions when subdivided into 5-year age groups, as is now customary in cancer statistics (1, 2). Other explanations than hormonal changes have been suggested for the "hook", but its occurrence in numerous situations suggests that its absence indicates deficiency of the underlying data (5).

It was already clear from the original analysis that the "hook" *irregularity* in the curve depends on only few cases. Inaccuracy in age given or pronounced variation in the delay before women seek medical attention will therefore tend to obscure the appearance of the menopause "hook". This seemed to be the case for data from the Upstate New York Registry (4) although the "hook" was quite obvious elsewhere in the USA (6).

There is a small, but not statistically significant, difference in the incidence of breast cancer between married and unmarried premenopausal women. The former have a slightly higher incidence and in Denmark this is more pronounced in the capital than in rural areas, where the "hook" is more pronounced (1, 2) (see Fig. 1). Since breast cancer is considerably more frequent in unmarried women after the menopause, the two age curves must cross at menopausal age logically with changes taking place on the married women's curve.

Jacobsen's data enabled Busk (7) to show that laterality of breast cancer may be genetically determined, i.e., relatives of women with affected right breasts are more likely to be similarly affected while the general risk is slightly higher for the left side. The differences may result simply from a hereditary asymmetry in the amount of gland tissue in the breasts and, despite some authors' beliefs (8), they cannot be accepted as evidence of the genetic role in susceptibility to breast cancer.

When these results had been attained it appeared that similar observations about the age distribution and the laterality of breast cancer could be made from data collected in the nineteenth century (9, 10), although on a less solid basis than registration (Fig. 2).

Fig. 1. Age distribution of incidence rates for breast cancer of women in Copenhagen and rural areas of Denmark, 1953–57, according to marital status

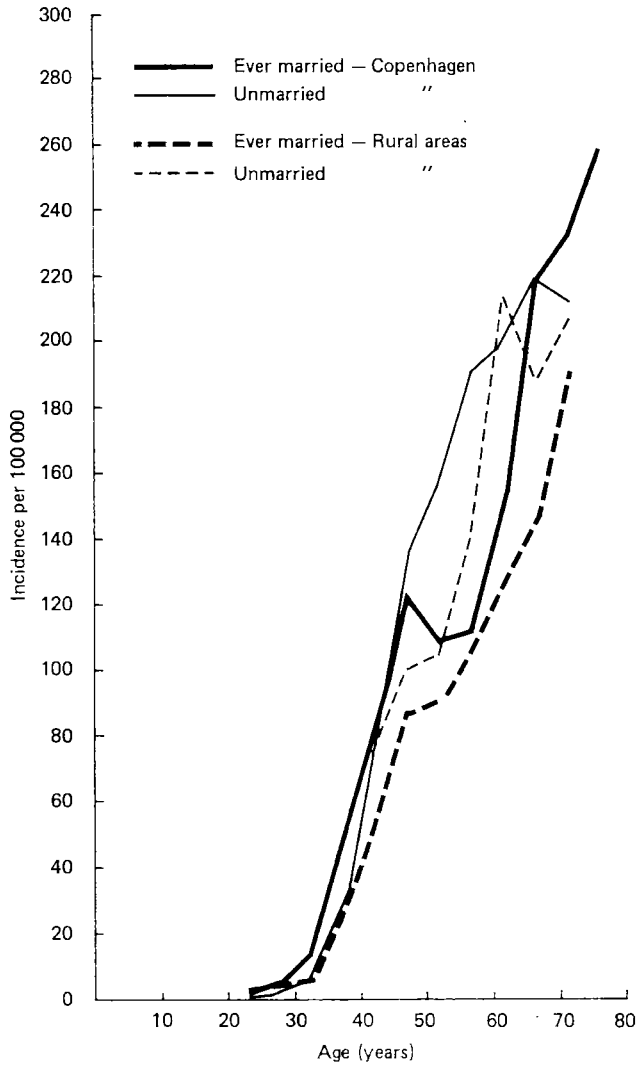
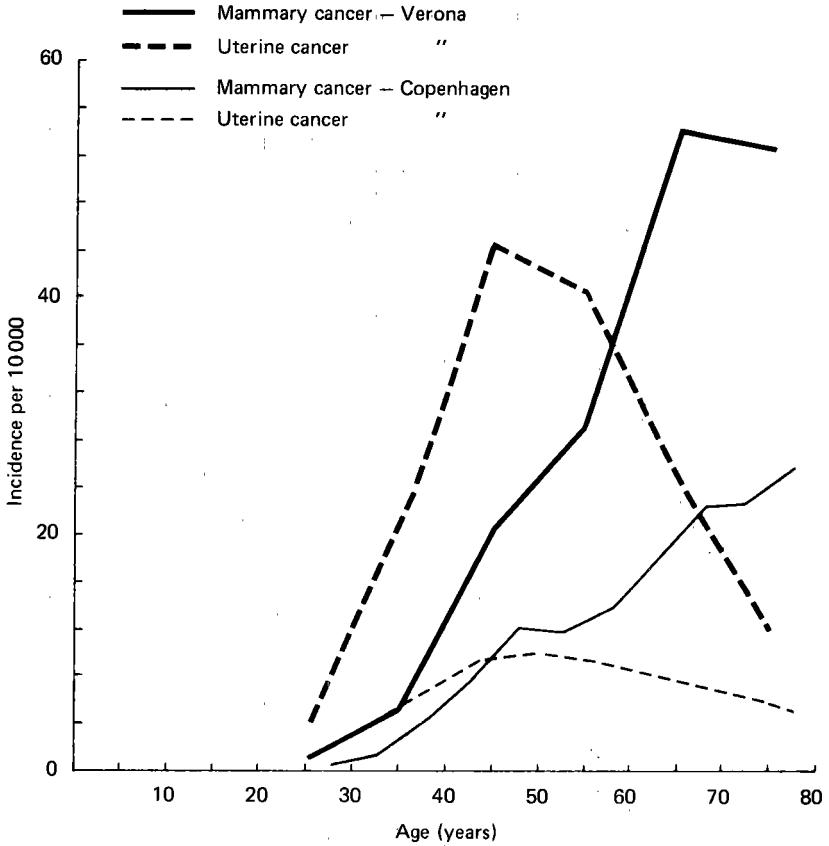


Fig. 2. Age distribution of mortality rates for breast and uterine cancer in women, Verona, 1760–1839,^a compared with incidence rates for breast and cervical cancer in women in Copenhagen, 1943–57^b



^a Data from Stern (9, 10).

^b Data from Danish Cancer Registry (1, 2).

Applications

The use of registration data suggested improvements in statistical techniques and encouraged the study of hormonal aspects of breast cancer.

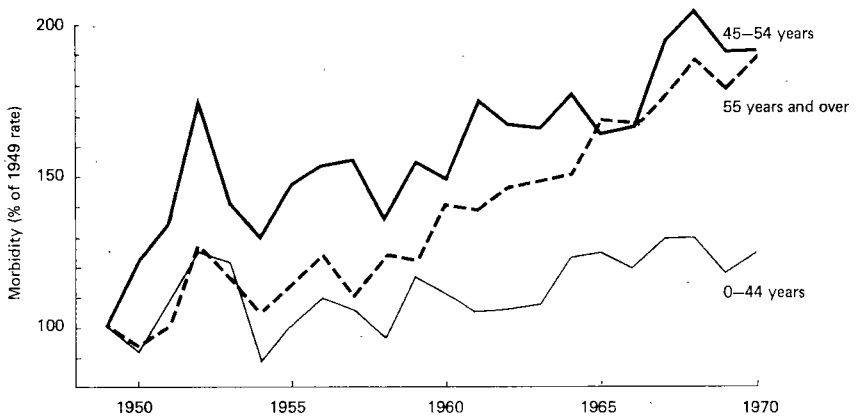
Self-examination campaign

From 1951 to 1954 a campaign to encourage women to examine their breasts was carried out in Denmark under the auspices of the National Cancer Society. Women were encouraged to seek medical advice and accept the need

for histological examination when breast nodules were found. The campaign was planned with a view to later statistical evaluation and some years later, when there was increasing scepticism about the effectiveness of cancer campaigns, it was possible to provide evidence of improved prognosis for breast cancer as a result of this campaign (1).

A successful campaign for early diagnosis will first result in an increase in the number of cases diagnosed, followed later by a corresponding reduction. This applied to the present campaign and it was also found that the resulting improvement in 7-year survival rates was limited to women under 55 years; survivals for older women remained the same (Fig. 3). However, a rise in incidence of breast cancer in women aged 45 years and over resulted in an overall reduction in survival rate because of the poorer prognosis for these age groups.

Fig. 3. Incidence of breast cancer in Danish women at various ages, 1949–70, as percentages of the 1949 rates



Survival estimates

Cancer registration has also been useful in the overall estimate of survival rates for a region or a country. Such estimates will give a long-term measure of the overall efficiency of a treatment system, while an exact evaluation of some specific form of therapy usually requires uniform and specified information from one or a small group of uniform centres.

On the basis of the gross staging possible from reports to the cancer registry, it has been possible in Denmark (1960–66) to compare relative survival rates for women treated for breast cancer by surgery with surgery followed by radiotherapy. Comparisons were made for cases described as stage I (isolated tumours only) and for stage II (with lymph node metastases).

Stage III showed too few cases for analysis. The cases in each stage were subdivided according to age below 55 years and 55 years and over (Fig. 4).

It appeared from 5-year survival rates that it made no difference to cases in stage I whether surgery was followed by radiotherapy or not. For stage II cases it was found that radiotherapy following surgery did improve survival rates; the results are in accord with three earlier studies (11-13). In practice, this means that about 600 radiotherapy treatments per year might be used for other purposes and that this number might increase if more cases could be diagnosed while still in stage I.

A further prognosis of practical use will, however, not be possible, because, as illustrated in Fig. 3, the incidence of breast cancer will depend on the age distribution of the population at the given time together with marital status and age at first delivery.

LUNG CANCER

The age pattern of lung cancer is quite different from that of breast cancer. Like cancer of the bladder and pancreas, the increase in incidence appears related to subsequent birth group cohorts (Fig. 5). In 1953 this observation was used by Clemmesen et al. (14) to calculate the future incidence of lung cancer among men in Copenhagen, although based on the number of deaths that at the time was very close to the total. It was assumed that the age curve for the 1905 birth cohort would represent the highest incidence of lung cancer, which seems to have been justified (Fig. 5). Furthermore, it was assumed that prevention and the general age distribution would remain unchanged to 1990, and so the age curve for the youngest cohort could be drawn parallel to that of the oldest.

Table 1 shows a comparison of projected incidence with the actual numbers from the cancer registry, and the projections seem to have been fairly accurate if total cases are counted. It should be mentioned, however, that changes in age distribution and in smoking habits will have influenced rates.

Applications

Many younger physicians who do not remember when cancer frequency could only be vaguely assumed from a general cause of death such as "abdominal cancer" will tend to take for granted the results from cancer registration, and administrators planning services for cancer patients may well use the data from cancer registries with little idea of how that information was compiled.

For cancer registration to be of use in the study of causation it must be maintained over long periods since exposures to certain occupational agents during 20, or even 40, years may be involved. This long-term aspect is particularly relevant to preservation of personal files on employees in present and

Fig. 4. Five-year survivals of Danish women treated for breast cancer in 1960-66 by means of surgery and surgery combined with radiotherapy, as crude and relative survival rates

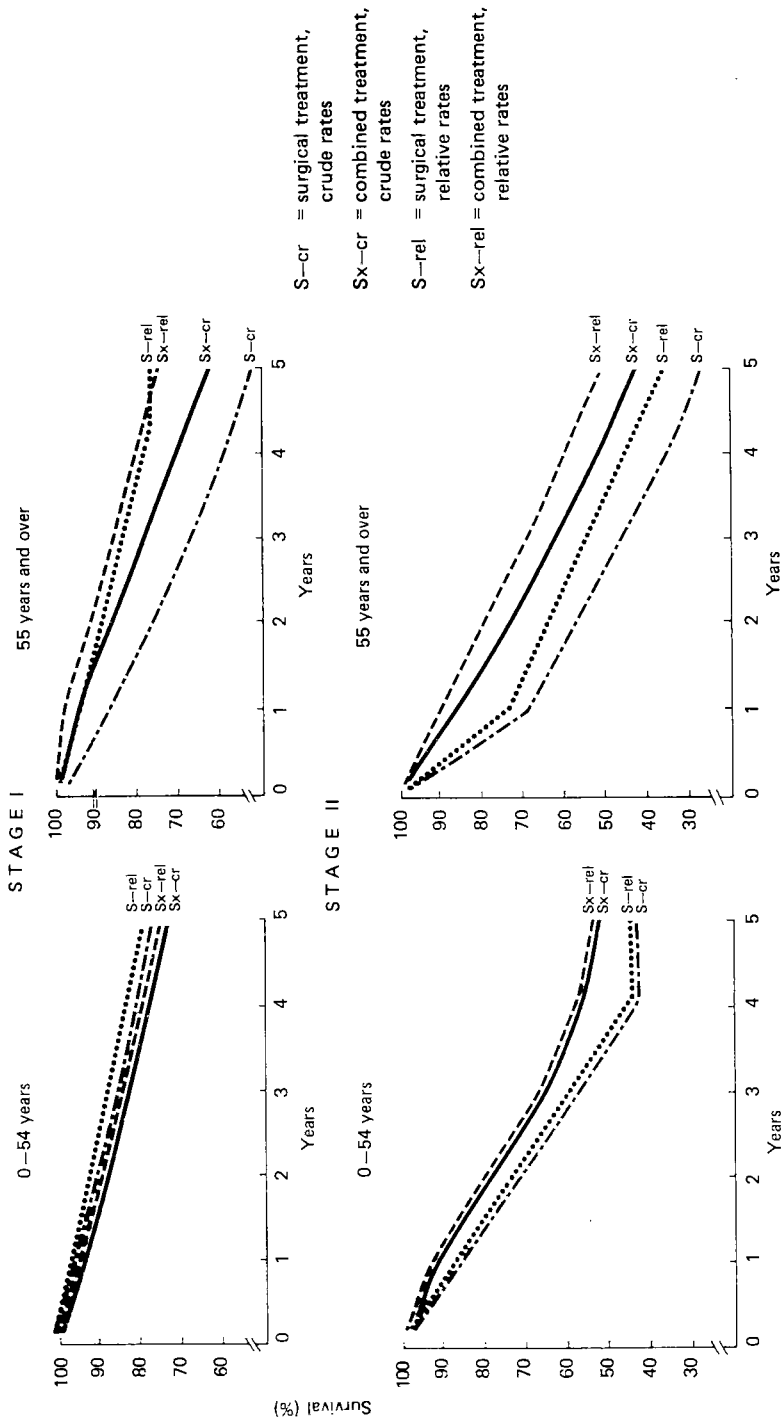
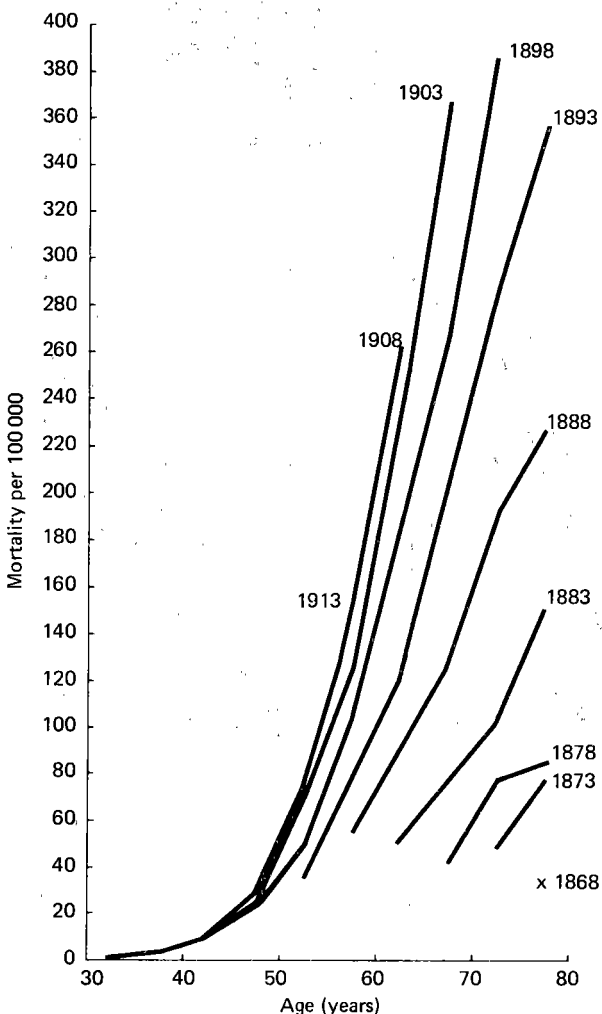


Fig. 5. Age distribution of incidence rates of lung cancer in Danish men, according to 5-year birth groups, 1943–72



past occupations. With ever more potentially harmful agents being used in industry, there is increasing need for monitoring of all cases in a population, and only by rational use of computers will this be feasible. A combination of cancer registration with computerized employee registration for all industries is therefore being prepared in Denmark.

Such long-term policies have been lost sight of in various places. Because the *International statistical classification of diseases, injuries, and causes of*

Table 1. Prognosis for the occurrence of lung cancer among men in Greater Copenhagen^a

Deaths expected		Cases registered	
Period	No.	Year	No.
1951–55	241	1951	235
		1952	236
		1953	229
		1954	262
		1955	277
1956–60	362	1956	341
		1957	347
		1958	358
		1959	438
		1960	407
1961–65	507	1961	480
		1962	405
		1963	426
		1964	464
		1965	500
1966–70	660	1966	480
		1967	557
		1968	599
		1969	590
		1970	562
1971–75	808	1971	540
		1972	600
1976–80	919		
1981–85	979		
1986–90	1 007		

^a Calculated in 1953 from a prolongation of the age distribution curve for deaths among the cohort born in 1905, as deaths expected under the assumption of a constant age distribution of the population at risk.

death is revised every 10 years, it cannot fail to influence the accuracy of statistics of cancer incidence over several decades, which are necessary for appropriate conclusions. The three volumes on *Cancer incidence in five continents (15–17)* show how statistics over several decades can enable us to draw relevant conclusions, provided that revisions in nomenclature and coding do not affect comparability over the years.

For the first time since 1728, when cancer registration was first attempted, the final goal may at last be within reach if the prerequisites can be assured. If this is achieved, cancer registration may soon be taken as a matter of course.

REFERENCES

1. Clemmesen, J. Statistical studies in malignant neoplasms I. *Acta Pathologica et Microbiologica Scandinavica*, supplement, 174 (1965).
2. Clemmesen, J. Statistical studies in malignant neoplasms V. *Acta Pathologica et Microbiologica Scandinavica*, supplement, 261 (1977).
3. Jacobsen, O. *Heredity in breast cancer: a genetic and clinical study of two hundred probands* (thesis). Copenhagen, Busck; London, H.K. Lewis, 1946.
4. Clemmesen, J. Carcinoma of the breast: results from statistical research. *British journal of radiology*, 21: 583–590 (1948).
5. MacMahon, B. et al. Etiology of human breast cancer. *Journal of the National Cancer Institute*, 50: 21–42 (1973).
6. Dorn, H. & Cutler, S. Morbidity from cancer in the United States. *Public health reports (Wash.)*, 29 (1955).
7. Busk, T. Some observations on heredity in breast cancer and leukemia. *Annals of eugenics*, 14: 213–229 (1948).
8. Petrakis, N.L. Genetic factors in the etiology of breast cancer. *Cancer (Philadelphia)*, 39: 2709–2715 (1977).
9. Stern, R. Fatti statistici relativi alla malattie cancerose [Statistical facts relating to cancerous diseases]. *Giornale per servire al progressi della patologia e della terapeutica* [Diary designed to assist in the progress of pathology and therapeutics], 2: 507–517 (1842).
10. Stern, R. Nota sulle ricerche del dottor Tanchou intorno la frequenza del cancro [Note on Dr Tanchou's research on the occurrence of cancer]. *Annali Universali di Medicina*, 110: 484–503 (1844).
11. Fischer, B. Postoperative radiotherapy in the treatment of breast cancer. *Annals of surgery*, 172: 711–732 (1970).
12. Høst, H. & Brennhovd, I.G. Combined surgery and radiation therapy versus surgery alone in primary mammary carcinoma. *Acta Radiologica (Stockholm)*, 14: 25–32 (1975).
13. *British medical journal*, 1: 1035–1038 (1976) (Report on an international multicentre trial supported by the Cancer Research Campaign).
14. Clemmesen, J. et al. Mortality and incidence of cancer of the lung. *Acta Union Internationale Contre le Cancer*, 9: 603–636 (1953).
15. Doll, R. et al., ed. *Cancer incidence in five continents. I*. Heidelberg, Springer, 1966 (on behalf of the International Union Against Cancer).
16. Doll, R. et al., ed. *Cancer incidence in five continents. II*. Heidelberg, Springer, 1970 (on behalf of the International Union Against Cancer).
17. Waterhouse, J. et al., ed. *Cancer incidence in five continents. III*. Lyons, International Agency for Research on Cancer, 1976 (IARC publication No. 15).

AN EVALUATION OF SOCIAL THERAPY IN CHRONIC ALCOHOLISM^a

Joan Collins, A.L. Cochrane, & W.E. Waters

INTRODUCTION

The Seebohm Committee (2) described the personal social services as large-scale experiments in ways of helping those in need. The Committee emphasized that it is both wasteful and irresponsible to set experiments in motion and to omit to record and analyse the results. The evaluation of social work presents many problems and is sometimes thought to be impossible. Often, the social workers' enthusiastic conviction that they are doing good makes evaluation difficult. Each case a social worker undertakes is an experiment since each is unique. This situation is comparable with many encountered in medical practice and can be examined scientifically by means of a randomized clinical trial.

METHODS AND RESULTS

A randomized clinical trial was set up to compare the effectiveness of two approaches to the treatment of patients with chronic alcoholism. The possibility of conducting the trial arose when some members of a unit for the treatment of chronic alcoholics realized that an intensive approach by a team of workers, concerned not only with the patient but also with his family, might be more effective than the usual treatment which sometimes includes referral by a ward staff member to a social worker in either hospital or local authority services. The trial was planned as a pilot to see if the extra time and money that the intensive approach required was likely to improve the effectiveness of treatment.

Cases admitted to a unit for the treatment of chronic alcoholism entered the trial if they (1) lived within 24 km of the hospital, (2) were currently living as a member of one of a number of previously defined social groups, and (3) had not been in the unit in the previous two years. All suitable patients

^a Example based on Collins (1).

were interviewed within 3 days of admission by the research social worker using a schedule which included the patient's version of his personal history, the history of the drinking problem, its magnitude and effects, and his view of how the situation might be improved. If the patients were willing, their families were visited and a similar schedule completed. A narrative report was compiled from these schedules and filed in the ward medical records. The patients were then allocated at random, by opening a sealed envelope, to either the "intensive" or the "traditional" approach.

The patients and families in the intensive approach group were discussed by the team at a succession of weekly meetings. All policy decisions, which were recorded, were team decisions, but one worker acted as agent for the team in dealings with the patient and various organizations.

Whether or not individuals in the traditional approach group were referred for social help was decided by the ward sister. In fact, one man was already on probation, another had a social worker in his family, and two became known to a hospital social worker and two to a local authority social worker. The others were not referred at all.

The research social worker was not involved in this stage, but 2 months after admission to the trial all patients and families who were still in contact with the unit were interviewed again using a schedule that attempted to measure change for the better or worse. It sought verifiable facts about various aspects of the problem, obtained the views of both patient and family about any changes in the situation and set these against an assessment of risk based on the pre-admission history.

Since the research social worker was aware of the grouping there was a possibility of bias; hence an independent assessment of benefit was made in the following week by a physician who was unaware of the details of the two methods being compared and who did not know which method had been used in individual cases. This independent assessment also used a standard questionnaire and a clinical interview for both the patient and the family, seen separately in their homes.

The standard questionnaire concerned attitudes to alcohol, an assessment of the amount of alcohol taken in the previous week, and relationships with the treatment unit and within the family. This was scored on arbitrary scales so that the higher the score the greater the apparent benefit.

The clinical interview was also scored for alcohol intake, patient's insight into his condition, and future outlook; a total score was obtained by addition. An overall clinical impression was also recorded and, before the code was broken, the cases were ranked in what appeared to be a descending order of benefit from treatment.

Intake interviews before randomization showed no important differences between the group that received the intensive social therapy and the remainder (i.e., the traditional approach group). In general, the assessments of the social worker and the independent assessor were similar. The independent assessor saw both the patient and the patient's family in 12 cases but, for various reasons, failed to make a complete evaluation in 8 cases; 4 were treated by each method.

The scores for the independent assessors' standard questionnaire and for the clinical interview are given in Table 1. The overall clinical impression of

Table 1. Treatment scores for standard questionnaire and clinical interview (independent assessor)

Patient number	Group ^a	Standard questionnaire (max. = 44)	Clinical interview (max. = 56)
1	T	35	46
19	In	34	44
9	T	31	43
13	In	27	36
18	In	26	31
17	T	24	27
10	In	17	21
14	T	17	19
20	T	15	22
11	In	13	19
3	In	13	21
4	T	2	3

^a T, traditional treatment; In, intensive approach to social therapy.


benefit is shown in Table 2. In the week immediately preceding the independent assessor's visit, the alcohol consumption in the two groups was similar (Table 3).

The geographical and social conditions imposed in an attempt to reduce the variables between the two groups meant that only about 1 in 5 patients admitted was eligible for consideration and accumulating the series took 6 months. Also, because there were no precedents for this work numerous discussions with various interested parties and a prepilot run to try out methods and gauge patient and staff reaction to an apparently arbitrary distinction between two groups of people needing help were necessary.

APPLICATIONS

This study, although based on small numbers, showed that the randomized clinical trial can be used in evaluating social work. Although obviously not conclusive, it has considerably modified the views of many about the benefits of the collaborative approach to social therapy. The study has much wider implications, which have been discussed elsewhere (1, 3).

Table 2. Overall clinical impression of apparent benefit at independent assessor's interview

Overall clinical assessment	Patient number	Group ^a	
<i>descending to</i> 	9	T	
	1	T	
	19	In	
	13	In	
	3	In	
	20	T	
	11	In	
	18	In	
	10	In	
	17	T	
	14	T	
	Little effect	4	T

^a T, traditional treatment; In, intensive approach to social therapy.

The deficiencies in aftercare have often been attributed to fragmentation of the social services and inadequate liaison between institutions and the community. In this study there was close cooperation between individuals and also between institution and community services, but this led to no obvious increase in benefit to the patients although the social workers felt that they had derived considerable benefit for themselves.

Perhaps the most important conclusion from this study is that attempts such as this to measure the effectiveness of social work by randomized clinical trials are feasible. More specifically, it outlines an approach that will show whether enthusiastic therapists really are doing good, i.e., whether they are, in fact, altering the natural history of the condition in their patients. Since this survey began, in January 1970, Goldberg (4) has reported on the effectiveness

Table 3. Number of days in week before independent assessor's visit on which alcohol was consumed

Number of days	"Traditional" treatment group	"Intensive" approach to social therapy group
0	2	2
1-6	1	1
7	3	3

of social work with elderly clients, using a random allocation method. Use of the randomized clinical trial enables assessment of the effectiveness of social work. As the Seebohm Committee (2) said, to refuse to analyse the benefits of the personal social services "makes no sense in terms of administrative efficiency and, however little intended, it indicates a careless attitude towards human welfare".

REFERENCES

1. **Collins, J.** Group treatment for alcoholism. *British hospital journal*, **81**, part II, p. 1661 (1971).
2. **Committee on Local Authority and Allied Personal Social Services 1965–1968.** *Report*. London, Her Majesty's Stationery Office, 1968 (Cmnd 3703).
3. **Waters, W.E. et al.** Evaluation of social therapy in chronic alcoholism. *British journal of preventive and social medicine*, **26**: 57 (1972).
4. **Goldberg, E.M.** *Helping the aged: a field experiment in social work*. London, Allen & Unwin for the National Institute of Social Work, 1971.

DEFINITION OF HOSPITAL CATCHMENT AREA POPULATION AS DENOMINATOR FOR MORBIDITY INDICATORS DERIVED FROM HOSPITAL DATA IN DEVELOPING COUNTRIES^a

H.J. Diesfeld

INTRODUCTION

Exact measurement of ill health can be achieved only by intensive cross-sectional or longitudinal prevalence studies based on a sound sampling technique. For many reasons, mainly of a technical or financial nature, such surveys can only be carried out in relation to special health problems.

A widely used source of information for retrospective studies is hospital records, but these have to be treated with considerable scepticism and are often thought to be totally useless. However, they may be the only source of data from which basic information on the magnitude of a health problem can be obtained. For mortality and morbidity statistics a functioning diagnostic and recording system is often taken for granted, but in countries with a highly developed and dense network of health services this is often insufficient and it is particularly true of countries with low-density health services.

Frequently, though incorrectly, disease frequencies from hospital records are taken as "prevalence" or "incidence" but these terms should be reserved for completely recorded frequencies in a defined "population at risk"; for hospital statistics this denominator is not known.

Disease frequencies are often presented in the form of rates, using either a dependent denominator such as the total number of cases, or an independent denominator, i.e., an administrative section of a census population. Again there are strong reservations from the statistical point of view.

In many English-speaking African countries annual returns of disease exist for hospital inpatients and outpatients, with a list of main diagnoses according to the ICD (2). They are usually compiled annually for an entire country and published in annual health reports which means that most of the inherent information is lost.

In this paper a method is presented which may help to improve the value of morbidity indicators derived from hospital statistics by taking the inherent systematic and mathematical bias into consideration.

^a Example based on a paper presented at the Sixth International Scientific Meeting of the International Epidemiological Association in Primosten, Yugoslavia, in 1971 (Diesfeld (1)).

THE HOSPITAL RECORDING RATE

The systematic bias of hospital statistics can be divided into an external component and an internal component:

- (1) the *external systematic bias* is due to generally unknown and unmeasurable factors which operate because hospital patients are a heavily selected sample of the total population;
- (2) the *internal systematic bias* operates from when the patient visits the hospital to the time final diagnosis is made and entered in the records.

These two components of systematic bias vary according to the hospital catchment area, hospital size, quality of medical care, and the patient's approach to disease and modern medicine. This makes it extremely difficult to compare hospital records on a regional or national basis, much less on an international basis.

It is suggested that a new concept of a defined *hospital recording rate* (HRR) should be introduced; this indicates the limitations and also defines the type of information conveyed by hospital records.

The HRR is defined as the number of cases of a given diagnosis per time period entered in a hospital record, divided by the defined population of the hospital catchment area. The formula would be:

$$\text{HRR} \quad \lambda = \frac{\sum xi}{ni}$$

where λ indicates the frequency with which a certain disease (x) has been observed in a hospital (i) within a defined population (n) within its catchment area, assuming that only one diagnosis according to the ICD is entered into the hospital return.

The symbol λ is used instead of p for frequency because this kind of frequency distribution is best approximated by a Poisson distribution (3). The frequency with which a diagnosis is made and entered into the hospital records represents a "rare event" compared with the large number of "possible events" in the population of the catchment area. This means that the probability of a certain disease occurring in the population and being entered in the hospital records is comparatively low.

The proportional frequency with which cases from the denominator population are recorded in hospital statistics never exceeds 0.01. Since these data follow a discrete distribution they have to be classified for use in statistics. Classification is best done by ranking hospitals according to the hospital recording rate on a logarithmic scale or in a stanine rank transformation in nine ranks of ascending order. Stanine ranking differentiates rather better than the logarithmic ranking without giving an unjustified impression of accuracy by adhering to relative frequencies (Table 1).

Thus for each hospital and its estimated catchment area a disease pattern can be drawn which can be compared on a national or regional scale, as a measurement of health or ill health.

In contrast to the incorrect use of the terms prevalence or incidence, the term "hospital recording rate" by definition admits that there is an "information

Table 1. Stanine and logarithmic ranking of hospital recording rates

(a) Stanine ranking

1st = 4%	}	of all recording hospitals in an ascending order of disease recording rate per population in catchment area
2nd = 7%		
3rd = 12%		
4th = 17%		
5th = 20%		
6th = 17%		
7th = 12%		
8th = 7%		
9th = 4%		

(b) Logarithmic ranking

0 : i = 0	}	cases per 100 000 population in catchment area
1 : i = 0.1 – 0.9		
2 : i = 1.0 – 9.0		
3 : i = 10.0 – 99.0		
4 : i = 100.0 – 999.0		
5 : i = 1 000.0 – 9 999.0		

gap". It implies that within the hospital catchment area there might well be diseases which are not seen or recorded in hospital because of the external and internal systematic bias.

DEFINITION OF THE HOSPITAL CATCHMENT AREA AND ITS POPULATION

In using the term HRR it is crucial to define the population of the hospital catchment area, i.e., the denominator for the recorded number of cases. Theoretically, the hospital catchment area can be described as a circle (πr^2) with the radius r being the maximum distance people will travel to hospital.

Within the prevailing pattern of communication and travelling in East Africa, where this study was carried out, the theoretical catchment area can be described by three zones (Table 2):

- (1) Zone 1 with an inner radius r_1 of 10 km, which is the distance that people are prepared to walk in East Africa if transport is not available;
- (2) Zone 2 with a middle radius r_2 of 20 km takes the existence of public and private transport or the walking distance of nomadic people into consideration;

Table 2. The theoretical zones of hospital catchment area, East Africa

Radius	$r_1 = 10 \text{ km}$	$r_2 = 20 \text{ km}$	$r_3 = 30 \text{ km}$
Zone-area	$Z_1 = 314.159 \text{ km}^2$	$Z_2 = 942.477 \text{ km}^2$	$Z_3 = 1\,570.795 \text{ km}^2$
Catchment area (C)	$C = Z_1 = 314.159 \text{ km}^2$	$C = Z_1 + Z_2 = 1\,256.636 \text{ km}^2$	$C = Z_1 + Z_2 + Z_3 = 2\,827.431 \text{ km}^2$

(3) Zone 3 with an outer radius r_3 of 30 km describes the maximum distance from which patients may still be able to reach hospital.

Two main factors determine the functional relationship between the hospital and its surrounding population and may modify the theoretical zonal outline of the catchment area. In brief, these are:

- (1) the *attractiveness* of a hospital in terms of size, staffing, equipment and reputation; and
- (2) its *accessibility* in terms of its topographical situation and its position within the network of health services.

Christaller's theory of central places (4) provides a method for the qualitative and semi-quantitative measurement of the functional relations between hospitals within one region by using a scoring system, thus delineating the catchment area of each. Any suitable and generally available criterion that helps to identify and measure the functional qualities of a hospital within its surroundings may be used. The sum total of the scores of each hospital expresses its "degree of centrality" within the network of health services and determines the priority of one hospital over neighbouring ones and forms the basis for description of the functional boundaries of the catchment area.

Hospital services in terms of attractiveness and accessibility are estimated by scoring (from 0 to 4) certain criteria according to their availability in each hospital (Table 3).

Score S1 for attractiveness or hospital size

- S1.1, type of hospital (provincial, district, subdistrict)
- S1.2, number of beds;
- S1.3, availability of diagnostic facilities (laboratory, X-ray facilities, special units);
- S1.4, number of medical staff;
- S1.5, number of paramedical staff.

Score S2 for accessibility of the hospital

- S2.1, number of health centres referring patients to a certain hospital;
- S2.2, distance from the next hospital of higher priority; here the reciprocal score is used, i.e., the greater the distance from the next larger hospital the higher the score;

- S2.3, type of administrative centre at which the hospital is situated (centrality of the location);
 S2.4, position in relation to the road network of the area;
 S2.5, topographical situation (valley, hilltop, plain country, or plateau).

Table 3. Centrality score of 6 hospitals in the East African study region^a

Hospital	S1 Attractiveness						S2 Accessibility						ΣS1 + S2 (degree of centrality)
	.1	.2	.3	.4	.5	S1	.1	.2	.3	.4	.5	S2	
	Hospital type	No. of beds	Diagnostic facilities	No. of physicians	No. of medical personnel	Sum of S1	No. of health centres	Distance from hospital of higher priority	Type of administrative centre	Road network	Topographic situation	Sum of S2	
K09 Nandi Hills	1	1	1	1	2	6	0	1	2	3	3	9	15
K11 Eldoret	2	3	1	3	2	11	2	2	3	3	3	13	24
K16 Nakuru	3	4	3	3	3	16	3	2	3	3	3	14	30
K22 NAIROBI	3	4	3	3	3	16	3	3	3	3	3	15	31
K27 Nyeri	3	4	3	3	3	16	2	2	3	3	2	12	28
K31 Isiolo	2	1	0	0	1	4	0	1	2	1	3	7	11

^a The 6 hospitals have been selected from a total of 50.

The frequency distribution for each of the above criteria contributing to estimation of the degree of centrality was examined and scores given accordingly (Table 4).

Hospitals whose catchment area cannot be separated from each other and that use the same recording system can be pooled with respect to their denominator population "at risk".

In densely populated areas where there is more than one hospital or considerable overlap of hospital catchment areas, it is necessary to define the "effective denominator population" for each hospital within the total potential population of the common catchment area. Assuming that this potential population is able to make use of all the beds and services in the area the total population is multiplied by a factor F , which expresses the relationship between the number of beds in one hospital (B_1) and the total beds in all hospitals ($B_1 + B_2$) in this shared catchment area.

$$F = \frac{B_1}{B_1 - B_2} ; \quad P' = P \times F$$

Table 4. Distribution of marks of the scores among 50 hospitals in the Lake Victoria study area

<i>Score S1 for attractiveness</i>	No. of hospitals	Score S1	<i>Score S2 for accessibility</i>	No. of hospitals	Score S2
<i>Type of hospital</i>			<i>Number of health centres</i>		
provincial general hospital	8	3	10-13	7	3
district hospital	28	2	5-9	13	2
subdistrict hospital	14	1	1-4	14	1
	<u>50</u>		none	16	0
<i>Number of beds</i>			<i>Distance from hospital of next higher priority</i>		
more than 200	12	4	more than 160 km	9	3
101-200	12	3	56-160 km	18	2
51-100	16	2	32-56 km	15	1
up to 50	10	1	up to 32 km	8	0
	<u>50</u>			<u>50</u>	
<i>Diagnostic facilities</i>			<i>Situation in administrative centre</i>		
laboratory, X-ray, and special wards	8	3	at provincial town	12	3
laboratory and X-ray/or	8	2	at district town	27	2
laboratory and special ward	28	1	at subdistrict town	11	1
laboratory only	6	0		<u>50</u>	
nil	<u>50</u>		<i>Situation in road network</i>		
<i>Number of physicians</i>			at junction	32	3
more than 2	16	3	at bifurcation	8	2
2	12	2	at single road	10	3
1	14	1		<u>50</u>	
none	8	0	<i>Topography</i>		
	<u>50</u>		on plateau or plain	26	3
<i>Number of medical and clinical assistants</i>			on a slope	21	2
5 and more	17	3	on hilltop or in valley	3	1
3-4	24	2		<u>50</u>	
1-2	9	1			
	<u>50</u>				

P' then expresses the effective denominator population for each hospital and avoids bias in the total common population when hospitals share a common catchment area.

The variables used in either of these scoring systems are not statistically independent since hospitals are usually planned, or added to, in accordance with the needs and density of the population they serve. But many hospitals do not conform to the needs of their potential catchment areas. There are some discrepancies between the single components for each descriptive score and the total scores, S1 for attractiveness and S2 for accessibility: i.e., in some hospitals the number of available beds and the number of staff are not really related to size of catchment area, or they do not match the degree of centrality of the hospital; there are other hospitals at district level which are not located at the "headquarters" of a district or they may have no physicians; and although a hospital may be inaccessible it may be attractive for other reasons or vice versa.

Natural boundaries, lines of communication and other geographical factors, and links with institutions providing primary health care such as health centres have to be considered when delineating a hospital catchment area.

All these factors influence and modify the theoretical circular catchment areas of hospitals and permit a more refined grading of the functional relations between each hospital and its surroundings, or among different hospitals, than merely categorizing hospitals according to the number of beds, as is commonly done.

Catchment areas of hospitals necessarily overlap. The scoring system will provide some idea of which hospital has priority over a neighbouring one. Assuming that catchment areas were not determined by other factors, catchment areas of hospitals of higher degree of centrality will modify and lessen the catchment areas of hospitals of lower degree of centrality.

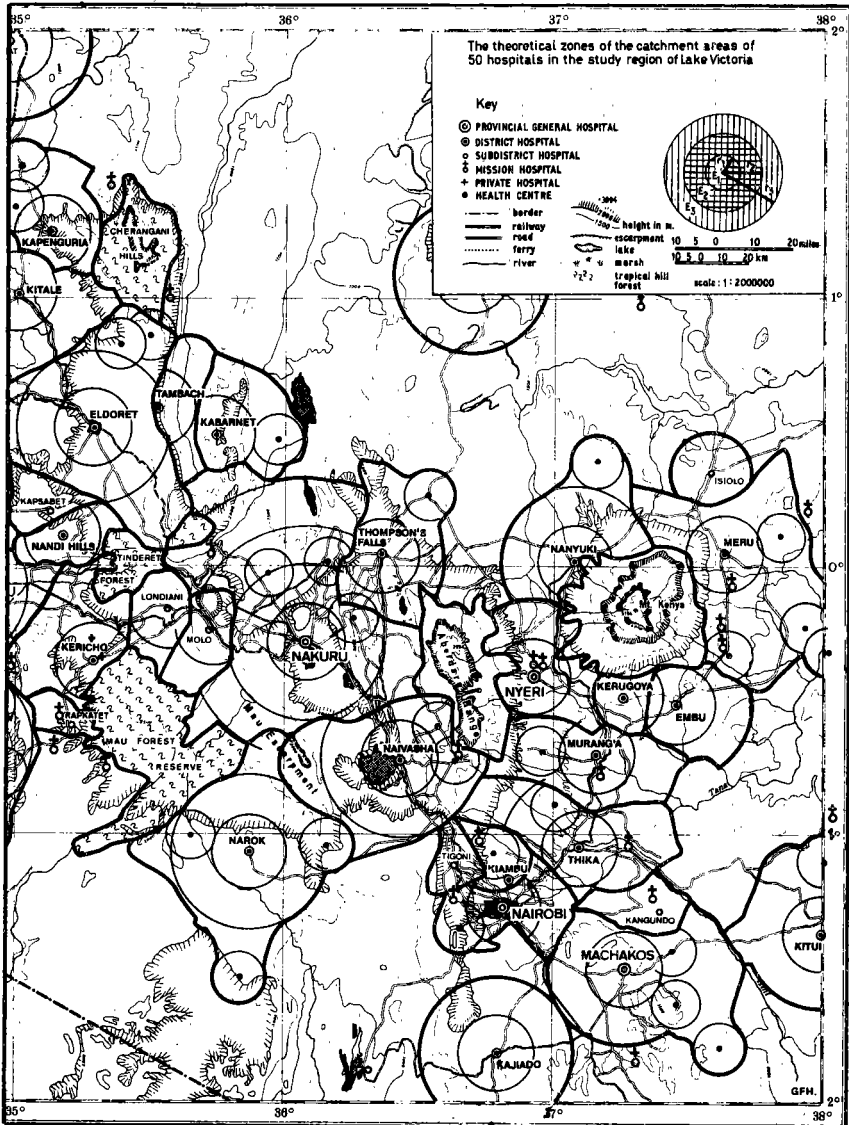
The descriptive part of this analysis is done on a topographical map, preferably on a scale of not less than 1:1 000 000, where the theoretical circular catchment areas of hospitals identified are drawn and modified by the degree of centrality attributed to each hospital and by natural or other types of boundary which can be found on the map (Fig. 1).

THE POPULATION OF THE CATCHMENT AREA

The final, most important, factor is the demographic aspect. Using census data on the smallest possible administrative level and population distribution maps the number of people living in the identified catchment area can be calculated.

The accuracy of this procedure depends on the accuracy of the census data and the maps, and to what extent the defined and drawn catchment area of the hospital is an acceptable approximation of reality. It should be remembered that this semi-quantitative demographic and geometrical approach is but a crude approximation of a denominator population for hospital statistics.

Fig. 1. Catchment areas of hospitals in the Lake Victoria study region East Africa



Reproduced by permission of Oxford University Press.

APPLICATIONS

From the enumeration of the thus defined population of the catchment area, as identified so far, and from the mapping of this area the physiological population density of the catchment area can be calculated; this can also be taken into account when estimating the degree of centrality of a hospital (Table 5).

Having arrived at a "population of the hospital catchment area" as an independent denominator for the recorded number of cases from hospital statistics the HRR for each recorded disease can be calculated.

From these weighted frequencies the disease pattern of a hospital can be calculated. By listing hospitals according to a logarithmic or stanine ranking of the disease frequencies in terms of hospital recording rates the level of ill health, as seen from hospital records, can be estimated.

Table 5. Distribution and classification of population, area, and population density per km² of catchment areas of 50 hospitals in the Lake Victoria study area

S3.1 Population (X 1 000)	No. of hospitals	Score S3.1
below 50	10	1
50-99	10	2
100-199	12	3
200-399	10	4
400-799	5	5
800 and over	3	6
	50	

S3.2 area (km ²)	No. of hospitals	Score S3.2
below 1 000	11	1
1 000-1 999	10	2
2 000-2 999	9	3
3 000-3 999	4	4
4 000-4 999	4	5
5 000 and over	12	6
	50	

S3.3 Population density per km ²	No. of hospitals	Score S3.3
below 30	10	1
30-59	14	2
60-119	12	3
120-149	6	4
150-179	4	5
180 and over	4	6
	50	

This method has been applied recently in the manner described, using models modified from earlier work (5), for a geomedical analysis of Kenya aiming at a more holistic approach (6).

REFERENCES

1. **Diesfeld, H.J.** The definition of the hospital catchment area and its population as a denominator for the evaluation of hospital returns in developing countries. *International journal of epidemiology*, 2: 47–53 (1973).
2. **World Health Organization.** *Manual of the international statistical classification of diseases, injuries, and causes of death, 1975 revision.* Geneva, 1977.
3. **Diesfeld, H.J.** *Biostatistische Auswertung von Krankenhausberichten als Grundlage einer geomedizinischen Analyse der Krankheitsverbreitung in tropischen Entwicklungsländern* [Biostatistical evaluation of hospital reports as the basis for a geomedical analysis of disease distribution in tropical developing countries]. Heidelberg, Habilitationsschrift, 1969.
4. **Christaller, W.** *Die zentralen Orte Süddeutschlands* [Central cities of southern Germany]. Jena, Fischer, 1933.
5. **Diesfeld, H.J.** Evaluation of hospital returns in developing countries. *Methods of information in medicine*, 9: 27–34 (1970).
6. **Diesfeld, H.J. & Hecklan, H.** Kenya, a geomedical monograph. In: Zusatz, K.J., ed. *Geomedical Monograph Series, No. 5.* Berlin & New York, 1978.

**THE USE OF THE CONTROLLED TRIAL TO
MEASURE NEW HEALTH CARE SYSTEMS:
MULTIPHASIC SCREENING AS AN
ADJUNCT TO THE UNITED KINGDOM
NATIONAL HEALTH SERVICE^a**

M.F. D'Souza

INTRODUCTION

Controlled trial experiments have become established as the method of choice in evaluating new drugs and treatments. In this study this methodology has been extended to evaluate the benefits of introducing a whole new health care system.

Multiphasic screening has had numerous advocates, particularly in the USA, and has been introduced into most of the developed countries of the world. This study was a nationally financed project designed to provide a clear answer to government on the value of instituting a national screening programme for the middle-aged as an adjunct to its other services.

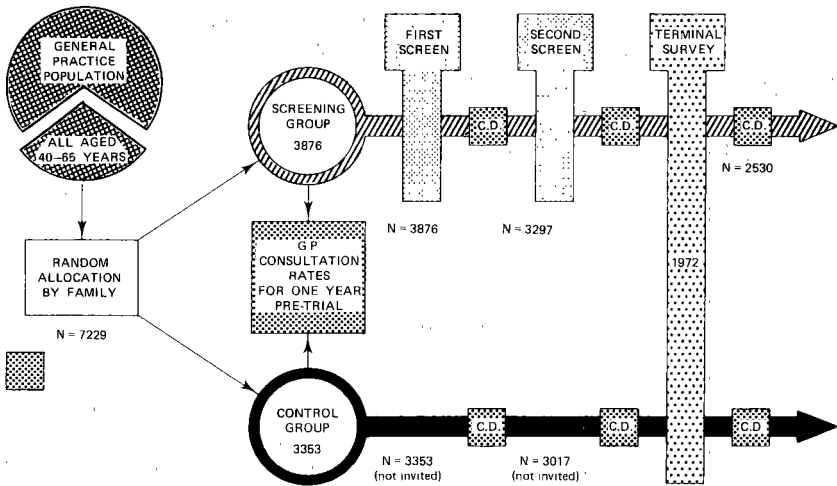
METHODS

The methodology of this study has already been described (2). Essentially, two large group practices in South London volunteered to take part. Using age-sex registers, all those aged 40–64 years in 1967 were identified and were then randomly allocated by family and within general practitioner list into two equal groups designated “control” and “screening” (Fig. 1). The screening group (total 3876) was then invited by personal letter from their general practitioner to be screened. Only 3297 were actually available for screening by the time the study began. The screening clinics operated an appointment system and were held in the evenings in a local infant welfare clinic. They were staffed by nurses and specially trained local housewives, and supervised by a physician. The screening tests performed are shown in Table 1. The initial screening was carried out from 1967 to 1968.

The overall response rate of those attending the screening was 2420 (73.4%). This proportion reflects the response rates from those actually contacted.

^a Example based on South-East London Screening Study Group (1).

Fig. 1. Screening study: overall plan^a



^a N = total number present at a particular point in time *invited* for screening. Changes reflect deaths and departures.

C.D. = consultation data: record of general practitioner consultation and National Health Service use collected every 6 months for both groups.

Reproduced by permission of the publishers of the *International journal of epidemiology*.

Because general practitioner lists are constantly changing, it is difficult to establish accurate denominators at any particular point in time.

Subsequent to the screening, all information was passed on to the appropriate general practitioner, who then made a physical examination of each subject and decided on further tests, diagnoses, and treatments. Although agreement was reached among the physicians as to which abnormalities in the clinical and biochemical tests they would like to have brought to their attention, there was no interference in subsequent management of the screened subject.

The Yield and Management of Disease at Initial Screening

On average 2.3 diseases were found per person screened; 53% of this morbidity was considered to be previously unknown by the general practitioners but 95% of these unknown abnormalities were of a minor nature, being neither disabling nor life-threatening. Of the serious disease discovered by screening 56.3% was already known about and in most instances being treated. Equally important was the observation that for the majority of abnormalities revealed by screening, with the notable exception of anaemia and hypertension, little therapeutic intervention was attempted. All those considered obese, however, were advised to slim and smokers were told to try to give up the habit. In

Table 1. Test used in multiphasic screening procedure

1. Self-administered symptoms questionnaire
2. Interviewer-administered questions on occupational data
3. Anthropometry — height, weight, skinfold thickness
4. Visual testing — near, distant, visual fields ^a
5. Audiometry
6. Chest X-ray
7. Lung function tests — peak expiratory flow rate, FEV ₁ , FVC
8. Electrocardiogram — 12 leads
9. Blood pressure
10. Blood tests — haemoglobin, packed cell volume, blood urea, random blood sugar, protein-bound iodine, serum cholesterol, serum uric acid
11. Stool for occult blood ^b
12. Basic physician examination ^a — skin, mouth, teeth, joints, abdomen for herniae, legs for varicose veins, breast and pelvic examination

^a Not carried out at terminal survey.

^b Only carried out at first screening.

general, considering the effort involved in the screening programme, the quantity and quality of the therapeutic intervention it engendered was disappointingly low, mostly because current medicine is limited in its ability to influence the course of most chronic degenerative diseases revealed by screening.

The “Rescreening” 1969–70

The same group was invited to reattend for screening in 1969–70. The response rate was only 65.6%, somewhat lower than before. Likewise, the yield of disease was lower than in initial screening. For example, 2.1% (50 people) were newly diagnosed hypertensive after the first screening, but only 0.5% (9 people) were newly diagnosed as such after the second screening (3).

The “Evaluation”

These two screenings, and all the interventions directly attributed to them, constitute the “treatment”, which was assessed in this controlled trial. A variety of variables were used in assessing change.

Throughout the study information was gathered every 6 months on all general practitioner consultations, periods of certified sickness absence, and hospitalizations. All deaths and departures from the study have also been carefully recorded in both the screening and control populations. For these morbidity and mortality variables complete information is available for control treatment comparison up to the time of death or departure. The difficulty of following up people who have left the area meant that no information was available about either the morbidity or mortality experience of the 20% from both control and screening populations who left during the course of the study.

Because of the high migration rate in these two, essentially suburban, practices, a further method of evaluation was adopted. This was intended to bring the study to a close before the residual populations became too small a proportion of the original totals to be representative.

This concluding evaluation took the form of a health survey of both the screening and the control groups. Essentially, this was a repeat of the previous screening procedures so that direct comparisons could be made of variables such as blood pressure levels between the screened population and the control group. However, this being a survey rather than a screening, efforts were made to encourage those who had refused the initial invitation to attend the clinic or at least give some health information at home. In one practice these efforts were exhaustive and involved home visiting of all those who did not attend the clinics. From this a response rate of 87% was achieved. In the other practice the general practitioners were concerned that such survey work might be a nuisance to their patients and only gave permission for repeated follow-up letters to be sent to those who refused implicitly or explicitly the initial invitation to attend the clinic. The response rate in this practice was only 74.9%.

RESULTS

Mortality

The mortality experienced by the screening and control populations is shown in Table 2. As can be seen, there are no significant differences in overall or specific death rates. More detailed analysis using survival curves has also failed to reveal any differences between these populations.

Sickness absence

There were no systematic differences in certified sickness absence between control and screening populations. There were large differences between men and women, but only because many women were not entitled to sickness certification.

Table 2. Death rates by cause: control versus screening, 1967-75

Cause of death ^a (first certified)	Control group (N = 3 132) (Total time at risk, 18 404.4 man/years)		Screening group (N = 3 292) (Total time at risk, 19 672.3 man/years) ^b	
	No. died	Death rate per 1 000 man/years at risk	No. died	Death rate per 1 000 man/years at risk
Neoplasms:				
ICD codes: 140-229 with 519.2 except 149, 166-169, 182-189, and 208-209	47	2.6	50	2.5
Central nervous system:				
ICD codes: 330-398 with 780-781.7 except 335-339, 346-349, and 358-359	13	0.7	17	0.9
Cardiovascular disease:				
ICD codes: 400-468 with 782.0-785.3 except 417-419, 423-429, 435-449, 454-459, and 784	52	2.8	84	4.3
Respiratory disease:				
ICD codes: 003.1 and 241, 470-527, with 782.3-783.7 except 476-479, 484-489, 494-499, 503-506, 508-509, and 527.1	37	2.0	28	1.4
All other causes	20	1.1	17	0.9
Total deaths (all causes)	169	9.2	196	10.0

^a According to the 1955 revision of the International Classification of Diseases (ICD) (5).

^b Time at risk is less than for previous analyses owing to delays in ascertaining cause of death.

Hospitalizations

Total hospital admissions for the screening and control populations were analysed by age and sex and also by percentage of admissions for specific disease groups. Again, no significant or systematic differences were observed which might suggest that screening reduces utilization of hospital services. If anything, the opposite was likely to be true (Table 3).

Consultation Rates

Higher overall consultation rates were observed in the screening population but the difference was not statistically significant ($t = 1.29$).

Direct Comparisons at Survey

The concluding health survey enabled direct morbidity comparisons to be made between the screening and control groups; no significant differences in either the health questionnaire or clinical tests were discovered. Some of the measurements used are illustrated in Table 4. A detailed analysis was undertaken of all outcome variables using a multifactor analysis taking into account age, sex, social class, smoking habit, blood pressure, blood sugar, serum cholesterol, and general practice group. These potential nuisance variables were analysed using the general linear model programme (GLIM) (4). However, even with such detailed scrutiny no consistent differences could be observed between the screening and control groups.

Costs

The crude average cost of this screening has been estimated retrospectively excluding adjustments for alternative uses of capital. The figure at 1968 prices came to only £5950 per 1000 persons screened, that is, a little under £6 per person. This figure at 1976 prices is calculated to have increased by over 100% and the best estimate of a similar multiphasic screening now would be £13 748 per 1000 persons screened.

Overall Findings

None of the measures of mortality, morbidity, or health service use has been shown to be improved by screening. However, the screening service itself was generally well received by the population to which it was offered, and this was reflected in the repeated response rates of over 70% of those invited. Most of the physicians taking part in the study were enthusiastic about screening at first but over time their attitudes changed and most of them felt that the procedures were onerous and usually unrewarding.

Table 3. Hospital admissions: control versus screening, 1967–76

	Control group (N = 3 132)		Screening group (N = 3 292)	
<i>(a) Number of admissions</i>				
Number of people admitted once or more, 1967–76		862		944
Rate per 1 000 man/years at risk ^a		49.6		50.7
Total number of admissions ^a per 1 000 man/years at risk		70.7		73.4
<i>(b) Hospital admissions by some principal diagnoses</i>				
Principal diagnosis at admission ^b	Control group		Screening group	
	Persons admitted once or more	Admission rate ^a per 1 000 man/years at risk	Persons admitted once or more	Admission rate ^a per 1 000 man/years at risk
Neoplasms:				
ICD codes: 140–229 with 519.2 except 149, 166–169, 182–189, and 208–209	185	9.1	217	10.0
Central nervous system:				
ICD codes: 330–398 with 780–781.7 except 335–339, 346–349, and 358–359	88	4.3	92	4.2
Cardiovascular disease:				
ICD codes: 400–468 with 782.0–785.3 except 417–419, 423–429, 435–449, 454–459, and 784	192	9.5	210	9.6
Respiratory disease:				
ICD codes: 003.1 and 241, 470–527, with 782.3–783.7 except 476–479, 484–489, 494–499, 503–506, 508–509, and 527.1	79	3.8	71	3.2
Digestive disease:				
ICD codes: 530–545, 560–561, 570–578, 580–587, 782.8, and 784–785	174	8.6	195	9.0
All other diagnoses	415	21.3	396	18.8

^a These rates have been calculated using different times at risk because once a particular event occurs the individual is no longer at risk of that event occurring. This means that the time the individual is at risk depends on the event in question; for total admissions, the individual is at risk until he dies or is lost to observation.

^b According to the 1955 revision of the International Classification of Diseases (ICD) (5).

Table 4. Some measures of morbidity: screening versus control groups at the concluding health survey in 1972-73, 5 years after initial screening

	Control group (%) N = 1 950 (max) ^a	Screened 1967-68 (%) N = 1 651 (max) ^a	Refused screening 1967-68 (%) N = 327 (max) ^a	Total screening group (%) N = 1 978 (max) ^a
A. Questionnaire measures of general health				
1. Claiming to have good or excellent health in the fortnight preceding the survey	56.5	53.2	56.7	53.6
2. Admitting to any major disability, e.g., inability to dress or undress self	1.8	2.0	5.0	2.5
3. Showing downward social mobility, i.e., a fall in social class over the preceding 5 years	27.4	27.4	27.1	27.4
B. Cardiovascular disease				
1. With evidence of angina on questionnaire ^b	22.4	21.9	21.4	21.9
2. With raised diastolic blood pressure \geq 105 mm Hg (Ph V)	3.1	2.7	2.4	2.8
3. With ischaemic changes on ECG ^c	16.6	17.6	21.0	17.9
C. Respiratory disease				
1. Still smoking	50.8	51.5	56.2	52.3
2. Complaining of any bronchitic symptoms ^d	30.6	28.4	34.9	29.0

^a Maximum means that this was the largest of the denominators used to derive the percentages below. These denominators varied according to the information available for the multifactor analysis.

^b Rose, G.A. (6).

^c Rose, G.A. & Blackburn, H. (7).

^d Medical Research Council Committee on Research into Chronic Bronchitis (8).

APPLICATION

The findings of this study have provided the answer for health planners; it is clearly inappropriate on health and economic grounds to introduce multiphasic screening for the middle-aged in the United Kingdom.

Although this was an expensive study, both in time and resources, the answer more than justifies its cost. The United Kingdom is among the few developed nations that have resisted spending large amounts of public money on screening its middle-aged population throughout the 10 years of this study. The completed study illustrates that such screening is not a necessary, or even a useful, public service.

This study also illustrates how controlled trial methodology can be applied to evaluating new, and even existing, public services. Provided an acceptable and measurable outcome can be developed, this experimental approach to societal innovation is clearly preferable to political advocacy. The use of such controlled trials need not be confined to the medical field; they could be set up anywhere where a long-term, expensive, public programme is envisaged. The main drawbacks, however, are in the choice of acceptable outcome measures and the length of time taken to perform accurate and significant trials.

REFERENCES

1. **South-East London Screening Study Group.** A controlled trial of multiphasic screening: results of the South-East London screening study. *International journal of epidemiology*, 6: 257–263 (1977).
2. **Trevelyan, H.** A study to evaluate the effects of multiphasic screening within general practice in Britain: design and methods. *Preventive medicine*, 2: 278 (1973).
3. **D'Souza, M.F. et al.** Screening for hypertension in general practice: the results of a long-term controlled trial. *Lancet*, 1: 1228 (1976).
4. **Nelder, J.A.** *General linear interactive modelling (GLIM) manual. (Release 2.* Distributed by the Numerical Algorithms Group). London, Royal Statistical Society, 1975.
5. **World Health Organization.** *Manual of the international statistical classification of diseases, injuries, and causes of death, 1955 revision.* Geneva, 1957.
6. **Rose, G.A.** The diagnosis of ischaemic heart pain and intermittent claudication in field surveys. *Bulletin of the World Health Organization*, 27: 645 (1962).
7. **Rose, G.A. & Blackburn, H.** *Cardiovascular survey methods.* Geneva, World Health Organization, 1968 (WHO Monograph Series, No. 56).
8. **Medical Research Council, Committee on Research into Chronic Bronchitis.** *Questionnaire on respiratory symptoms.* London, Medical Research Council, 1966.

PROXY MEASURES REQUIRED FOR DISTRIBUTION OF HEALTH RESOURCES IN ENGLAND^a

C. Graham

INTRODUCTION

In May 1975, the Secretary of State for Social Services in the United Kingdom appointed the Resource Allocation Working Party (RAWP), comprising officers of the Department of Health and Social Security (DHSS) and the National Health Service (NHS), to review the arrangements for distributing NHS capital and revenue to Regional Health Authorities (RHAs), Area Health Authorities (AHAs) and Districts, respectively, to establish a method of distribution which responds objectively, equitably, and efficiently to relative need and to make recommendations.

METHOD AND RESULTS

The report of RAWP was presented in September 1976 (*1*). The report also considered measurements for capital resources and for additional NHS service costs arising from the clinical teaching of medical and dental students but the substantial recommendations related to measurements for the allocation of revenue to the NHS. The following text is reproduced from "*Sharing resources for health in England*", pages 14–27.

"Measuring Need for Non-psychiatric In-patient Services

Age and sex. There is ample evidence to demonstrate that men, women, boys and girls of different ages place different demands on hospital in-patient services. The age/sex make-up of the population of different Regions has a significant effect upon the need of each population for resources. To reflect this, it is necessary to weight the population by the demands made by each age/sex group on services. The best available measure is the national utilisation

^a Example based on Department of Health and Social Security, United Kingdom (*1*).

of hospital beds. We recommend, therefore, that the population of each Region be weighted to reflect the difference in age/sex composition by reference to the national pattern of non-psychiatric hospital bed utilisation.

Morbidity. Need for hospital in-patient services is not, of course, a function of age and sex alone. Many other factors are known to play a part — social, occupational, hereditary, environmental, etc. The difficulty is not in determining which factors are likely to be influential, but in quantifying their influence and in eliminating overlap between them. Figures are available, for example, on relative population densities and on social class structures, but we have not found it possible to relate this information quantitatively to the need for health care. Furthermore, factors such as occupation, poverty, social class and pollution are likely to interact in ways which are not fully understood.

But it would not be necessary to take account of causal factors such as those mentioned above if it were possible to measure health care need directly. In our Interim Report we relied upon Regional in-patient and out-patient caseloads as an indicator of relative need over and above that arising from the age/sex structure of the population. We recognised that this had serious imperfections. Whilst numbers of cases clearly reflect need, they do so in terms of the available supply of services. Caseloads fail both to distinguish between degrees of need and to assess the extent to which need is unmet through lack of facilities. Waiting lists as one indicator of unmet need are also known to have questionable reliability. Moreover there is ample evidence to support the view that the level of supply has a significant influence on the level of demand. Need must, therefore, be measured by an indicator that is far less dominated by supply.

Statistics relating to payment of sickness benefit are more independent in this sense but do not apply to the whole of the population, important categories such as the elderly, children and many married women being excluded. There are also problems relating to the causes of incapacity as certified, and Regional differences may be partly attributable to industrial structure: the ability to continue to work despite the presence of morbid conditions may for example be influenced by the nature of employment. Moreover sickness absence does not necessarily imply a need for health care over and above that which can be provided by a GP. The General Household Survey provides evidence of differences in GP consultation rates and the prevalence of self-reported sickness between different parts of the country but the nature of the sampling frame does not permit compilation of statistics in terms of NHS boundaries. Self-reported sickness is not a direct measure of the need for health care resources and differences in the levels of reported sickness may be due in part to differences in the perception and reporting of sickness. Nor would the data on diagnostic category be sufficiently reliable for our purposes. Past ad hoc surveys of morbidity in various fields cannot usually be extrapolated to national level. The samples are usually small, the data rapidly become out of date and repetition would be a difficult and costly business.

The search for a reliable indicator, as independent as possible of supply, which could be used to assess Regional differences in need led us to examine the possibility of using mortality statistics as a proxy indicator of morbidity. Mortality statistics cover the whole population, are readily available and

permit compilation by place of usual residence. The quality of the statistics, including analyses by cause of death, is high. The crude death rate shows a considerable Regional variation (maximum exceeds minimum by 38% for both males and females). Even when allowance is made for age structure — which has a marked effect on comparative death rates — the residual variation is still as high as 28% for males and 21% for females. Fig. 1 illustrates the variations. The reasons for the pattern of differential Regional mortality are not wholly understood but it is believed that Regional differences in morbidity explain the greater part of it and that statistics of relative differences in Regional morbidity, if they existed, would exhibit the same pattern as those for mortality.

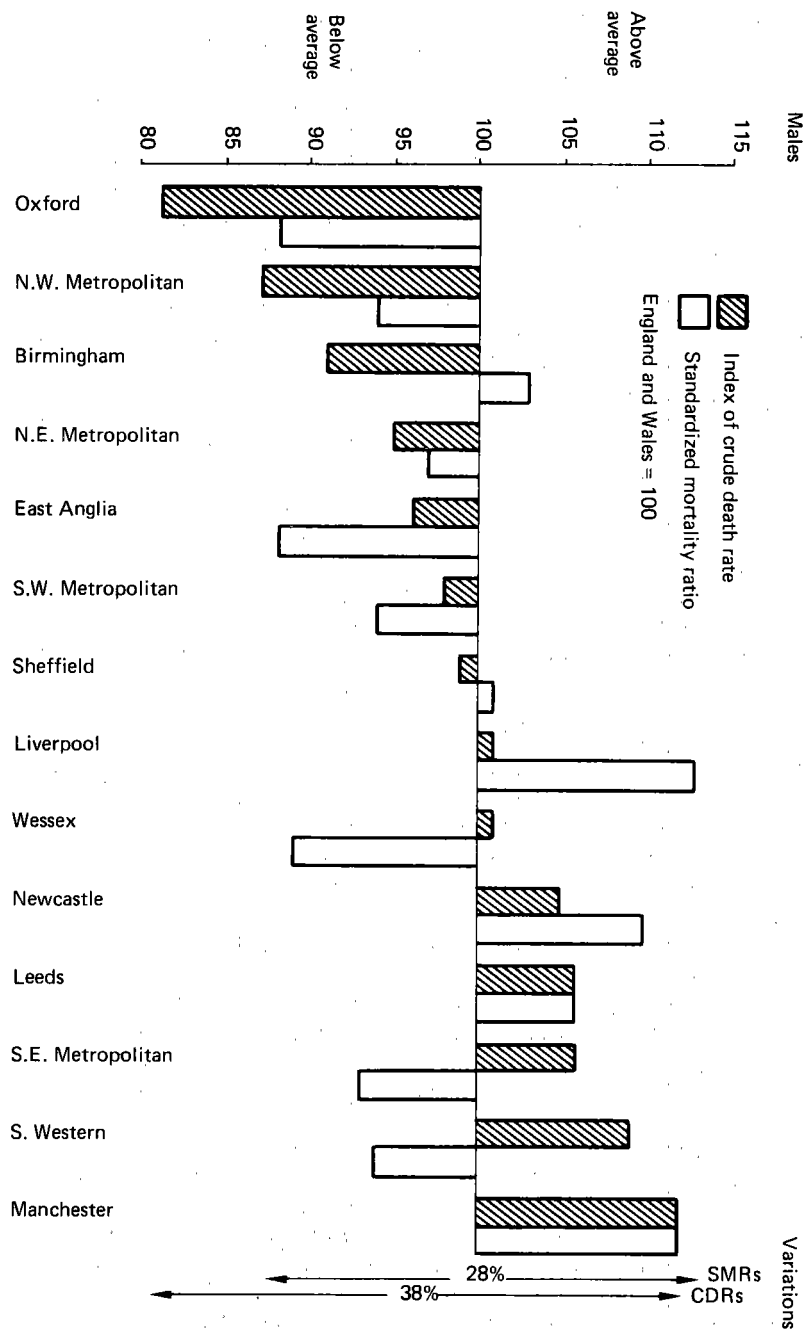
Some support for this assumption is provided by a comparison of mortality rates, adjusted to take account of age and sex differences, with such Regional morbidity-related data as exist, similarly adjusted. The comparison reveals significant positive correlations. The maps in Fig. 2–5 show the broad similarity between Regional differences in mortality and in data derived from sickness benefit statistics and the General Household Survey. Problems in using this information directly in the allocation process have already been outlined but these are less critical in the context of establishing geographical correlations with mortality because these can be calculated for standard statistical regions.

Mortality statistics also present an opportunity to relate differential morbidity to health care need by reference to conditions in a way that no other sources permit. It is possible to examine the variation in mortality between Regions by diagnostic conditions — using the underlying not the associated causes of death — grouping the conditions according to the 17 chapter headings of the International Classification of Diseases (ICD). The statistic used is the Standardised Mortality Ratio (SMR) which compares the number of deaths actually occurring in a Region with those which would be expected if the national mortality ratios by age and sex were applicable to the population of that Region. In this way the unique pattern of mortality in each Region can be established, calculated separately for each condition or group of conditions.

Many of the commonest conditions — including some which lead to death — place relatively little demand on health care services. Others require expensive care, perhaps over a long period. This relationship can be established by reference to the national figures of hospital bed utilisation for each condition category considered and incorporated in the calculation to provide the final link in the chain from mortality through morbidity to need for health services.

To each Region's population we have applied national age/sex utilisation rates for each individual group of conditions, calculated a standardised mortality ratio for each group and combined the two weighting factors for each condition. The effect of doing this is to produce a set of weighting factors independent of Regional differences in the supply of NHS facilities and reflecting morbidity differences between different parts of the country over and above those resulting from age and sex disparities. The method ensures that, in applying SMRs by condition, account is taken of the proportionate national

Fig. 1. Effects for each Regional Hospital Board of standardizing the 1971 crude death rates for age in males



bed utilisation for each condition. Fig. 6 shows how the results compare with those of a population weighting based on hospital utilisation analysed by age and sex alone.

As a result of the studies and analyses we have carried out, supported by the findings of research in related fields and expert advice, we have come to the conclusion that SMRs – adjusted in the way we propose – are the best available indicators of geographical variations in morbidity. And to ignore the considerable variations which this analysis displays would be to ignore a crucial factor in determining the relative needs for health care of different localities.

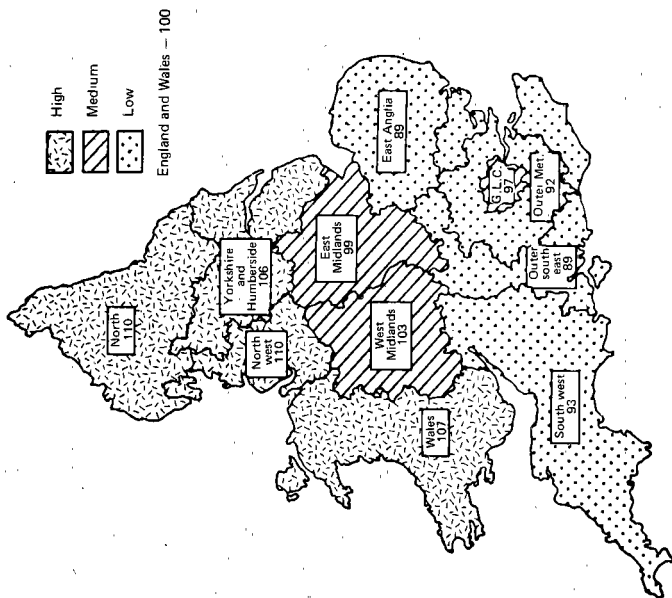
For certain conditions where mortality is very low – e.g. skin diseases and conditions of pregnancy – SMRs are unlikely to give a good guide to morbidity, and we have omitted these from our calculations. Age/sex weighting alone is a good indicator of the need for maternity services but it can be further improved by modifying the age/sex weighting for ICD Condition XI (conditions of pregnancy, childbirth, puerperium) to reflect fertility rates standardised for age, in the same way as other condition categories can be modified by SMRs.

There is evidence that other factors, e.g. marital status, are associated with the need for health care, but to add to those we have already proposed would strain the data to the point at which reliability was lost and incur a risk of double-counting. **We recommend that, in respect of acute nonpsychiatric hospital in-patient services, the population weighted for age and sex by national bed utilisation for each condition should be adjusted to take into account condition-specific SMRs for each Region. SMRs for conditions unlikely to lead to death, e.g. skin diseases, should not be used. For conditions of pregnancy, childbirth and puerperium, SMRs should be replaced by an index of fertility rates standardised for age.**

Cost-weighting. The cost of providing health care differs according to the condition being treated. In principle, therefore, the weighting system described above ought to be improved by attaching differential costs to the utilisation data. In practice, this is one of the areas where information is weakest. It is not at present possible to establish costs relative to ICD conditions. We have, however, set in hand a study which may enable a broad distinction to be made between conditions requiring predominantly acute or non-acute care, and the results, which will not be available until the autumn, may make it possible to apply a form of cost-weighting pending the availability of better information.

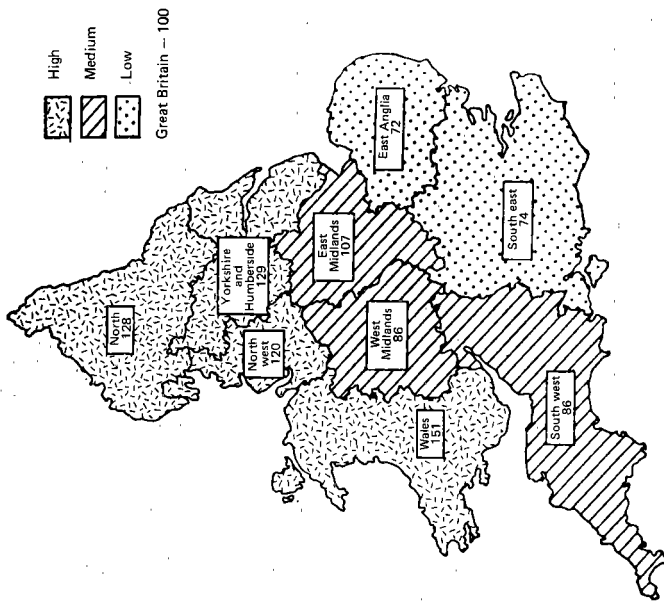
Movement of patients across administrative boundaries. Allocations must reflect the populations served, not simply those who reside within the administrative boundary. Where a RHA is responsible for managing services located outside its own geographic Region under a formal agency arrangement, the population there should be credited to the managing Region and not the Region of residence. Similarly, the costs of care for people who cross Regional boundaries to receive hospital in-patient services or are treated in hospitals managed by other RHAs (i.e. extra-territorially managed hospitals) should be credited to the Region providing treatment, and debited to the Region in which they live. The adjustment should have regard to the average

Fig. 2. Standardized mortality ratios in England and Wales, 1972: all persons



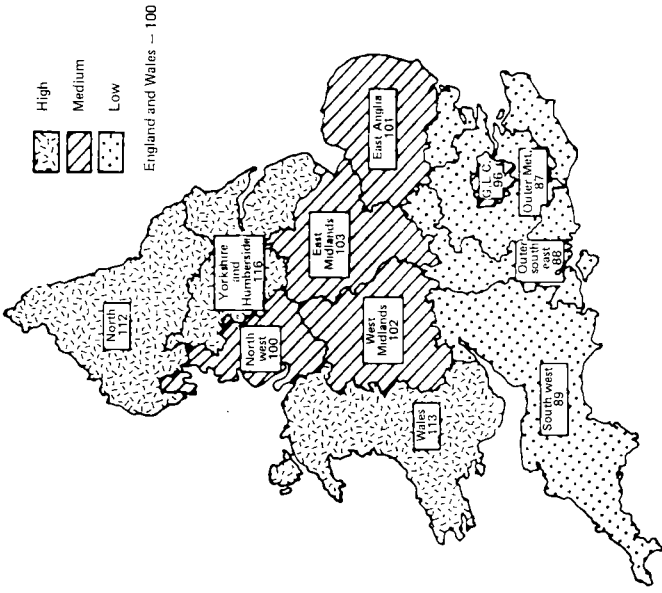
Note: Subdivisions of England and Wales are standard statistical regions.

Fig. 3. Certified spells of incapacity in England and Wales, 1972; males



Note: Females have been excluded because of the large proportion who are not insured for sickness benefit.

Fig. 4. Self-reported chronic illness in England and Wales standardized for age and sex, General Household Survey, 1972; all persons



Note: GHS data: Acute sickness is defined as restriction at any time during a 2-week reference period of the level of normal activity caused by illness or injury. Chronic sickness is defined as a state of long-standing illness, disability, or infirmity.

Fig. 5. Self-reported acute illness in England and Wales standardized for age and sex, General Household Survey, 1972; all persons

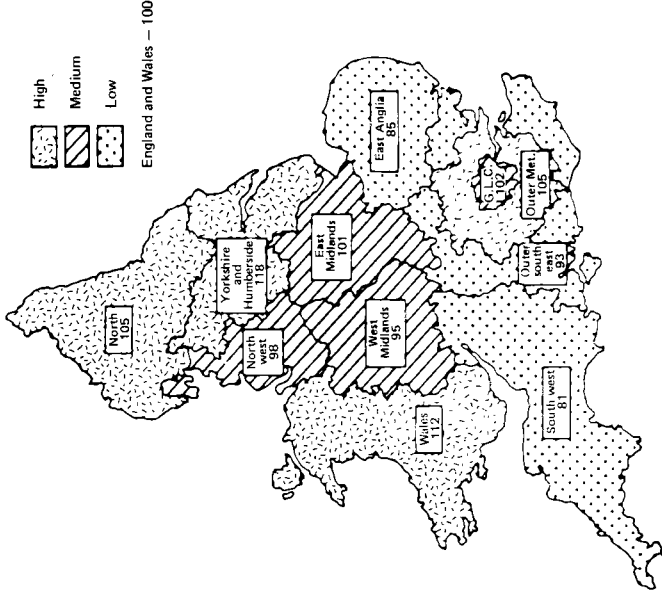
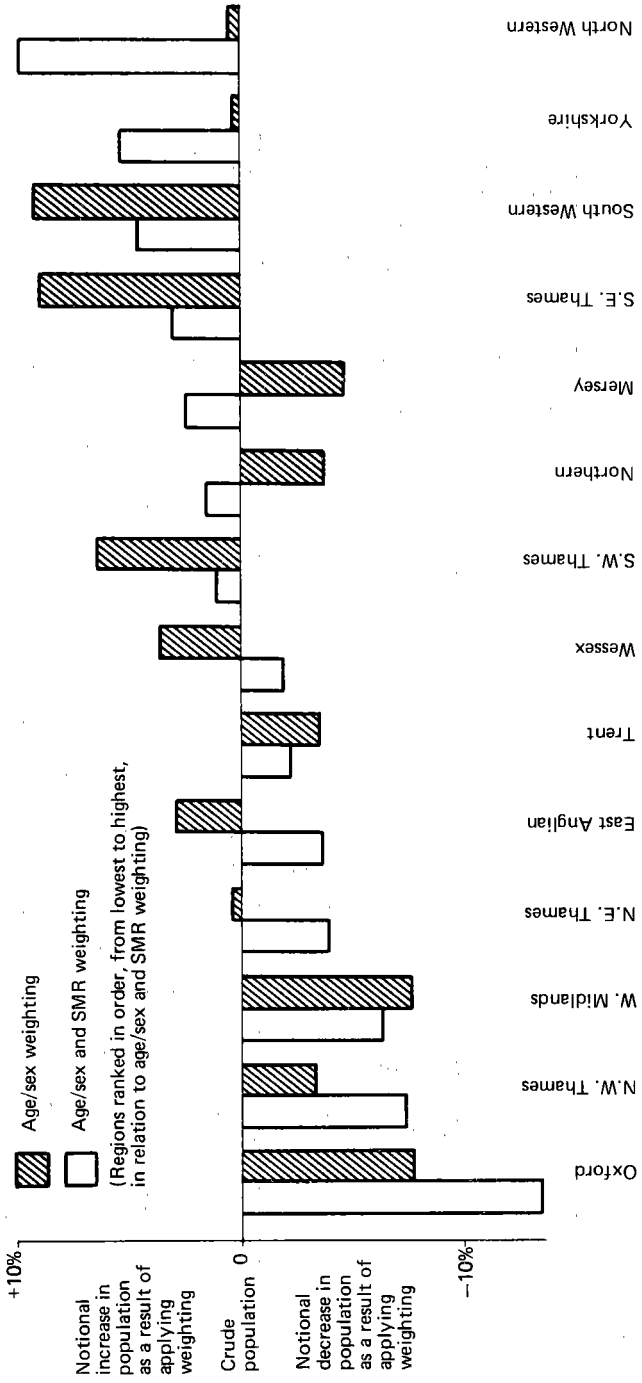


Fig. 6. Effects of age/sex weighting and age/sex/standardized mortality ratio weighting, applied to each region's crude population



national cost of the kind of services provided, rather than to local costs which may be influenced by local policy decisions. **We therefore recommend that adjustments to weighted populations should be made to take account of inter-Regional patient flows costed on a national specialty basis:**

Measuring Need for Day- and Out-patient Services

All of the principles described in the foregoing paragraphs in relation to in-patient services apply, *mutatis mutandis*, to day- and out-patient services also. The necessary differences are:

(1) *Age/sex weighting* should reflect national utilisation of services by out-patients and not that by in-patients, since the utilisation pattern is different between the two groups.

(2) *Morbidity* among non-psychiatric day- and out-patients is just as likely to be reflected by SMRs as it is among in-patients. Unfortunately it is not practicable to break down utilisation of these services by ICD condition so as to apply a utilisation weighting similar to that recommended for non-psychiatric hospital in-patients. It is, however, possible to modify the age/sex weighting for these groups by applying the overall SMR for each Region. We have established that overall SMRs when applied to in-patients give broadly the same result as the condition-specific SMRs. Though less sensitive, overall SMRs may therefore be applied to non-psychiatric day- and out-patients.

(3) *Movement of patients across administrative boundaries*. There are no statistics generally available to measure reliably the extent to which day- and out-patients cross Regional boundaries for services. We have searched for a different indicator which might be used as a proxy; but, after very full consideration, we have come to the conclusion that no assumptions as to the validity of any proxy can be made with any confidence. For example, whilst in-patient flows in some localities are thought to correspond approximately with out-patient flows, there is evidence to suggest that this correspondence is by no means general. We are not, therefore, in a position to make any recommendations save that we regard it as an urgent need to assemble better information about this group of patients, who represent a substantial and increasing proportion of NHS expenditure. Agency arrangements can and should be taken into account.

We recommend that, in respect of non-psychiatric day- and out-patient services, the population of each Region should be weighted to reflect the national pattern of utilisation of these services by age and sex, adjusted in the case of non-psychiatric patients to take into account SMRs for each Region. The weighted population should be adjusted to take account of agency arrangements.

Measuring Need for Community Services

The arguments and the data limitations applying to people using community services are similar to those for day- and out-patients. Age utilisation

patterns nationally are different from those for other services and this should be taken into account. No sex analysis is possible. We recommend that, in respect of community services, the population of each Region be weighted to reflect the national pattern of utilisation of these services by age, adjusted to take into account SMRs for each Region. The weighted population should be adjusted to take account of agency arrangements.

Measuring Need for Ambulance Services

Examination of the available data on use of ambulance services, which, as for other community services, are far from comprehensive, reveals that by far the most significant explanation of Regional variation is crude population. Age/sex weightings cannot be applied on existing information. Demands on the services will, however, be affected by variations in morbidity, and the use of SMRs in the same way as for other community services is therefore desirable. The Working Party examined very carefully the possibility of taking variations in population density into account as well, but the study did not point conclusively to the need to adjust Regional allocations on this account. We recommend that, in respect of ambulance services, the crude population of each Region be adjusted to take into account SMRs for each Region.

Measuring Need for FPC Administration

FPC administration consumes a tiny proportion of the available revenue funds. There is no way of relating the need for this service to detailed indices of morbidity. We therefore recommend that the index of need for FPC administration should be crude population.

Measuring Need for Mental Illness Hospital In-patient Services

Age and sex. As for other services, there is evidence that the pattern of utilisation of mental illness services nationally differs between age/sex groups. We recommend that the population of each Region be weighted to reflect the difference in age/sex composition by reference to the national pattern of mental illness hospital bed utilisation.

Morbidity. Mortality is clearly not an appropriate measure for psychiatric morbidity, over and above that explained by age/sex variations, since mental illness is rarely the direct cause of death. We sought expert advice on the best indicators of need for these conditions in addition to age and sex. As in the case of physical illness, many potentially relevant indicators – social class, poverty, social isolation and others – would need considerable research before it was possible to include them in any formula. There is, however, quantitative evidence that non-married people place heavier demands on mental illness services nationally than do married people and the age/sex weighting can be modified to take account of this. We recommend that, in respect of mental illness hospital in-patient services, marital status be used as an additional weighting factor to age and sex, pending the outcome of further research on other possible indicators.

Cost-weighting. As for other services, the necessary information is lacking and must be added to the long list of research and data requirements.

Movement of patients across administrative boundaries. The same arguments apply as those adduced above (see last paragraph, page 239) in relation to non-psychiatric hospital in-patients and lead to similar conclusions, though it is not possible to include a cost-weighting because of the data deficiencies already mentioned. In addition, however, estimates of the need now likely to arise for mental illness services do not and cannot reflect the presence of long-stay patients, often admitted many years ago, whose original homes are in many cases no longer known. Such patients tend to be concentrated for historical reasons in relatively few places. **We recommend that patient flows should be taken into account in respect of mental illness hospital in-patient services with the modification that long-stay mentally ill patients admitted to hospital before the date of the last MHE Census should as a temporary measure be credited to the Region in which they are receiving care by means of an adjustment to reflect the actual as against the expected number of such patients based on weighted population. The adjustment should be phased out as existing imbalances are corrected, as the purpose is to recognise those imbalances due to historic crossing of boundaries not reflected in current patient flows.**

Measuring Need for Mental Handicap Hospital In-patient Services

Similar arguments apply to mental handicap services as to mental illness services, except that we have been unable to identify any criteria of need available for application in the short term other than age and sex differences. **We recommend that the population of each Region be weighted to reflect the national pattern of mental handicap hospital bed utilisation by people of different ages and each sex, and that the recommendations in the preceding paragraph should be applied in respect of mental handicap hospital in-patient services.**

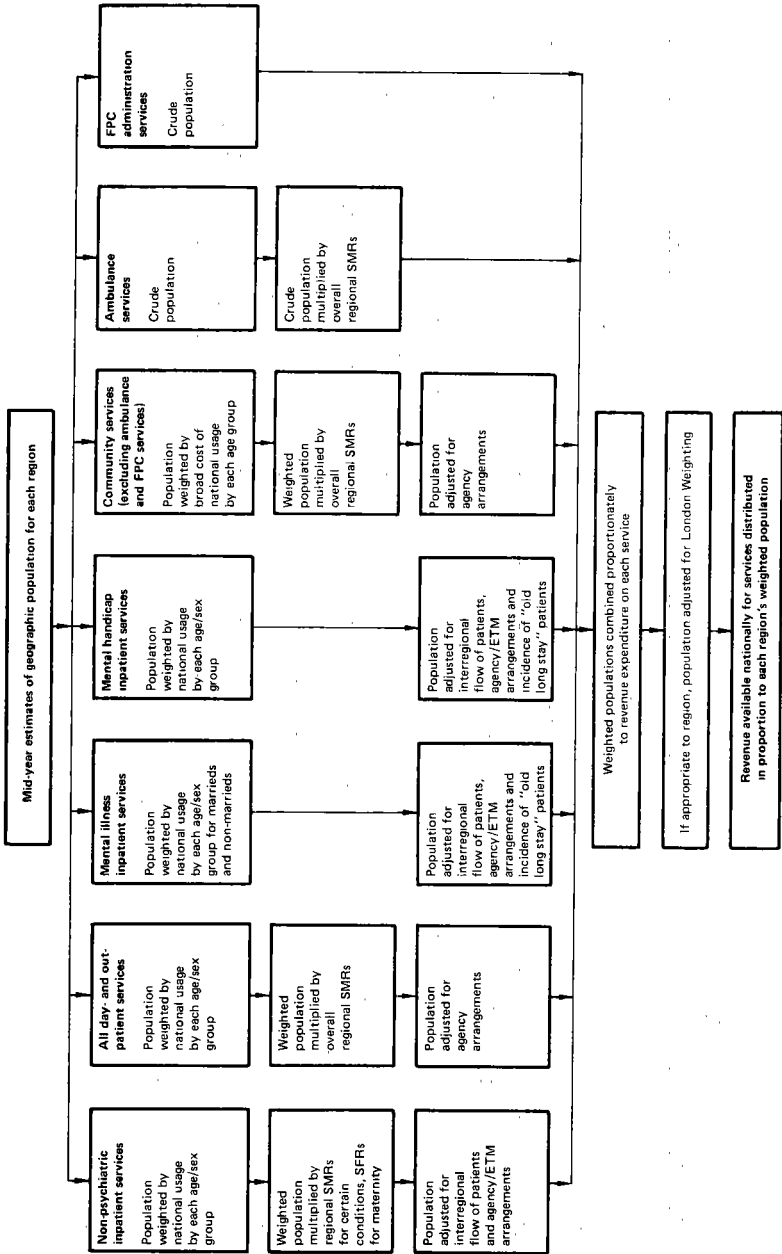
Caseload

With the improvements outlined above, the retention of caseload as a separate element in the formula can no longer be justified. It is not a satisfactory indicator of need since the number of cases treated must depend to a large extent on the availability of facilities to treat them. It was retained for 1976/77 only as an expedient to make up for the inadequacies in the age/sex weighting of population. **We recommend that the caseload measure should now be abandoned.**

Establishing Revenue Target Allocations

Applying the recommendations in the preceding paragraphs produces seven separate weighted populations for each Region. **We recommend that these should be combined into a single weighted population for each Region in proportion to the most recent information available on relative expenditure**

Fig. 7. The build-up of a revenue target



nationally on the services concerned. The revenue available for services nationally should then be notionally distributed in proportion to each Region's weighted population to arrive at the revenue target allocation for each RHA. The flow chart at Fig. 7 sets the whole process out in graphic form."

APPLICATION

The above recommendations were accepted by Ministers and have been implemented as far as possible in NHS revenue allocations for 1976–77, 1977–78, and 1978–79 and the process is expected to continue in future. Further research into some of the aspects mentioned will take place over the next few years in the hope of achieving a fairer distribution of resources by the end of the 1980s. An indication of the effect of action taken so far is that in 1976–77 before implementation of the recommendations the pattern of expenditure between regions revealed a spread of over 50% about a national average: after implementation that spread was reduced to 36%.

CONCLUSION

This example reveals the importance of blending the political requirement, the academic approach, the professional discipline and the administrative action to produce the required degree of change: it took just over a year to move from the political requirements through the stages shown to the administrative action. This perhaps offers an example for similar activities elsewhere.

REFERENCES

1. **Department of Health and Social Security, United Kingdom.** *Sharing resources for health in England: report of the Resource Allocation Working Party.* London, Her Majesty's Stationery Office, 1976.

