Total care usage of a defined population

THE BIRMINGHAM RESEARCH UNIT of THE ROYAL COLLEGE OF GENERAL PRACTITIONERS

SUMMARY. The North Staffordshire (Stoke) Morbidity Survey was set up to bring together data about the total use of general practice, hospital and local authority services by each individual patient in a representative sample of the population of Stoke, and to examine the associated problems of data linkage and confidentiality. Approximately 47·8 per cent (32·1 per cent if allowance is made for all measurable possible errors) of the patients attending hospital and 43·8 per cent of patients attending local authority services during the year, did not attend their general practitioner.

Primary automation of the registered list of patients at risk is essential to the success of any automated linkage study. Where the list consists of Hogben numbers to identify individuals, this also ensures the necessary confidentiality of sensitive data.

Introduction

THIS survey, originally called the North Staffordshire, was conducted by the Records and Statistics Unit of the Research Committee of the Royal College of General Practitioners. In association with the North Staffordshire Medical Institute it was the first of a series of surveys planned by the Research Committee of the College to study the various ways in which differing social and environmental factors influence morbidity and medical care.

The survey population was defined as the number of individuals registered under the NHS with 15 selected general practitioners in the city of Stoke. This amounted to 32,656, a 12·3 per cent sample of the population of the city. The main study began on 1 May 1964 and lasted for 12 months. It was preceded by a pilot study during March 1964.

Aims

The North Staffordshire Morbidity Survey was designed to provide information that had not been available previously about how much the general medical services as a whole were used by a defined population.

In particular it was designed to study the problems involved in linkage of data about the same patient from different medical care agencies; to enable a comparison to be made between those who have made use of any of the medical services and those who have not done so during the period of the survey; to develop a method by which information about total medical care might be collected from all relevant sources; and to ensure that the system preserved confidentiality for the patient.

The city of Stoke was chosen for several reasons. It formed a relatively closed community for medical care services and very few patients had to go elsewhere for specialist hospital services. The migration and immigration rates for the North Staffordshire area are low. There was keen interest among the general practitioners, the hospital services, and local authority staff to study the defined problems. The Birmingham Regional Hospital Board was planning studies involving sophisticated record systems in the city because of its unique demographic characteristics.

Method

A preliminary meeting of representatives of general practice, the public health department, and the hospital services was held in 1962 at which the general idea of the survey was put forward and discussed. As a result 15 practitioners from seven practices with lists totalling 32,656 patients (averaged during the survey year) volunteered to take part. Further discussions regarding the method of conducting the survey within the public health departments were continued with the medical officers of health for Staffordshire, Stoke-on-Trent, and the Borough of Newcastle-under-Lyme.

Recording systems

An age-sex register using the ledger system (College of

General Practitioners, 1963) was set up in each practice. Date of birth, sex, and address were recorded for all patients. It had been hoped to record also the map reference and rateable value of the dwelling at which the patient was registered, but this was not possible with the resources available. Each episode of illness was recorded in a standard Royal College of General Practitioners' disease index (RCGP Research Unit, 1971). There is a separate sheet for each rubric in the College classification of disease. Data about males are recorded on the front and data about females on the back of the sheet.

For each episode the following information was recorded on the appropriate sheet:

1. Date of first contact for the episode of illness.

2. First four letters of the patient's surname and his date of birth. This modified Hogben number (Hogben and Cross, 1948) formed an identification number for each patient in the survey.

3. Sex.

4. An 'O' was entered in the first consultation column (or an 'OP' if the first contact was by proxy) for a new episode of illness occurring within the survey year.

5. Subsequent consultations were recorded by a tick in the next vacant column (or by a 'P' where the contact was by proxy). If the disease episode being recorded began before the survey year, the O was omitted and replaced by a tick. Death was indicated by inserting 'D'.

6. A change of diagnosis was shown by entering the number of the new diagnosis in the spare columns at the end of the line.

7. Subsequent consultations were noted by entering ticks in the consultation columns.

8. If a patient was consulting for two simultaneous illnesses (e.g. thyrotoxicosis and a disclocated shoulder), both were recorded on separate sheets, but where a multiple diagnosis was due to complications of the same illness, only the principal diagnosis was recorded.

9. If a patient was referred to any other medical agency, a tick was inserted in the column 'R' of the sheet and the full details were entered up on a special referral form. These forms were sent to the records unit in batches and were used to provide a basis for preliminary search of hospital records. Occasionally, referral forms were completed without an entry in the disease index, particularly for out-of-hours emergency calls.

Public health records

All local authority units, clinics, and individual workers were issued with special cards upon which were entered the identification numbers of the patient as determined by the first four letters of his surname, his date of birth, the reason for contact, and the number of contacts. Completed cards were returned to the records unit and filed alphabetically, and at the end of the recording year they were linked with other records concerning each patient.

Hospital records

The referral form completed by each doctor at the time a patient was referred was used to obtain the hospital records from which extracts were made on a special form about the hospital services used by each of the patients concerned. The study had been planned to coincide with the introduction of automation of the inpatient and outpatient records in all Stoke hospitals.

In the event, this could not be implemented at the time because of the serious illness of one of the key staff. For this reason the data from this group of patients had to be acquired by manual search of all outpatient records and inpatient summary forms (BRHB Stat.6/1), and those relating to patients of the 15 general practitioners taking part in the study were identified. Relevant, but now highly selected, information was then transferred to the same special forms. This part of the project, which was less complete than originally planned, was also unavoidably delayed. The detailed data about the hospital services could not be included in the main analysis.

Confidentiality

No clinical or social information about patients should be accessible directly to anyone other than the *primary user* (Crombie, 1973). The primary user records, in a *primary file,* information legitimately acquired during his professional relationship with a patient or client.

Provided data which could be used to identify any individual are removed from this file, it can be legitimately used by others. Such a file is defined for this purpose as a *secondary file.* A file in which each individual is identified only by a Hogben number is still a secondary file, for the individual can be identified only by cross-reference from the Hogben number to the age-sex registers of named registered patients of the general practitioners.

In this study all linkage procedures were carried out on secondary files in which individuals were identified by Hogben number only.

Data linkage

The coded information from these four sources (disease indexes, practice referral forms, hospital records, and local authority records) was then transferred to 80-column punch cards and entered onto computer tape. This part of the project, the basis for an M.SC thesis, was carried out by a postgraduate student (J. Peel), from the Department of Engineering Production, University of Birmingham, using the university computer, an English Electric KDF9. The language used was EGDON (a variant of Fortran IV but offering more facilities than Fortran IV). Data about each individual patient from the various primary sources were then merged, using the Hogben number as the identifier.

Table 1. Survey population by age/sex showing percentage distribution compared with that of England

	Males							*Females*	
	<1	1-4	5-14	15-44	45-64	65+	Total	<1	1-4
Study population	56	1,066	2,345	6,656	4,333	1,597	16,053	49	961
Percentage distribution	*0.3*	*6.6*	*14.6*	*41.5*	*27.0*	*9.9*	*100*	*0.3*	*5.8*
Percentage distribution of England and Wales	*1.9*	*7.0*	*15.0*	*42.2*	*24.5*	*9.4*	*100*	*1.7*	*6.3*

Disc storage was not readily available at this time and the procedures involved direct matching and merging from tapes.

The first phase in this process was the production of an integrated general practice tape produced by matching and merging the disease index material with the practice referral forms. This integrated general practice file was the basis for analysis of general practice activities.

Perhaps the most important single defect of the system design was the absence of an automated age-sex register. The implications of this will be considered later.

Results

The registered survey population (Table 1) is representative of the population of England and Wales, except for a reduced proportion of patients under one year of age. This deficit reflects the delay in registration of the newly born. Often the need for the first consultation also stimulates registration.

Data from general practice

The consulting patterns based on disease index information are given in Table 2.

The data from practice 3 suggest that only one of the two partners was recording all the patients attending. Apart from practice 3, the mean proportion of all patients attending (73·19 per cent) is higher than the comparable figures from either the first or second National Morbidity Surveys. Notwithstanding this, the other rates are very comparable with the Second National Morbidity Survey undertaken in 1970-71. The First National Morbidity Survey was carried out in 1955-56. The data from the second study relates to 1970-71.

Data linkage study

On completion of the main data-linkage programme, it was clear that:

Table 2. Consulting patterns.

	Stoke Survey 1964-65 Practice number						
	1	2	3	4	5	6	7
Practice/survey size	3,463	2,273	6,912	5,052	3,372	4,276	7,308
Patients attending in year	2,467	1,592	3,223	4,296	2,656	3,504	6,162
Attendance %	*71.24*	*70.04*	*46.63*	*85.04*	*78.77*	*81.95*	*84.32*
No. of consultations	10,957	5,703	11,172	26,097	15,296	15,529	29,695
No. of consultations/patient in practice/survey	3.16	2.51	1.62	5.17	4.54	3.63	4.06
No. of consultations/patient attending	4.44	3.58	3.47	6.07	5.76	4.43	4.82
No. of episodes (illnesses NMS 1)	5,005	2,623	4,941	11,085	6,548	7,326	12,797
No. of consultations/episode	2.19	2.17	2.26	2.35	2.34	2.12	2.32
No. of episodes/patient in practice/survey	1.45	1.15	0.71	2.19	1.94	1.71	1.75
No. of episodes/patient attending	2.03	1.65	1.53	2.58	2.47	2.09	2.08
General practitioners per practice	1	1	2	3	2	2	4

nd Wales as at June 1964.

					Total population						
5-14	15-44	45-64	65+	Total	<1	1-4	5-14	15-44	45-64	65+	Total
2,156	6,392	4,468	2,577	16,603	105	2,027	4,501	13,048	8,801	4,174	32,656
13.0	*38.5*	*26.9*	*15.5*	*100*	*0.3*	*6.2*	*13.8*	*40.0*	*27.0*	*12.8*	*100*
13.5	*38.8*	*25.0*	*14.6*	*100*	*1.8*	*6.7*	*14.2*	*40.4*	*24.8*	*12.1*	*100*

1. There were *on average over the year* 32,656 separate identified patients in the non-automated basic practice age-sex registers.

2. Of these, 23,900 consulted their general practitioners during the survey year and were recorded as separate individuals in the practitioners' disease indexes.

3. Of the 4,338 patients (2,406 + 195 + 1,737) recorded as separate individuals on referral forms, 297 did not appear among the 23,900 recorded in disease indexes.

4. A further 2,382 individuals attended hospital during the year and were not recorded in disease indexes or on referral forms.

5. A further 1,477 individuals who were recorded as attending local authority services during the year did not appear in any general practice or hospital records.

Before examining the implications of these findings in detail, the problems encountered in the data linkage programme will be considered.

Problems in the data linkage programme

The problems of data linkage using the Hogben number as an identifier were explored using the data from the disease indexes only. This restriction was necessary because of the inevitable delays already described in obtaining the hospital data. The results confirm those already identified by Acheson (1966) in the Oxford Linkage Study, but different conclusions are to be drawn.

Errors affecting the linkage process can occur at various stages:

1. The date of birth and/or the name may be incorrectly given by the patient on different occasions.

2. When correctly given, it may be incorrectly recorded on documents.

3. When correctly given and recorded on documents, it may be incorrectly entered onto tape.

4. There may be computer errors.

			First National Morbidity Survey 1955-56			*Second National Morbidity Survey 1970-71*		
Total	Males	Females	Total	Males	Females	Total	Males	Females
32,656	16,053	16,603	382,829	180,060	202,769	292,247	140,346	151,901
23,900	11,076	12,824	256,595	114,294	142,301	196,292	88,529	107,763
73.19	*69.00*	*77.24*	*67.03*	*63.48*	*70.18*	*67.17*	*63.08*	*70.94*
114,449	50,782	63,667	1,436,155	609,571	826,584	879,554	358,890	520,664
3.50	3.16	3.83	3.75	3.39	4.08	3.01	2.56	3.43
4.79	4.58	4.96	5.60	5.33	5.81	4.48	4.05	4.83
50,325	22,698	27,627	509,446*	217,929*	291,517*	528,562	215,327	313,235
2.27	2.24	2.30	2.67*	2.74*	2.62*	1.66	1.67	1.66
1.54	1.41	1.66	1.33*	1.21*	1.44*	1.81	1.53	2.06
2.11	2.05	2.15	2.0*	1.9*	2.1*	2.69	2.43	2.91
15	—	—	106	—	—	53	—	—

*These figures exclude prophylactic procedures and routine antenatal episodes.

Table 3. 'E' book records—(data type 1) detected error by practice.

Practice	Number of errors	Number of entries	Errors %	Verified	Errors in dates of birth accepted (later rejected at Keele)
1	106	4,912	*2.16*	No	1
2	11	2,573	*0.43*	Yes	1
3	172	4,730	*3.64*	No	0
4	97	11,212	*0.87*	Yes	0
5	57	6,474	*0.88*	Yes	0
6	92	7,246	*1.27*	No	5
7	65	12,613	*0.52*	No	2
	600	49,760	Mean *1.21*		9
Number unverified	435	29,501	*1.47*		
Verified	165	20,259	*0.81*		

Other records were added after this statistic had been extracted, hence the differences recorded in the number of records (episodes) 49,760 here and 50,325 elsewhere.

Errors in disease indexes

The influence of errors in 3 was greatly reduced by verification procedures. These took the form of double punching of all coded data. This double punching was carried out in three of the practices involving some 20,259 disease index records (40·7 per cent of the total).

It is difficult to separate the errors in 1 and 2 but Table 3 gives the 'logical' errors detected by the computing programme. Out of a total of 49,760 cards, there were 600 detected logical errors (1·2 per cent of the total). There is a considerable variation between practices but the verified cards have a lower mean error rate (0·814 per cent compared to 1·475 per cent).

In this study the combined effects of coding and computing errors were trivial and can be dismissed.

Other errors may occur and remain undetected. The extent of some of these can be estimated from the following analysis carried out by the recorder for practice no. 4, by cross-checking the practice diagnostic index (Table 4).

The error rate for the recording of sex is relatively high. It is a direct result of the recording method. Males and females are entered on the front and back respectively of the same diagnostic index sheet. The only errors affecting linkage, however, were those due to the date of birth, and the name; 45 in 11,212 or 0·40 per cent occurred in practice no. 4. Errors found in a name, outside the first four letters of that name, were ignored in the analysis. Most errors occur in clusters; for example, the name or the whole date of birth may be wrong. The number of resulting mismatches is always less than the number of total errors. If the above figures are representative, the estimated total number of these 'non logical' errors in all practices will be 566 (1·124 per cent of 50,325). These would have resulted in approximately 201 mismatches (0·40 per cent of 50,325).

The error rates for the more complex coded cards appear at approximately the same rates of errors per column. A more detailed presentation of the verification and validation programme is given elsewhere (Peel, 1967).

The figures shown above are based on the total number of errors per card; therefore, to derive an error rate per column the figures should be divided by 24. This will be a rough guide since errors were more prevalent in some columns than in others.

This relatively high error rate, approximately 0·097 per cent

$$\left(\frac{1{\cdot}21 + 1{\cdot}12}{24} \text{ per cent} \right)$$

of all columns, would affect approximately 0·40 per cent of all records for linkage purposes. Mismatching of this order, low compared with the figure of 15 per cent which Acheson found in the Oxford Linkage Study (1966), is acceptable for statistical purposes but quite unacceptable if the linked records are to be used for actual patient care. The lower rates in the general practice recording in this study were achieved for two

Table 4. Disease index recording errors (practice no. 4).

Record type	*Number of errors*	*Error %*
Name	6	*0.054*
Date of birth	39	*0.348*
Sex	81	*0.722*
Total	126	*1.124*

Table 5. Comparison of patients attending general practitioner, taken from patients in integrated general practice file.*

	Practice number							
	1	2	3	4	5	6	7	Total
Number of patients attending general practitioner only	1,766	1,323	2,614	3,649	2,220	3,059	5,228	19,859
Number of patients attending general practitioner and hospital	410	116	218	647	381	177	457	2,406
Number of patients attending general practitioner and local authority	282	161	384	55	103	287	465	1,737
Number of patients attending general practitioner, hospital and local authority	50	12	22	20	23	16	52	195
Total	2,508	1,612	3,238	4,371	2,727	3,539	6,202	24,197
Number of patients recorded in diagnostic index as attending general practitioner	2,467	1,592	3,223	4,296	2,656	3,504	6,162	23,900
Difference	41	20	15	75	71	35	40	297
Error %	*1.6*	*1.2*	*0.5*	*1.7*	*2.6*	*1.0*	*0.6*	*1.2*

*Integrated general practice file consists of linked data about patients recorded in disease indexes and practice referral forms.

main reasons. All the participants knew that the patient's name and full date of birth were to be used for linkage, and therefore much more care was taken with the initial recording and transcription. The 'logical' validation procedures were remarkably effective.

There were 23,900 patients recorded in the disease indexes as attending their general practitioners during the survey year. The equivalent number in the primary integrated general practice file is 24,197. The excess of 297 individuals (1·2 per cent) in the integrated file includes patients who attended hospital during the survey year after referral by their practitioners but were not recorded in the disease index (Table 5). This applies particularly to emergency and out-of-hours' admissions, but not to those who consulted their practitioner before the recording period. The equivalent error in a validation procedure from the first year of the Second National Morbidity Survey was 35 in 1,000 representative entries in the clinical records (OPCS, RCGP and DHSS, 1974). The total errors in these surveys due to non-recording in the disease indexes are not known.

Errors in hospital derived data

The Hogben numbers in this set of data were derived from ordinary clinical records and were therefore liable to the high error rates found by Acheson at Oxford. No attempt was made to estimate this error but it was assumed that, as at Oxford, 15 per cent of Hogben numbers derived from hospital clinical records would have contained an error which would have prevented data linkage. In populations of the size with which we are concerned here, such errors would be most likely to result in non-matching and only occasionally, if at all, in mismatching (Crombie, 1973).

Other sources of errors

Approximately eight per cent of the registered population of Stoke change to another general practitioner during any one year. When the numbers leaving are roughly balanced by the numbers of new patients registering, as in the study practices, no important change in the age-sex structures of the practices results. This loss of patients from the practice during the year would not account for any of the 2,382 patients attending hospital who did not apparently consult their general practitioner during the study year. If any patient attending hospital changed his doctor during the course of the year and attended his new doctor, he would be recorded in the hospital records under the name of his new doctor.

A further 1·8 per cent of the population is accounted for by children born during the study year. All children born in hospital and their mothers appear in the hospital data as separate registered patients; however, only a small proportion of the children are registered with their general practitioners during the first year of life. We have already noted (Table 1) that the proportion of children under one year of age *at risk* to their general practitioner in the general population is

1·8 per cent (an estimated 588 children) while the equivalent proportion registered is only 0·3 per cent (an estimated 105 children). If the practices contained the expected 1·8 per cent of the population under one year of age there would have been 588 children registered instead of the actual number, 105. This discrepancy is obvious in the totals for this age group (Table 6), where the summed rates for use of all medical resources exceeds 1,000. However, of this estimated potential excess of 483 children at risk to the practitioner, but not yet registered, only 65 attended hospital and 76 were attended by local authority services (Peel, 1967). If these 65 children are removed from the 2,382 patients attending hospital but not their general practitioners, there remain 2,317 in this category.

In the age group 65 + there were 4,744 separate individuals receiving services while only 4,174 were registered with their general practitioner. This excess probably includes patients in long-stay institutions whose hospital records still carry the name of their last general practitioner, but whose names have been routinely withdrawn from his list of registered patients. This excess of 570 unregistered patients must be accounted for by the 432 patients in this age group recorded as attending hospital only, and the 378 receiving local authority services only. Even if all the 432 elderly patients in hospital in this age group were in long-stay institutions and now no longer registered with their general practitioners, this would only reduce the previous estimate of 2,317 to 1,885.

There is a final source of error to be considered. It is already known (Acheson, 1966) that the error rates in Hogben numbers generated from ordinary hospital records not specifically designed for this purpose can be as high as 15 per cent. In other words, among those 1,885 hospital attenders apparently not attending their general practitioner (Figure 1), there could have been some 15 per cent (283 patients) who did in fact attend their general practitioner but whose matches have been missed because of errors in the Hogben numbers. This still leaves a minimum of 1,602 patients in this category unaccounted for. If all these estimates of maximum error actually existed, the percentage of patients attending hospital who did not consult their general practitioner in any one year would have been reduced from 47·8 per cent (2,382/4,983) to 32·1 per cent (1,602/4,983). The equivalent percentages as proportions of the registered practice populations are 7·3 per cent (2,382/32,656) and 4·9 per cent (1,602/32,656).

As stated above, one of the major defects in the system used in this survey was the absence of an automated age-sex register of the practice patients. If this had been available, non-matching of the Hogben numbers from the hospital data with those from the practice data would have been immediately evident and could have been investigated at the time. The automation of age-sex registers was, of course, an essential feature of the Second National Morbidity Survey.

Linked data

This section deals with statistics derived when considering the usage by individual patients of the various medical services, bearing in mind the possible errors and problems outlined above (Table 6). This usage is presented as a series of exclusive sets in Figure 1.

Of the 4,338 (2,406 + 195 + 1,737) patients (13·3 per cent of the registered population at risk) attending the general practitioner and hospital or local authority services, 3,402 (10·4 per cent) were actually referred by the general practitioner during the recording year. The remaining 936 (2·9 per cent) shared patients not referred during the recording year may well have been referred before the study began. The total referral rate of 10·4 per cent (3,402/32,656) of patients at risk is

Table 6. Analysis of usage of medical facilities by age and sex during survey year. Number of patients/

	Males							*Females*	
	<1	1-4	5-14	15-44	45-64	65+	Total	<1	1-4
General practitioner only	607	510	524	535	585	829	575	388	481
Hospital only	625	57	55	43	65	130	62	592	63
Local authority only	732	98	111	21	4	63	41	714	113
General practitioner and hospital	143	41	51	48	75	102	61	82	25
Hospital and local authority	0	4	2	0.15	0.23	0	0.68	0	1
General practitioner and local authority	250	225	154	17	5	39	51	184	225
General practitioner, hospital and local authority	18	22	15	2	1	8	6	20	12

Total survey population 32,656
(12.3 per cent of population of Stoke)
Total of patients attending general practitioner 24,197
(74 per cent of patients at risk)
Total attending hospital 4,983
(15 per cent of patients at risk)
Total using local or authority services 3,438
(11 per cent of patients at risk)
Total using any service 28,056
(86 per cent of patients at risk)
Total referred by general practitioner 3,402
(10 per cent of patients at risk)
Total shared 4,338
(13 per cent of patients at risk)
∴ Total shared but not referred 936
(3 per cent of patients at risk)
(22 per cent of patients shared)

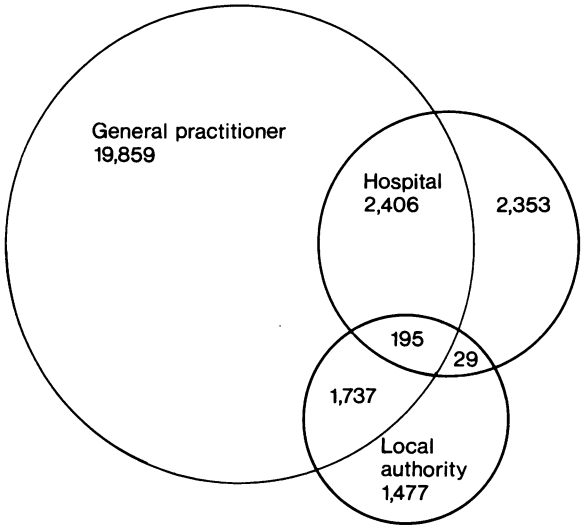

Figure 1. *Record Linkage Study (Stoke) 1964-1965.*

similar to results in other surveys. What is more surprising is the relatively large number of 2,382 (7·3 per cent) additional patients attending hospital departments, who apparently made no contact with the practitioner during the recording year. As has been shown, even when all possible sources of error are taken into account, the rate is still as high as 4·9 per cent of the registered practice population or 32·1 per cent of the patients attending hospital.

Whatever the actual proportion of patients attending hospital and not their general practitioner, it can be inferred that these patients receive virtually their whole medical requirements from the hospital services. There is also an unknown number of patients attending hospital for dental problems who might have had no other reason for attending their general practitioner during the survey year.

The equivalent proportions of patients attending the local authority services are cause for less concern, since they are often specialized services intended to complement those of the general practitioner.

The various totals of patients in each of the categories are given in more detail by age and sex in Table 6.

Discussion

So far as we know this is the first presentation of data of this kind, and it can come only from a linkage study. Where studies are based on records from any one of the main health care agencies, the unshared patients will be excluded. It is clear from this study that these form a relatively large proportion of the patients attending the hospital or local authority services, though a relatively small proportion of any practice population.

The implications of the high ratio of unshared

1,000 registered with the general practitioners.

					Total						
5-14	15-44	45-64	65 +	Total	<1	1-4	5-14	15-44	45-64	65 +	Total
530	625	653	747	630	505	501	529	583	624	792	608
33	106	58	85	79	619	61	45	75	63	103	72
112	19	7	108	49	724	105	112	20	6	91	45
38	111	78	68	81	114	36	47	82	78	85	74
1	1	0	3	1	0	2	2	1	0.11	2	1
148	29	7	63	56	219	225	151	23	6	54	53
8	5	2	11	6	19	17	12	4	2	10	6

hospital patients has been touched on. Fry (Royal College of General Practitioners, 1973), using other data, has also identified this possibility.

In a study of a one in five representative sample of North Lambeth in London, Palmer and his colleagues (1969) estimated that 18·12 per cent of the population attended hospital (including casualty departments) in 1966. This is close to the Stoke figure of 15·3 per cent, but there was no information about the population who did or did not consult their general practitioner.

Morrell and his colleagues (1971), in a study involving three practitioners, noted that 16 per cent of the practice in the survey year referred themselves to hospital without consulting their general practitioner.

Butterfield and Wadsworth (1966) showed that only 68 per cent of the patients attending outpatient services at Guy's Hospital in 1961 were referred by their general practitioner. A further 7·7 per cent arrived first via casualty, 14·2 per cent were internal referrals from other parts of the hospital, and 4·7 per cent were from other hospitals.

Unreasonable use of casualty departments can only be overcome if the hospitals themselves take the initiative, if only because one out of every four 'casual' patients attending casualty departments has made any contact with a general practitioner (Crombie, 1959; O'Flanagan, 1976). Those who have bring with them a written letter from their general practitioner. It seems that if a large proportion of the 47·8 per cent of all patients attending hospital who do not also see their general practitioner are attending hospital unnecessarily, then only the hospital service is in a position to take action.

A further important conclusion to be drawn from this study concerns the basic design of systems for a linkage programme. It is clear that when the error rates are high, the system must be designed so that they are made apparent at the earliest possible moment. The only solution which bypasses the problems encountered in this survey, and also in Acheson's in Oxford (1966), is for all linkage to be made by matching to a primary automated file in which every member of the population studied is identified. Every new item of data for which the patient is inadequately identified is immediately rejected as a non-match and the error can be observed and remedied immediately.

For population studies in health care, the only manageable populations are those represented on general practitioners' lists which, on average, consist of about 2,500 individuals. For populations of this size a Hogben number ensures the minimum number of mismatches and adequate confidentiality of the associated data (Crombie, 1973).

Most primary health care usage involves the general practitioner. On changing their addresses patients will register when primary care is first needed. Apart from children under one year of age, the general practitioners' lists will therefore be as sensitive to population movements as is necessary for health care administration and planning. They will always be more up to date than files or registers based on census data.

As a result of the experience gained in the North Staffordshire Survey, the Second National Morbidity Survey, which began in November 1970 (Crombie, 1973; OPCS, RCGP and DHSS, 1974), was based on an updated automated age-sex register. Hogben numbers were used to identify all patients and as a basis for matching and merging new data as it arrived from the practices where it had been recorded in standard disease indexes. They have also been used to link morbidity data from the Second National Morbidity Survey data file with census data about the same individuals. We believe they could be the basis now for linking records from hospital and general practice. This system has proved to be extremely satisfactory in operation and has bypassed the problems encountered in the Oxford and North Staffordshire Surveys. It has also satisfied the criteria of confidentiality quite rightly demanded.

References

Acheson, E. D. (1966). *Medical Record Linkage.* London: Oxford University Press.

Butterfield, J. & Wadsworth, M. (1966). *Problems and Progress in Medical Care.* London: Oxford University Press.

College of General Practitioners (1963). *Journal of the College of General Practitioners,* **6**, 196-197.

Crombie, D. L. (1959). *Journal of the College of General Practitioners,* **2**, 346-356.

Crombie, D. L. (1973). *Journal of the Royal College of General Practitioners,* **23**, 863-879.

Hogben, L. & Cross, K. W. (1948). *British Journal of Social Medicine,* **2**, 149-152.

Morrell, D. C., Gage, H. G. & Robinson, N. A. (1971). *Journal of the Royal College of General Practitioners,* **21**, 77-85.

Office of Population Censuses and Surveys, Royal College of General Practitioners & Department of Health and Social Security (1974). *Morbidity Statistics from General Practice, Second National Study 1970-71.* London: HMSO. *Preliminary Report* (1973). London: HMSO.

O'Flanagan, P. (1976). *Journal of the Royal College of General Practitioners,* **26**, 154-160.

Palmer, J. W., Kasap, H. S., Bennett, A. E. & Holland, W. W. (1969). *British Journal of Preventive and Social Medicine,* **23**, 91-100.

Peel, J. (1967). *A Study in the Integration of Information from the Various Sources of Medical Care.* Thesis.

Pinsent, R. J. F. H. (1968). *Journal of the Royal College of General Practitioners,* **16**, 127-134.

Royal College of General Practitioners (1971). *Journal of the Royal College of General Practitioners,* **21**, 609-613.

Royal College of General Practitioners (1973). *Present State and Future Needs of General Practice.* 3rd Edition. *Reports from General Practice No. 16.* London: Royal College of General Practitioners.

Acknowledgements

We wish to thank Drs H. W. K. Acheson, Arthur, Boyle, Daly, Fulton, John, Kay, Levens, Robinson, Scott, Simpson, Sweetnam, Wood and Yates F. & G., who recorded practice morbidity, using a diagnostic index, and contributed this data for analysis; Mr John Peel, who developed the linkage system and computer programme, and Mrs Grundy who assisted with the computer analysis.

We also wish to thank, particularly, Dr H. W. K. Acheson who, as recorder of the North Staffordshire (Stoke) Morbidity Survey, was heavily involved in the planning and co-ordination of the projects, and whose contribution was crucial to its success.