Report

THE CONTINUING OBSERVATION AND RECORDING OF MORBIDITY

A Study by the Research Committee of the College of General Practitioners.

Since the end of the war the term " epidemiology " has come to embrace the observation of all disease as it affects the community. In the past the records of mortality were used as indices of the state of health of a community. Such records, however, only gave indirect evidence concerning the diseases which led to death. It was not possible to gain a clear picture of morbidity—the amount of disability in a community at any given time. Many diseases disable but do not kill; their influence on the health and wellbeing of a people can only be estimated by studies of morbidity.

In some communities the pattern of medical practice includes the general medical practitioner, who is the first trained person consulted by a patient who believes that he is ill. The threshold of tolerance to a disease or a symptom varies from patient to patient, but illness may be said to have reached a significant level when an individual consults his family doctor. It has been shown that there is much unreported morbidity below this threshold (Backett, 1953, Spence, 1954) but no studies of this are possible, even though this minor morbidity may contain the seeds of more severe illness.

The field of morbidity which can be observed by the general practitioner begins at the level of the minimal significant symptom and extends through ascending levels of severity until death. Some of his patients—perhaps not more than twenty per cent in any year—may attend hospital wards or departments, where they remain with the knowledge of their doctor, though not under his clinical care. Thus the illnesses which the general practitioner sees do truly represent the total morbidity of significance. His task must be to find means of measuring it.

Previous work on Morbidity Statistics

Attempts have been made to obtain information on morbidity at general-practitioner level from other sources, but these have failed for lack of completeness. A limited number of infectious diseases are statutorily notifiable by Acts of Parliament (1931, 1938) although there is no means of checking the accuracy of these reports. Conclusions may be drawn from causes of disability notified under National Health Insurance regulations, but this information comes from a limited section of the population and the standard of diagnostic accuracy encountered may vary considerably. Reliance can

seldom be placed on material collected for one purpose when attempts are made to use it for another. Social surveys by Gallup Poll techniques have been tried (Gen. Reg. Office 1949), but lay interpretation of medical diagnostic terms makes these unsuitable for serious study.

The field is clear, then, for studies of morbidity by the general practitioners who encounter it themselves in their daily work. The first studies were by individuals who independently observed and recorded in their own practices. (Pinsent, 1950; Fry, 1952; Horder, J. and Horder, E., 1954; McGregor, 1950; Crombie, 1955. They reported studies of morbidity in their differing types of practice. Sometimes recorded observations were comparable between practices but more often they were not, and the next development was the observation and recording by small groups of general practitioners. Logan (1953) was the first to conduct a survey of general-practice morbidity in a group of 8 practices in which material was collected over a two year period for central analysis. This work showed that reliance could be placed on records made in general practice, and that mechanical means could be applied to their interpretation. A survey of mental illness, using a group of general-practitioner observers was carried out by Desmond O'Neill (Finlay, B., *et al.*, 1954) and other techniques of obtaining information by sampling were tried out in the United States. The Washington Sickness Survey and the study of general medical practice in North Carolina (Osler D. Peterson *et al.*, 1956) both sought to relate general-practice morbidity to the International Classification of Diseases and Causes of Death prepared by the World Health Organization.

The success of the pilot survey by Logan (1953) in a small number of general practices led to the larger-scale National Morbidity Survey of 1955-56 conducted, as a combined undertaking, by the General Register Office and the College of General Practitioners. In this study the observer field was extended to include 108 general practices—120 doctors in all—and records were made of every item of service carried out during one year. A record card was designed which fitted into the National Health Service medical record envelope and an entry was made at each attendance. Doctors were invited to use diagnostic terms to which they were accustomed, and their observations were coded and checked later by a large clerical staff. A simultaneous investigation was made concerning the occupations of patients whose illnesses were recorded during the year. The large quantity of material—which it is believed will give information statistically representative of the population of England and Wales—is now being analysed. The picture which will

be built up following analysis of this National Morbidity Survey will show the pattern of illness in England and Wales in great detail but only for one year—the period covered by the survey. A base-line for study will have been drawn which should be of great value, but the one-year survey can tell us little about the way in which evolution is taking place in human morbidity.

Changes are taking place in both disease and health care all the time. To find an indication of the change in types of diseases, and the rate of their change, surveys of this kind need to be repeated at intervals which might, in Britain, be related to the decennial census. Repeated one-year surveys of total morbidity on this scale would be costly and their results less revealing than a simpler but continued study. Had our predecessors been required to determine the course of pneumonia from temperature readings made on the first and tenth days of the disease, the two forms of recovery, crisis and lysis, would never have been differentiated, and so we may learn lessons from continuing observation and recording of the morbidity of the community, which may escape a series of separated studies even though these may be comparable between themselves.

The continued observation of morbidity has been undertaken in a few general practices for some time. Fry, 1952; McGregor, 1950; and Hodgkin, 1957, have independently designed methods for day-to-day recording of morbidity and means of analysing it by the use of punch-card systems. These practices, in some of which records have been maintained for several years, are among the few in which we can observe the ebb and flow of diseases of different kinds.

Two important lessons were learned from the conduct of the National Morbidity Survey. The first is that an observer corps of general-practitioner workers could be obtained which was broadly representative of the general-practitioner population of the country, and the second that these practitioner observers were able to keep accurate records without detriment to the medical care which their patients received. These family doctors showed that the continuing study of morbidity was practicable.

Present Study

INTRODUCTION

This paper is concerned with the problems of definition and of diagnosis. It is the report of a trial of method rather than of a pilot morbidity survey, but considers also the classification of the diagnoses used in general practice. This latter problem will be dealt with in greater detail in a second paper.

In 1953 the Research Committee of the Council of College of General Practitioners set up a working party to consider the

establishment of a Records Unit as part of a statistical advisory service to field research workers. This working party considered the problems which might be met in a continuing study of morbidity by a group of doctors. Basic units of measurement had to be decided upon; whether the illness-experience of an individual, the doctor-patient contact, or the episode or " spell " of illness, should be used. A diagnostic classification, a method of recording of data compatible with practice routine, and a central organization to receive, analyse and " process " material were necessary.

METHOD

After full consideration the episode of illness was chosen as the unit of measurement for the trial. These episodes must be identified and a diagnostic code was required, since complete freedom of nomenclature would lead to unworkably bulky material. Diagnostic codes used by individuals, other research organizations, and by the World Health Organization were examined, and as none seemed appropriate to the circumstances of general practice in this country a new and provisional adaptation of the W.H.O. code under 71 headings, was devised (see Table V).

The number of headings chosen was a compromise between the minimum of 17 used by Fry (1947) and the 240 used by Logan (1953). Where the diagnosis is coded by coding clerks and the practitioner has only to insert the diagnosis, the number of headings is dictated only by the analytical capacity of the central organization. Where the practitioner has also to insert the code number of the disease process a further limitation is placed on the number of headings. In the event there were certain obvious omissions, such as rubella and mumps, which would have to be included in any future classification.

This limitation was also necessary because much morbidity in general practice presents as symptom-complexes which, while they can be adequately treated, cannot be classified easily as diseases. Multiple diagnoses are also often made in general practice, such as when an acute illness supervenes upon a chronic. The need to find out more about these inter-relationships was recognized.

It was realized that " between-observer " error would be hard to eliminate, and that there would be need for an assessment of a degree of accuracy which could be achieved within the terms of the chosen diagnostic classification. It was further recognized that the end of an illness is often not so easy to determine as its beginning.

The Research Committee Working Party co-opted Dr K. W. Cross of the Department of Medical Statistics, Birmingham

University, and with his help and with the help of the panel of twelve a record card was designed. The same practitioners later took part in the pilot survey. The cards were " self-coding " for analysis by Powers-Samas machinery. They were distinctively coloured and shaped to fit the National Health Service medical record envelopes. They contained spaces for the recording of basic information relating to both practitioners and patients, and columns in which a maximum of five entries might be made by one stroke at each consultation. On completion of the illness data as to its duration and the other diagnoses made during the episode were inserted, and the card was returned to the centre for punching and analysis. A serial number on the card and on the Medical Record Envelope ensured continuity of recording of patients through consecutive illnesses. (Figures 1 and 2.)

All the practitioners taking part in this experiment had had experience of collection of morbidity material from their practices either in the National Morbidity Survey or in their own investigations. Every patient presenting with a new condition in a calendar month in the spring of 1957 was allotted a serial number and a card on which each attendance was recorded until the episode ended. No further registration of new cases took place in the second month, but episodes beginning in the first were watched and notes made until the end of the survey—three months from the start. At the end of each episode the card was completed, the serial number recorded on the medical record envelope and the card despatched to the analysis centre. Space on the card allowed the doctor to note difficulties in its completion, at the time at which these were experienced.

One practitioner withdrew from the trial during the first month on account of the administrative difficulty he found in reconciling the survey card routine with his own current punch-card routine, the smooth running of which he did not wish to disturb. One practitioner undertook to relate his diagnosis to the full W.H.O. classification of diseases, as a special study of the validity of the diagnostic terms selected. The comments of all who took part were consolidated at the end of the trial. Each practice was analysed separately so that differences between observers would become apparent.

FINDINGS

The survey is a trial of methods of recording and a study of between-observer errors and differences of interpretation. The comparisons made are between practitioners in their approach to the problem of recording information and not between the character of their practices, about which only general comments have been

made. Such figures as are given are concerned with the qualitative rather than quantitative aspects of morbidity.

The number of illnesses cared for by each of the eleven doctors during the survey is given in Table I, together with the numbers of surgery attendances and visits. The ratio of surgery attendances

APPENDIX

COLLEGE OF GENERAL PRACTITIONERS

MORBIDITY SURVEY—*RECORD CARD FOR PILOT SURVEY*

To be used at each consultation and completed for each episode of illness.

Doctor's Name and Code Number ..... 1. ☐

Patient's Name ..... 2-4. Serial Number ☐☐☐

		Single	Married	Widowed	Sept. or Div.	Not Known
5 Sex and Marital Status	Male	1	3	5	7	9
	Female	2	4	6	8	10

6 Age-group (years)								
Under 1	1-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39
0	1	2	3	4	5	6	7	8
40-44	45-49	50-54	55-59	60-64	65-69	70-74	75 and over	
9	10	11	12	13	14	15	16	

7-8 Date of first consultation for this episode ☐☐ Day Mth.

To be completed at end of episode

9 Number of attendances at surgery ☐ 10 Number of visits ☐

11-12 Date of last consultation for this episode. (If episode not finished at end of survey, *leave blank*) ☐☐ Day Mth.

13-14 Duration of episode of illness (days). (Code 00 to 99 ; if over 99 days code as 99 and specify .....) ☐☐

15 Number of different *primary* diagnoses recorded during episode ☐

REMARKS .....

.....

NOTE.—If more than 8 consultations are made for *one episode* continue with a second card. Complete items 1-8 on both cards but leave items 9-15 blank on second card. Please staple together cards relating to same episode.

Figure 1. Front of record card

to visits varied considerably from practice to practice; the ratios for the two Birmingham practices agreed closely and, as would be expected, far exceeded those of practices in less densely populated areas. On the other hand the mean number of consultations per episode of illness was surprisingly constant.

To be completed at each Consultation

NOTE.—Except for coding the primary and other diagnoses (and the specifications), indicate the answers by placing a tick in the appropriate space.

Number of Consultations	1st	2nd	3rd	4th	5th	6th	7th	8th
Place of Consultation	Col. 16	Col. 22	Col. 28	Col. 34	Col. 40	Col. 46	Col. 52	Col. 58
Surgery 1								
Patient's Home 2								
Other 3								
Primary Reason for Consultation	Col. 17	Col. 23	Col. 29	Col. 35	Col. 41	Col. 47	Col. 53	Col. 59
Advice or treatment for pathological condition 1								
Certificate—medical 2								
Certificate—non-medical 3								
Repeat prescription 4								
Prophylactic procedures 5								
Relatives' anxiety 6								
Other (specify .....) 7								
Primary Diagnosis	18 19	24 25	30 31	36 37	42 43	48 49	54 55	60 61
When code number prefixed by X specify actual condition in space provided								
Accuracy of Primary Diagnosis	Col. 20	Col. 26	Col. 32	Col. 38	Col. 44	Col. 50	Col. 56	Col. 62
No diagnosis 0								
Tentative diagnosis 1								
Exclusion of serious alternative diagnosis 2								
Firm diagnosis 3								
Other Diagnoses	Col. 21	Col. 27	Col. 33	Col. 39	Col. 45	Col. 51	Col. 57	Col. 63
None 0								
Number of: One 1								
Two 2								
Three or more 3								
Insert in spaces provided the code numbers of these diagnoses								

Figure 2. Back of record card

TABLE I

NUMBER OF EPISODES, ATTENDANCES AND VISITS BY DOCTORS DURING PERIOD OF SURVEY

Practice	*Practice population*	*No. of episodes*	*No. of survey attendances*	*No. of visits*	*Total no. of consultations*	*Mean no. of consultations per episode*	*Ratio S.A. to V.*
Ilford	Urban 3163	437	532	297	829	1.6	1.8
Birmingham 'A'..	Urban 6700	203	299	52	351	1.7	5.8
Redcar	Mixed 2417	496	917	354	1271	2.5	2.6
Ipswich	Mixed 1680	264	356	225	581	2.2	1.6
Burley	Rural 2400	292	321	286	607	2.1	1.1
Worcester ..	Mixed 6000	262	322	232	554	2.1	1.4
Alford	Rural 3600	144	200	98	298	2.1	2.0
Birmingham 'B' ..	Urban 2500	203	369	59	428	2.1	6.0
Plymouth ..	Urban 3000	365	644	187	831	2.3	3.4
Felsted	Rural 2500	262	339	190	529	2.0	1.8
Ibstock	Mixed 3000	236	358	158	516	2.2	2.3
All practices ..		3164	4657	2138	6795	2.15	2.2

Almost half the total number of episodes required a single consultation, and in a further quarter only two consultations were necessary. Table II shows that there was a high degree of consistency between most of the doctors. The proportions of episodes decreased rapidly as the consultation rate increased, and form a rough geometrical progression with a ratio of half.

In general practice some items of service are carried out on grounds of administrative need or convenience and these may influence the amount of " morbidity " recorded. Medical certification is a particular source of inflation. The distribution of episodes of illness according to the primary reason for the first consultation was examined.

Table III shows that the primary reason for the first consultation in 91.4 per cent of episodes was for advice or treatment. The proportions, varying from 85.4 per cent to 96.1 per cent, show

TABLE II

DISTRIBUTION OF EPISODES ACCORDING TO NUMBER OF CONSULTATIONS

		Number of consultations								*Total episodes*
		1	2	3	4	5	6	7	8 *or over*	
Ilford	Nos.	239	109	39	20	13	7	3	6	436
	%	*54.8*	*25.0*	*8.9*	*4.6*	*3.0*	*1.6*	*0.7*	*1.4*	*100.0*
B'ham 'A'	Nos.	127	38	19	11	4	2	1	1	203
	%	*62.6*	*18.7*	*9.4*	*5.4*	*1.9*	*1.0*	*0.5*	*0.5*	*100.0*
Redcar	Nos.	137	173	90	32	21	15	4	17	489
	%	*28.0*	*35.4*	*18.4*	*6.5*	*4.3*	*3.1*	*0.8*	*3.5*	*100.0*
Ipswich	Nos.	130	47	34	23	12	9	1	4	260
	%	*50.0*	*18.1*	*13.1*	*8.8*	*4.6*	*3.5*	*0.4*	*1.5*	*100.0*
Burley	Nos.	156	58	31	22	5	6	6	6	290
	%	*53.8*	*20.0*	*10.7*	*7.6*	*1.7*	*2.1*	*2.1*	*2.1*	*100.0*
Worcester	Nos.	110	67	36	18	7	6	6	2	252
	%	*43.6*	*26.6*	*14.3*	*7.1*	*2.8*	*2.4*	*2.4*	*0.8*	*100.0*
Alford	Nos.	69	38	20	6	5	2	3	1	144
	%	*47.9*	*26.4*	*13.9*	*4.2*	*3.5*	*1.4*	*2.1*	*0.7*	*100.0*
B'ham 'B'	Nos.	80	68	26	16	9	3	1	—	203
	%	*39.4*	*33.5*	*12.8*	*7.9*	*4.4*	*1.5*	*0.5*	—	*100.0*
Plymouth	Nos.	165	89	40	27	15	8	3	13	360
	%	*45.8*	*24.7*	*11.1*	*7.5*	*4.2*	*2.2*	*0.8*	*3.6*	*100.0*
Felsted	Nos.	141	61	21	13	11	3	2	5	257
	%	*54.9*	*23.7*	*8.2*	*5.1*	*4.3*	*1.2*	*0.8*	*1.9*	*100.0*
Ibstock	Nos.	124	41	32	11	9	6	3	7	233
	%	*53.2*	*17.6*	*13.7*	*4.7*	*3.9*	*2.6*	*1.3*	*3.0*	*100.0*
All practices	Nos.	1478	789	388	199	111	67	33	62	3127
	%	*47.3*	*25.2*	*12.4*	*6.4*	*3.5*	*2.1*	*1.1*	*2.0*	*100.0*

consistency, and indicate the comparative unimportance of all other reasons for consultation. Of the episodes seen in Plymouth, nearly 5 per cent were primarily for medical certification; while the Birmingham 'B' and Felsted practices had corresponding figures of 3.5 per cent and 3.9 per cent. The differences in the proportions from doctor to doctor probably reflected differences in the assessment of the necessity for the patient's attendance rather than in any difference in the proportions of patients receiving medical certificates of different kinds. The anxiety of relatives was the primary reason for consultation in 2 per cent of all episodes, although the Alford

TABLE III

DISTRIBUTION OF EPISODES ACCORDING TO PRIMARY REASON FOR FIRST CONSULTATION

		Primary reason for first consultation							*Total episodes*
		Advice or treatment	*Med. Cert.*	*Non-Med. Cert.*	*Rep. Presc.*	*Prophylactic proc.*	*Relatives anxiety*	*Other*	
Ilford	Nos.	416	6	3	—	5	2	4	436
	%	*95.4*	*1.4*	*0.7*	—	*1.2*	*0.4*	*0.9*	*100.0*
B'ham 'A'	Nos.	179	5	2	5	4	3	5	203
	%	*88.1*	*2.5*	*1.0*	*2.5*	*1.9*	*1.5*	*2.5*	*100.0*
Redcar	Nos.	448	5	1	6	7	8	14	489
	%	*91.6*	*1.0*	*0.2*	*1.2*	*1.4*	*1.6*	*2.9*	*100.0*
Ipswich	Nos.	239	2	2	2	10	3	2	260
	%	*91.9*	*0.8*	*0.8*	*0.8*	*3.8*	*1.1*	*0.8*	*100.0*
Burley	Nos.	265	1	—	—	10	11	3	290
	%	*91.4*	*0.4*	—	—	*3.4*	*3.8*	*1.0*	*100.0*
Worcester	Nos.	238	2	2	—	6	3	1	252
	%	*94.4*	*0.8*	*0.8*	—	*2.4*	*1.2*	*0.4*	*100.0*
Alford	Nos.	123	1	—	4	1	13	2	144
	%	*85.4*	*0.7*	—	*2.8*	*0.7*	*9.0*	*1.4*	*100.0*
B'ham 'B'	Nos.	178	7	1	2	10	4	1	203
	%	*87.7*	*3.5*	*0.5*	*1.0*	*4.9*	*1.9*	*0.5*	*100.0*
Plymouth	Nos.	320	17	—	4	6	11	2	360
	%	*88.9*	*4.7*	—	*1.1*	*1.7*	*3.1*	*0.5*	*100.0*
Felsted	Nos.	227	10	—	2	9	4	5	257
	%	*88.3*	*3.9*	—	*0.8*	*3.5*	*1.6*	*1.9*	*100.0*
Ibstock	Nos.	224	2	3	—	1	2	1	233
	%	*96.1*	*0.9*	*1.3*	—	*0.4*	*0.9*	*0.4*	*100.0*
All Practices	Nos.	2857	58	14	25	69	64	40	3127
	%	*91.4*	*1.9*	*0.4*	*0.8*	*2.2*	*2.0*	*1.3*	*100.0*

practice had a proportion as high as 9 per cent. Prophylactic procedures accounted for 2 per cent.

Corresponding distributions relating to the last consultation were calculated, and these differed very little from the above.

Observer error

In any accurate study of morbidity, knowledge of the existence and extent of between-observer error, and error in diagnoses by the same observer at different times is essential. A part of the card

was designed to investigate this, and the accuracy of the primary diagnosis was indicated by four degrees of precision: 0—no diagnosis made; 1—tentative diagnosis; 2—exclusion of serious alternative diagnosis; 3—firm diagnosis.

These degrees of precision were decided on, after much discussion, as being those most appropriate to the conditions of general practice, and to the habits of thought in regard to diagnosis which are acquired by general practitioners.

Table IV shows that the diagnosis remained unaltered in *91 per cent of all episodes* while multiple primary diagnoses were recorded in 5 per cent of all episodes. Episodes in which " no diagnosis " was recorded accounted for 3.6 per cent of the total and this proportion varies considerably from practice to practice, due partly at least to differences in interpretation of terms used. In the Ilford, Redcar and Alford practices there were about 10 per cent of episodes with multiple primary diagnoses or in which the original diagnosis was later changed, whereas the remainder had very few.

The large number of episodes in which the diagnosis remained unaltered indicates that continued reporting of morbidity in terms of one primary diagnosis only, might be reliable providing that the occasions on which " no diagnosis " is made are recorded. This would greatly simplify the completion and analysis of the cards. The analysis of cards on which it was recorded that the primary diagnosis was later changed showed that many were pregnancies in due time changed to " delivery ". In the majority of the remaining episodes where a change was made in the primary diagnosis an entirely new diagnosis was entered, and in only a few instances was a secondary diagnosis at the first consultation elevated to the primary position on a subsequent consultation.

Almost 30 per cent of all diagnoses made at the first consultation were tentative, and 55.5 per cent were firm. No diagnosis was made in 8 per cent of all episodes. The use of the category " exclusion of serious alternative diagnosis " differed considerably from practice to practice. None of the diagnoses of the Ibstock practice was so classified, and in three other practices there were proportions of under 3 per cent. In most practices about half of the diagnoses were " firm " at the first consultation; in the Birmingham 'A' practice only 25.6 per cent were, while in Redcar (72.4 per cent) and Ipswich (69.2 per cent), the figures were much higher.

Accuracy of Diagnosis

The maximum degree of accuracy achieved at the first consultation by any of the doctors, as measured by his " firm " diagnoses, was 72.4 per cent. All the practitioners were meeting a very similar sample

TABLE IV

DISTRIBUTION OF EPISODES ACCORDING TO NUMBER OF PRIMARY DIAGNOSES, ACCURACY OF DIAGNOSIS AT FIRST CONSULTATION AND NUMBER OF OTHER DIAGNOSES AT FIRST CONSULTATION

		No. of P. Diagnoses				*Accuracy*					*Other Diagnoses*				
		0	1	2	3+	0	1	2	3	*N.S.*	0	1	2	3	*N.S.*
Ilford	Nos.	1	391	39	5	24	153	34	225	—	389	43	4	—	—
	%	*0.2*	*89.7*	*8.9*	*1.2*	*5.5*	*35.1*	*7.8*	*51.6*	—	*89.2*	*9.9*	*0.9*	—	—
Birmingham 'A'	Nos.	9	188	4	2	14	123	11	52	3	156	39	6	2	—
	%	*4.4*	*92.6*	*2.0*	*1.0*	*6.9*	*60.6*	*5.4*	*25.6*	*1.5*	*76.9*	*19.2*	*2.9*	*1.0*	—
Redcar	Nos.	12	425	48	4	11	84	28	354	12	389	95	4	1	—
	%	*2.5*	*86.9*	*9.8*	*0.8*	*2.2*	*17.2*	*5.7*	*72.4*	*2.5*	*79.6*	*19.4*	*0.8*	*0.2*	—
Ipswich	Nos.	16	239	5	—	35	38	7	180	—	249	11	—	—	—
	%	*6.2*	*91.9*	*1.9*	—	*13.5*	*14.6*	*2.7*	*69.2*	—	*95.8*	*4.2*	—	—	—
Burley	Nos.	8	274	7	1	44	60	28	155	3	233	50	6	1	—
	%	*2.8*	*94.5*	*2.4*	*0.3*	*15.2*	*20.7*	*9.7*	*53.4*	*1.0*	*80.4*	*17.2*	*2.1*	*0.3*	—
Worcester	Nos.	2	235	15	—	28	84	3	137	—	220	31	1	—	—
	%	*0.8*	*93.5*	*5.9*	—	*11.1*	*33.3*	*1.2*	*54.4*	—	*87.3*	*12.3*	*0.4*	—	—
Alford	Nos.	3	123	16	2	13	48	11	70	2	113	28	2	1	—
	%	*2.1*	*85.4*	*11.1*	*1.4*	*9.0*	*33.3*	*7.6*	*48.6*	*1.4*	*78.5*	*19.4*	*1.4*	*0.7*	—
Birmingham 'B'	Nos.	16	181	6	—	17	71	13	99	3	188	11	2	—	2
	%	*7.9*	*89.2*	*2.9*	—	*8.4*	*34.9*	*6.4*	*48.8*	*1.5*	*92.6*	*5.4*	*1.0*	—	*1.0*
Plymouth	Nos.	18	341	1	—	28	99	10	223	—	342	16	2	—	—
	%	*5.0*	*94.7*	*0.3*	—	*7.8*	*27.5*	*2.8*	*61.9*	—	*95.0*	*4.4*	*0.6*	—	—
Felsted	Nos.	23	232	2	—	19	52	40	135	11	237	16	2	1	1
	%	*8.9*	*90.3*	*0.8*	—	*7.4*	*20.2*	*15.6*	*52.5*	*4.3*	*92.2*	*6.2*	*0.8*	*0.4*	*0.4*
Ibstock	Nos.	6	222	5	—	14	113	—	106	—	178	53	2	—	—
	%	*2.6*	*95.3*	*2.1*	—	*6.0*	*48.5*	—	*45.5*	—	*76.4*	*22.7*	*0.9*	—	—
All Practices	Nos.	114	2851	148	14	247	925	185	1736	34	2694	393	31	6	3
	%	*3.6*	*91.2*	*4.7*	*0.5*	*7.9*	*29.6*	*5.9*	*55.5*	*1.1*	*86.1*	*12.6*	*1.0*	*0.2*	*0.1*

of morbidity, and the variations between this upper figure and the lowest of 25.6 per cent represent differences both in the interpretation of the clinical picture and in the value of the diagnostic label attached to it. The fact that the average for the proportion of " firm " diagnoses was only 55.5 per cent, and the maximum achieved was 72.4 per cent, indicates the inadequacy of our diagnostic armamentarium, and hence of any conclusions based on its use. We must, it seems, examine the diagnostic habits of doctors with great care, and bear these wide differences in mind when we interpret any figures collected from general practice.

There is clearly a difference of opinion on what constitutes a firm diagnosis, and on the degree of accuracy of clinical observation and pathological investigation as carried out and interpreted by different doctors. While absolute accuracy in diagnosis is as unobtainable as exact standardization between observers, a degree of consistency can be maintained within the standards chosen. Further work on this aspect of the problem is necessary.

In addition there is the separate problem of the accuracy of the diagnosis itself. While a " chill " may be an acceptable diagnosis to a layman, " respiratory tract infection " will mean a little more to a trained observer, and " febrile catarrhal sinusitis of haemolytic streptococcal origin " may be a more detailed and more accurate definition of the same disorder. An attempt was made, therefore, to test the accuracy in use of a diagnostic classification adapted from the W.H.O. code. Symptom-complexes are often sufficient in themselves to suggest a rational therapy, and it is a quality of general practice that at times the diagnostic process may not be carried beyond this point.

The *Concise English Dictionary* defines " diagnosis " as " identification of disease by means of patient's symptoms, etc., formal statement of identification ". In terms of this strict definition a diagnosis is a label. A pathologist would probably insist on a more clear-cut definition, while general practitioners regard a " diagnosis " as a stepping stone towards the better management of the clinical problem facing them. To the pathologist the establishment of a diagnosis is an end in itself. Consequently, conditions encountered by pathologists will be capable of clearer definition than those which are not. Most conditions encountered at hospital level belong to the first category; those most commonly met in general practice belong to the latter. The W.H.O. code is concerned primarily with morbidity at hospital level and therefore reflects the closer definition which more advanced investigation makes possible.

There are two points concerning the accuracy with which any given clinical condition can be labelled. Firstly there is the general

accuracy in terms of aetiology and pathology implied by a diagnostic label, and secondly the accuracy with which this label is applied. In this study accuracy has been interpreted as the summation of these two and it is obvious that the importance attributed to each of them varied from practitioner to practitioner and from one diagnosis to another. For instance the Birmingham 'A' and 'B' and Ibstock practitioners classed most of the conditions labelled " cough " as tentative while the Redcar practitioner classed them as " firm ". The former were relating accuracy to the strict meaning of the diagnostic label employed, while the latter was describing the accuracy with which the label represented the clinical condition. It can be assumed however, that the highest proportions for " firm " diagnosis given in the survey—the Redcar practice 72.4 per cent—represents the maximum degree of accuracy attainable by any interpretation, while 55 per cent forms a good working average. Table V gives the average proportions of firm diagnoses of all practitioners for each diagnostic category.

Distribution of Episodes by Diagnosis at First Consultation

Episodes of illness seen by each doctor during the month of the survey were analysed by " diagnosis ". Table V gives the results for all practices combined, and also shows the number of consultations for the episodes specified in each diagnostic category. The mean number of consultations per episode varies from 1.4 for " depression " to 4.5 for " chronic bronchitis " (ignoring those mean values based upon very small numbers of episodes). Other diagnostic categories with ratios of four or more were:—delivery, pneumonia, and diseases of the liver, biliary system and pancreas.

The percentage distribution of episodes recorded shows that five diagnostic categories: acute nasopharyngitis (10 per cent); skin disorders (6.5 per cent); other E.N.T. conditions (6.4 per cent); psychoneurotic disorders (5.1 per cent), and cough (4.6 per cent) together account for almost one-third of all episodes; acute bronchitis also makes a relatively large contribution. E.N.T. conditions account for one-fifth, skin disorders for nearly one-tenth (9.4 per cent), and injuries for 7.5 per cent of the total episodes. Symptom-complexes such as " cough," " headache," " senility," but excluding " backache " and " abdominal pain," accounts for 9.2 per cent of the total. Of the 71 categories only 23 contributed more than 1 per cent to the total number of episodes.

Accuracy of Diagnosis at First Consultation

The distribution of all episodes seen by each practice with respect to accuracy of diagnosis has been given on Table IV. Corresponding

TABLE V

DISTRIBUTION OF EPISODES AND CONSULTATIONS BY DIAGNOSIS (ALL PRACTICES)

Diagnostic Code No.	*Diseases or conditions*	*No. of episodes*	*No. of consultations*	*No. of consultations per episode*	*% Distribution per episode*	*% Distribution of consult'ns*	*Proportion of firm diagnoses*
01	Tuberculosis of respiratory system	3	24	(8.0)	0.1	0.4	(0.0)
02	Tuberculosis, other forms	1	5	(5.0)	0.0	0.1	(0.0)
03	Whooping cough	7	13	(1.9)	0.2	0.2	(28.6)
04	Measles	46	128	2.8	1.5	1.9	69.6
05	Malignant neoplasms	5	14	(2.8)	0.2	0.2	(20.0)
06	Other neoplasms	5	10	(2.0)	0.2	0.2	(60.0)
07	Asthma	17	44	2.6	0.5	0.7	70.6
08	Obesity (not specifically as of endocrine origin)	28	57	2.0	0.9	0.9	78.6
09	Metabolic disorders (including gout)	13	35	2.7	0.4	0.5	61.5
10	Pernicious anaemia	2	3	(1.5)	0.1	0.1	(100.0)
11	Other types of anaemia	10	31	3.1	0.3	0.5	(40.0)
12	Other diseases of the blood	—	—	—	—	—	—
13	Psychoses	7	25	(3.6)	0.2	0.4	(57.1)
14	Psychoneurotic disorders	161	303	1.9	5.1	4.5	60.9
15	Other psychogenic illness	29	64	2.2	0.9	1.0	69.0
16	Vascular lesions affecting C.N.S.	8	35	(4.4)	0.3	0.5	(75.0)
17	Migraine	13	27	2.1	0.4	0.4	46.2
18	Other disorders of C.N.S.	15	43	2.9	0.5	0.6	53.3
19	Sciatica, disc, lumbago and backache	41	90	2.2	1.3	1.3	12.2
20	Unspecified neuralgia and neuritis	19	33	1.7	0.6	0.5	21.1
21	Diseases of the eye	90	134	1.5	2.9	2.0	62.2
22	Otitis media	67	151	2.3	2.1	2.3	88.1
23	Acute nasopharyngitis	312	545	1.8	10.0	8.1	55.4
24	Acute tonsillitis and scarlet fever	70	186	2.7	2.2	2.8	75.7
25	Other E.N.T. conditions	200	362	1.8	6.4	5.4	73.5
26	Influenza	56	153	2.7	1.8	2.3	21.4
27	Pneumonias (and pneumonitis)	17	72	4.2	0.5	1.1	41.2
28	Bronchitis—acute	116	342	3.0	3.7	5.1	75.9

Continued on pages 122 *and* 123

TABLE V—*continued*

Diagnostic Code No.	*Diseases or conditions*	*No. of episodes*	*No. of consultations*	*No. of consultations per episode*	*% Distribution per episode*	*% Distribution of consult'ns*	*Proportion of firm diagnoses*
29	Chronic bronchitis	11	49	4.5	0.4	0.7	72.7
30	Rheumatic fever	—	—	—	—	—	—
31	Chronic rheumatic heart disease	1	6	(6.0)	0.0	0.1	(0.0)
32	Coronary disease and angina pectoris	5	8	(1.6)	0.2	0.1	(40.0)
33	Other diseases of heart	12	39	3.3	0.4	0.6	58.3
34	Hypertension	19	47	2.5	0.6	0.7	68.4
35	Varicose veins	24	50	2.1	0.8	0.7	91.7
36	Haemorrhoids	7	8	(1.1)	0.2	0.1	(85.7)
37	Peptic ulcer	25	64	2.6	0.8	1.0	40.0
38	Appendicitis	6	20	(3.3)	0.2	0.3	(33.3)
39	Hernia of abdominal cavity (diaphragmatic)	4	8	(2.0)	0.1	0.1	(100.0)
40	Disease of liver, biliary system and pancreas	33	131	4.0	1.1	2.0	42.4
41	Abdominal pain and other abdominal disorders	80	185	2.3	2.6	2.8	31.3
42	Chronic nephritis and nephrosis	4	14	(3.5)	0.1	0.2	(25.0)
43	Acute urinary tract infection	35	94	2.7	1.1	1.4	28.6
44	Disorders of menstruation	16	36	2.3	0.5	0.5	56.3
45	Menopausal symptoms	10	22	2.2	0.3	0.3	(70.0)
46	Other gynaecological conditions	25	61	2.4	0.8	0.9	56.0
47	Pregnancy	39	120	3.1	1.3	1.8	71.8
48	Delivery	11	46	4.2	0.4	0.7	90.9
49	Puerperium	4	15	(3.8)	0.1	0.2	(75.0)
50	Boil and carbuncle	47	94	2.0	1.5	1.4	93.6
51	Cellulitis of finger and toe	21	44	2.1	0.7	0.7	90.5
52	Other cellulitis and abscess	23	51	2.2	0.7	0.8	78.3
53	Skin disorders	204	368	1.8	6.5	5.5	69.6
54	Rheumatoid arthritis	10	28	2.8	0.3	0.4	(50.0)
55	Osteo-arthritis	19	31	1.6	0.6	0.5	63.2
56	Arthritis, unspecified	12	23	1.9	0.4	0.3	33.3

Continued on page 123

TABLE V—*concluded*

Diagnostic Code No.	*Diseases or conditions*	*No. of episodes*	*No. of consultations*	*No. of consultations per episode*	*% Distribution per episode*	*% Distribution of consult'ns*	*Proportion of firm diagnoses*
57	Rheumatism (unspecified) fibrositis	52	102	2.0	1.7	1.5	26.9
58	Congenital malformations	5	8	(1.6)	0.2	0.1	(60.0)
59	Certain diseases of early infancy	8	10	(1.3)	0.3	0.2	(62.5)
60	Cough	143	226	1.6	4.6	3.4	36.4
61	Diarrhoea (age 2 years plus) and/or vomiting	35	65	1.9	1.1	1.0	17.1
62	Depression	14	20	1.4	0.5	0.3	50.0
63	Headache	11	16	1.5	0.4	0.2	0.0
64	Senility	10	29	2.9	0.3	0.4	(60.0)
65	P.U.O.	43	133	3.1	1.4	2.0	4.7
66	Dyspepsia	28	41	1.5	0.9	0.6	3.6
67	Sprains and strains, etc.	98	198	2.0	3.1	3.0	50.0
68	Lacerations, superficial injury, contusion abrasion and crushing	90	190	2.1	2.9	2.8	93.3
69	Burns	7	14	(2.0)	0.2	0.2	(85.7)
70	All other injuries including fractures	39	104	2.7	1.3	1.6	53.8
71	All other illness	359	778	2.2	11.5	11.6	56.8
	None	120	195	1.6	3.8	2.9	—
	All diseases	3127	6720	2.2	100.0	100.0	55.5

(*The figures in brackets are based on* 10 *or less cases*)

distributions were calculated for each diagnostic category and each doctor, and those relating to the nine most common diagnoses have been examined separately. However, even though the overall numbers are 90 or more, those relating to single doctors are often too small to merit separate analysis.

As would be expected, a diagnosis of " laceration," " superficial injury," etc., was a firm one in all but very few such episodes; whereas " sprains and strains " was recorded as a firm diagnosis in half of the episodes so diagnosed, 41 per cent being tentative.

" Acute nasopharyngitis " was a firm diagnosis in 56 per cent of all such episodes and 37.5 per cent were noted as being tentative; the proportion in which a serious alternative was excluded was 5 per cent. As regards " cough," the accuracy was not recorded in 14 per cent of such episodes and almost equal proportions (37 per cent and 36 per cent) fell into the categories tentative and firm respectively. In 12 per cent of the episodes, a serious alternative diagnosis was considered to have been excluded. As regards both " other E.N.T. conditions " and " acute bronchitis," a firm diagnosis was made in about three-quarters of the episodes so diagnosed and a tentative one in about one-fifth of the episodes.

The category " exclusion of serious alternative diagnosis " was used more often, relatively, in the case of " psychoneurotic disorders " than for any other of these diagnostic categories, the proportionate contribution being 14 per cent. Of the episodes so coded, 21 per cent and 61 per cent were considered to be tentative and firm diagnoses.

Similar proportions of tentative and firm diagnoses were obtained for " diseases of the eye," although the proportions of such episodes in which a serious alternative was excluded was 9 per cent.

Table V also gives the proportion of firm diagnoses recorded by all practitioners combined for each diagnostic category. Proportions based upon less than 10 episodes are in parenthesis. Mention has previously been made of the differing interpretations put upon accuracy by the doctors concerned, but these proportions show which diagnoses are uniformly interpreted either as having a low or high specificity. It is high in the following:—otitis media, varicose veins, delivery, boil and carbuncle, cellulitis of finger and toe, lacerations, etc.—conditions which can be clearly defined on sight. On the other hand, vague states such as " sciatica," " disc," " lumbago " and " backache," " unspecified neuralgia," " influenza," " abdominal pain " and " other abdominal disorders," " acute urinary tract infections," " arthritis " and " rheumatism " and other symptom-complexes, have exceptionally low values.

Duration of Illness

It had been hoped that an estimate could be made of the duration of disabilities observed during the course of the survey, but this was not achieved. The beginning of an episode of illness may be easy to identify either as the date of the first symptom or the date on which the first attendance took place. The end of an episode, however, was more often than not impossible to define. Recovery is a gradual process; return to work, to school, or to resumption of full activity being no more than stages.

Practical difficulty arose also after consultations at which the patient was advised to return to the doctor if he felt no better. This advice is frequently given in general practice, and a convention was adopted whereby episodes consisting of a single consultation were regarded as being of one day's duration only. Estimates based on the days between first and last attendance were also, it was realized, inaccurate. It seems that duration of illness is not a suitable yardstick against which to measure disability in general practice.

Discussion

Before an organization can be set up to undertake the continuing observation of morbidity in the community, there are certain pre-requisites:—there must be a network of observers distributed in such a way as to represent the community with statistical accuracy; there must be methods of recording which are accurate and yet simple, so that accuracy may be maintained during periods of working stress, and there must be means for rapid analysis of recorded material and its translation into useful terms. While the present survey was mainly concerned with the second of these, consideration was given to the other factors also.

The network of general-practitioner observers would be made up from those on the Research Register of the College of General Practitioners (Watts and Cross, 1957). While some observers would contribute records representing the whole of the morbidity they met, others would observe and record in greater detail a limited aspect of disease in which they were particularly interested. There is now an increasing number of general practitioners with experience of morbidity study, and their co-operation will be encouraged by care in the design of the methods of recording they will be asked to use.

It is envisaged that cards would be returned to the co-ordinating centre at predetermined intervals, and that the centre would be equipped and staffed to handle a continuous flow of information. The central organization would be sufficiently flexible to allow

special studies of limited scope but greater detail to be carried out simultaneously with the collection and analysis of routine morbidity figures. The amount of recorded material would be great, and the organization would need to use modern electrical equipment. It might be possible to complete the analysis of the returns for one week by the end of the week after, thus allowing quick publication of information in suitable medical journals.

In any continued study of morbidity the difficulties encountered in this pilot enquiry would be taken into account. Many of these difficulties arose because several items were put on the card for experimental purposes; and will not necessarily be included in future.

An important decision to be made in the light of the pilot study is the best unit of observation. It may be more satisfactory to have one card per patient instead of one card per episode of illness. The advantage would be that the patient's number, age, sex and marital status would be recorded once only during a year. This procedure would also avoid the necessity of bringing together all episodes referable to the same patient. It is sometimes difficult to decide in a chronic illness whether a consultation after an elapse of time should be recorded as a new episode or a re-attendance for the same episode as previously indexed.

It is evident from the pilot survey that it is very difficult to determine the end of an episode accurately and it is concluded that the duration of episodes is not worth recording. Many episodes involved only a single consultation and the duration of other episodes varies considerably. Attempts to assess duration were inaccurate and the collection of the information was time consuming.

Certain amendments and improvements in the design of the card was suggested. In any continuous enquiry the doctors' code numbers would be pre-printed on the cards, and, if possible, repetition of patients' age, sex and marital state on cards for different episodes would be avoided. Provision would be made for the doctor to enter the date at each consultation. The total number of consultations would be recorded on the card, in addition to the number of attendances and the number of visits. Provision could also be made for recording the admission of patient to hospital and, possibly, other referrals also.

It was sometimes difficult to decide upon the primary diagnosis where more than one condition was being treated. The diagnostic code used in the pilot survey will have to be revised in the light of the results, and many useful observations were made by the doctors. It might be possible to avoid repetition of a code number where

there is no change in primary diagnosis during an episode. Another important point to be decided is whether the coding should be carried out by the participating doctors or done centrally by lay staff from specified diagnoses.

To increase man's knowledge of himself he must study the circumstances of his life as well as those which cause his death. Mortality statistics have so far been our major clue to the morbidity preceding death, and the accuracy of figures derived from this source is now being questioned. Schemes of statutory notification of defined communicable diseases are of less importance now that interest is focussing on less well-defined and often unnamed virus infections and the degenerative diseases. We must find means to measure the terms in which we diagnose disease, and our own accuracy in the use of those terms. We must learn to record, analyse, and to interpret. When we have learned to do these things we can harness for the common good the vast potential for observational research which lies in general practice in this country, in the Commonwealth, and wherever else high standards of practice are to be found.

Summary

A study of methods of continued recording of the morbidity met in general practice has been carried out by the research committee of Council of the College of General Practitioners. Twelve general practitioners took part.

A trial was made of the use of a diagnostic classification adapted from the World Health Organization code, and an assessment was made of the accuracy with which this could be used.

Certain differences between the interpretations of the observing practitioners have been noted. These are concerned with the accuracy with which a diagnostic label describes a clinical condition. It is proposed to consider this further in a later paper.

It is believed that an organization could be set up to study continuously the morbidity seen in general medical practice, and that information gained in this way would be invaluable in the study of the natural history of disease and in the planning of means to mitigate it.

This study was designed by, and undertaken on behalf of, the Research Committee of Council of the College of General Practitioners.

Those taking part were:—R. N. Compton Smith (Ilford), D. L. Crombie* (Birmingham), J. Fry (Beckenham), G. H. K. Hodgkin (Redcar), R. Horn (Ipswich), C. R. G. Howard (Burley), A. J. Laidlaw (Worcester), W. J. H. Lord (Alford), R. J. F. H. Pinsent* (Birmingham), R. Simpson-White (Plymouth), P. A. Walford (Felsted), C. A. H. Watts* (Ibstock), and Dr K. W. Cross* (Department of Medical Statistics, Birmingham University).

*Planning Working Party.

REFERENCES

Backett, M., (1953), *Proc. roy., Soc. Med.*, **46**, 707.

Spence, Sir J., "A Thousand Families in Newcastle-on-Tyne." J. Spence, W. S. Walton, F. J. Miller and S. P. M. Court. London, Sumberlage for the Nuffield Foundation, 1954.

Public Health Act, 1931. Food and Drugs Act, 1938.

Registrar General's Statistical Review of England and Wales (1949). Supplement on General Morbidity, Cancer and Mental Health.

Pinsent, R. J. F. H., (1950). *Lancet*, **1**, 917.

Fry, J., (1952). *Brit. med. J.*, **2**, 249.

Horder, J. and Horder, E., (1954). *Practitioner*, **173**, 177-187.

McGregor, R. M., (1950). *Edin. med. J.*, **57**, 433.

Crombie, D. L., (1955), M.D. Thesis, University of Birmingham.

Logan, W. P. D., (1953). General Practitioners Records G.R.O., H.M.S.O., London.

Finlay, B., *et al.*, *Practitioner*, **172**, 183.

Peterson, Osler L., *et al.*, (1956). An Analytical Study of North Carolina general practice. *J. med. Educ.* Vol. 31, No. 12, Part 2.

Hodgkin, G. K., (1957). Personal communication.

Watts, C. A. H. and Cross, K. W., (1957). *Brit. med. J.*, **1**, 1112.

Grundy, Prof. F. and Fanning, Lewis (1957). Morbidity and Mortality in the first year of life. Eugenics Soc., Lond.

New Preparations

Ames Company (London) Limited announce the introduction of two new simple diagnostic tests.

"ALBUSTIX" Reagent Strips. The first colour test for urine protein. An impregnated paper stick is dipped into the urine—an immediate colour indicates a positive test and this can be compared with the colour scale as a guide to the amount of protein present. Even cloudy urines can be tested without filtration. "ALBUSTIX" is ideally suited for routine urine testing as one test only takes a few seconds.

Supplied in bottles of 60 sticks.

"HEMATEST" Reagent Tablets. A standardized o-tolidine test for occult blood in faeces. The sensitivity is of the same order as that of the Gregerson test, but avoids the use of benzidine, now regarded as carcinogenic. Where a more sensitive test is required "OCCULTEST" Reagent Tablets are recommended. Both tests are suitable for use in clinics and wards as well as in laboratories.

Supplied in bottles of 50 tablets.