WORLD HEALTH ORGANIZATION

Regional Office for the Eastern Mediterranean

ORGANISATION MONDIALE DE LA SANTÉ

Bureau régional pour la Méditerranée orientale

SEMINAR ON THE APPLICATION OF PSYCHIATRIC EPIDEMIOLOGY
Khartoum, 17 - 21 February 1975

EM/SEM.APPL.PSY.EPID./7

ENGLISH ONLY

METHODOLOGY OF DATA-COLLECTION

UTILIZATION OF ROUTINE STATISTICS FROM MENTAL AND OTHER HEALTH SERVICES

by

Miss Eileen M. Brooke *

1. What are statistics ?

People talk about the collection and utilization of statistics, but may have very vague ideas of what is actually involved. Doctors, in particular, tend to be disappointed that statistics do not tell them about individual patients. Statisticians often show a tendency to be satisfied if they have got some numbers, without worrying too much about their accuracy or completeness, and may regard their task as finished when they have produced a set of tables, irrespective of how, when and where the latter may be used, by whom and for what purpose.

It must be stated :

(i) that statistics is a classifying science

(ii) that it tells us about groups, namely those which have been produced by classifying the basic material.

Statistics are therefore particularly useful in relation to the health services, which have to be planned for groups of people and not for every single individual separately. The treatment and care of a single patient depends upon the way in which what has been planned for a group can be slightly modified locally to fit individual needs.

2. The need for economy in statistics

A physician does not take a temperature with a faulty thermometre, nor weigh the ingredients of a medicine with a balance, part of which is missing. Statistical tools call for the same precision, but do not always receive it. The effects of services are often measured with tools,

(*) Chief, Department of Medical Statistics, University Institute of Social and Preventive Medicine, Lausanne, Switzerland

parts of which are missing, as for example the private psychiatric services where they exist. A surgeon will not cut you open with a blunt knife, but the question of the sharpness of a stisistical tool for any given purpose is not so often discussed. Steam engines may quite likely be used to crack nuts, while nutcrackers may be ineffectively applied in an effort to crush rocks. Neither system is economical. When one looks at the statistical tables which are so often prepared on a routine annual basis, one is glad that the statisticians responsible are not doctors, for they would give you the entire contents of the medicine chest instead of the one or two medicaments necessary for your condition.

Statistics, like drugs and instruments,are expensive. For some countries, even the cost of paper for printing data-sheets is a consideration. Costs come in the form of paper, printing, dispatching data-sheets to the services, training those who are going to fill them in, paying people for doing this job, sending parcels back to the statistical office, checking, processing, printing, advertising and selling reports. There is the cost of machine equipment and the cost of time. Consequently there is also the need for the greatest economy at all stages of the data-collection and processing system.

Hence the first essential for data-collection, which is again the first step in the production of a statistical tool, is to define very clearly what you want the tool to do when you have it. This will enable you to decide the best materials, shape and size for the purpose it will also enable any superfluous parts to be eliminated. A kidney machine is excellent for a certain purpose, but is not necessarily improved by receiving a pianola attachment. The more parts there are to a tool, the more there are to go wrong. And it is useful to remember the law which says 'If anything can go wrong it will !' Computers are constantly our of action whereas there are not many people who have lost one of their ten fingers. It is also easier to teach people to count with their fingers than with a computer. Let economy be the watchword, and on the same principle that you do not buy an extra coat at sale-times because 'it may come in useful later', do not collect extra data -- like the coat they will probably be out of fashion by the time you come to use them.

3. Formulation of the questions to be asked

In this paper we are concerned with mental and other health services; what do we want routine statistics to tell us about them ? And by routine we no doubt mean annual statistics, although for the sake of economy we would do well to consider whether they might not be periodi

some data being collected only every third year, for example. In the first place a plan for the services will have been formulated. Hence we should like the statistics to hold up a mirror to the services, and show exactly what they are doing. This means that it is not sufficient to collect data about patients, but also about services. So that while we may make patients the basis of some statistics, and describe them in terms of various forms of classification, we shall also have to make treatment facilities the basis of other statistics and describe them in terms of who uses them and who works in them, and thirdly we shall have to make personnel the basis of yet other statistics and describe them in terms of who they care for and where their work is carried out.

However, the patients have not become ill, nor have the services been developed, without reference to an environmental setting. And while we have given a great deal of care and attention to the purification of water supplies, the eradication of mosquitoes, and the detection of pathogenic elements in industry, for example, the study of the physical and living environment in the production of mental disorder remains, except in a few instances such as lead encephalopathy, particularly underdeveloped. The statistical study must therefore be complemented by environmental ones. Nor, when we come to the second important stage, that of interpreting the statistics, can we confine ourselves to the present moment. The services we have to-day are largely the product of systems which have developed over more than a hundred years. In Europe they have in many cases left us with hospital systems whose present shape is reminiscent of that of a prickly pear plant -- a basic root with bits stuck on here and there. In other parts of the world either copies have been made of the European system or there are folk medicine and traditional healers whose effects have not been sufficiently evaluated for us to affirm that they are doing a good job but which we hesitate to jettison in case they should be doing better, for their patients, than some of the so-called Western systems.

Our questions then, to be answered if possible by routine statistics, are :

(i) what are the services doing ?

(ii) are they carrying out the plan which was formulated ?

(iii) now that we see the results do we want -- or rather are we obliged -- to change the plan ?

Under these three headings we might consider the problems of heroin addiction and methadone maintenance. A letter to the Journal of the American Medical Association (229:8, 19 Aug., 1974)

from Dr P. Baconsfield of the Royal Free Hospital of Medicine, London begins : "The problem of treating heroin addicts continues to go unresolved. The supporters and detractors of methadone maintenance programmes produce more and more evidence to support their divergent views. The sobering fact remains that we are no nearer success in the treatment of heroin addiction than we were a decade ago".

This brings us back to the question of the need for economy in statistics, and for thinking carefully in advance of data collection, what items to include and how to classify them. If we are considering services, we are unlikely to get much valuable epidemiological data since the patients at whom we are looking are the ones who come to the services. Genetic questions may be among those which are better left out of routine statistics, since although, for example the first-born in families might appear to be a high risk group for some pathological condition, it is too late for us to do anything about changing that circumstance now. We would therefore be wiser to limit the statistic to those items which do influence the type of service to be given, and also to things which it is in our power to alter, and among these would be forms of service, and types of personnel and their training.

Variables which might produce group classifications in routine statistics are sex, since desegregation has made relatively little progress; age, because while it is in many places thought wise to treat children and old people apart from adults, there is the possibility of carrying out controlled experiments to see whether these two special groups do in fact do better separated from each other and from the rest of patients, or whether by putting all ages together we do not produce an interaction which is beneficial to all. Marital status may have a claim to be considered, not with a view to trying to find epidemiological high risk groups, which it is unlikely to do without supplementary questions, but because we may wish to find out what other services, at present unprovided or not utilized, are necessary to maintain some people 'hospitalised at home', for which the existence of a healthy and willing spouse, might form a good basis.

When we come to such items as occupation, education and religion, so often included in data-collection systems, we are on less firm ground. In the first place these three items may not be independent of each other; certain religions may preclude certain occupations and certain types of education, notably mixed schools for boys and girls, Level of education and social class of occupation may be positively correlated for some illnesses, but inversely

correlated for others. In economy of data collection we may place these variables, not in items to be included in routine statistics of services, but among those many data which the psychiatrist or social worker needs for his work with the individual patient. Another drawback of these variables is that they are usually classified in a way suitable for preparing demographic statistics of entire populations, most of whom are free from any kind of pathology, psychiatric or other, and there is nothing to suggest that the same grouping will be suitable for classifying persons with rheumatism, cardiovascular disease or anxiety neurosis. In fact a large scale research project is being carried out in Lausanne to try to find the best way of recording occupation for mental patients. Nor is it possible to decide upon a relation of cause and effect, since we do not know whether a personality, predisposing to mental disorder, for example, influenced the choice of a particular occupation in the first case.

Where admission to and discharge from services is subject to law, as is usually the case with psychiatric disorders, it can be useful to include a question on the various sections of the law, to see how they are working. In the revised British legislation a section was included making it possible to detain a person for 72 hours on the advice of one doctor only -- this breathing space was intended to apply to acute cases, breaking up the existing situation and enabling the psychiatrist to decide on the next move, which would normally be either to detain the patient for 28 days or persuade him to accept treatment on a voluntary basis. Routine statistics, however, showed that in one hospital region, these violent acute patients were largely in their eighties and nineties, thus suggesting that this section of the Act was being used to get an old person into a mental hospital bed.

Another question which it may be useful to consider in relation to patients, is who sent them to the psychiatric services. In countries where the traditional place of treatment is a mental hospital, the tendency may be to continue with this practice even when other services, such as units in general hospitals, out-patient clinics and day care have been introduced, so that the hospitals remain overcrowded, while the new services are under-utilized.

This emphasizes also the need for studying the utilization of all services simultaneously, rather than in isolation, and on a regional basis. Where the services themselves are organized on a regional basis, this will be relatively easy. In other cases it may be necessary to try to divide up the territory into suitable units and to group the services arbitrarily for study purposes, using natural boundaries, such as rivers or mountains, which it is difficult to cross.

Very often the same administrative units are being used for general purposes, for fiscal purposes and for various types of medical services. However, if we are going to place a team skilled in treating burns, or in treating facture cases, in a regional unit, different considerations will influence the determination of the catchment area than if we are providing for tuberculosis of psychiatric cases. By studying the geographical distribution of cases and the urgency of getting them to hospital, useful data can be obtained for the siting of new services or the conversion of existing buildings for new purposes.

4. Routine statistics relating to services

One of our questions here will no doubt be 'which categories of patients are using different types of services ?' and when we have the answer to this, we may ask whether this distribution is the best possible. And here the word 'best' is a relative term. Very of the burden of numbers and lack of finance reduces us to the need for striking a balance between what is best for the patient, who should be the first person to be considered, and his family, whose point of view must also be taken into account, and what is best for the hospital, the doctors and the nurses -- giving rise, for example, to a discussion on whether labour should be induced in order to regulate the flow of cases or to allow women to spend Christmas with their families.

In particular, for psychiatric services, it is important to have simple regional comparisons of the kinds of patients that are being cared for or treated in residential or ambulatory services, especially where it is the declared policy to treat people as far as possible in the community and to use the residential services as 'back-up' beds. In this case we must have data on the community services which must be provided for the purpose, such as special equipment, non-slip floor covering, community nurses, meals on wheels, or laundry services. Although this sounds an expensive procedure, it may be less expensive than maintaining a person in an institutional setting, and by using the patients' families after giving them some initial training and small payments, we may be surprised at the results. For the collection of routine statistics, important questions here will be : (a) do you consider this patient is in the right service ? (b) If not, where do you think he should be placed ? (c) What services would have to be provided to enable him to remain in his own home ?

For ambulatory services, an important question is the times during the week when the service is open to receive patients. Here again there is an opportunity for a controlled

experiment, by changing the hours of opening and observing the effect on patterns of utilization of the services. Whereas hospitals may admit patients at any hour of the day or night, out-patient clinics may tend to restrict their hours of opening to mornings and afternoons. The effect of evening opening might be that more patients could remain at work during the daytime. Anti-smoking clinics have been known to increase their clientèle by this means.

5. Statistics of Personnel

Here one of the most important problems that we have to face in setting up a routine data-collection system is that of classification. Are we going to classify people according to the qualifications they present in terms of diplomas, experience, etc., or are we to classify them in terms of the kind of work they do ? The former method may be unfair to older workers who began their careers at a time when there were no diplomas issued; the latter involves knowing, or more likely finding out, the job content of each category of personnel. Then there is the question of what weighting to give to length of practical training or length of experience working in the field. Placing these people in relation to the services in which they are employed and the type of patients with whom they can work may be enormously important in countries suffering from a shortage of trained personnel. For example, how much responsibility can be given to nurses where there is an acute shortage of psychiatrists ? We have to find (a) simple definitions which are easily applicable for classifying personnel, (b) simple statistics which show what kind of patients the different categories of personnel can deal with and (c) simple patient statistics, enabling us to compare the outcome of their work.

6. Steps in setting up a routine data collection system

(i) As already stated, the first thing is to decide what we want the statistical tool to do; what questions have we, what areas do we want to investigate. Here everything may be gained by getting away from our traditional, administrative-epidemiology types of statistics and re-thinking the whole process. Countries which have not got a statistical system may be envied, as they have nothing to destroy before building afresh.

(ii) Definitions and classifications are essential and at present this is one of our weakest areas. A glossary of psychiatric terminology has been devised, which should yield great dividends in the form of the better classification of psychiatric diagnosis (WHO, Geneva, 1974). However, when it comes to services, such as hospital, out-patient clinic , day hospital, day centre, and to personnel such as qualified psychiatrist,

occupational therapist, health visitor, social worker, we do not yet enjoy a similar means of standardization.

(iii) Decision on the minimum amount of data necessary to answer the questions we have posed.

(iv) Development of data-collection sheets : a task which sounds much easier than, in fact, it is. Points to be remembered are that the data should be easily obtainable, that they should be objective and hence definable, as stated in (ii). So far as possible, the subject matter of individual boxes on the data sheet should be independent, while the items listed in any box should be also independent of each other, but cover all possibilities. The lay-out of a data sheet is important; it should be as easy to fill in as possible. Precoded items are easy to tick and the information can be transferred straight away to punched cards; do not clutter up the data sheet with unnecessary details, for example by putting squares to be checked when you have supplied code numbers as well :

e.g.	✓ 1 Masculine	and not	☒	Masculine 1
	2 Feminine		☐	Feminine 2,

Be sure to state on the data sheet the service from which it comes, for example Ministry of Health, or Abercorn Psychiatric Hospital, otherwise it may not find its way back to you.

(v) Writing instructions for those who are going to be asked to fill in the data-sheets is another essential task. It is often when you come to write the instructions that you find out the weaknesses of your Data sheet. With the instructions correctly written it should be possible for anyone to fill in the same data in the same way, thus getting uniformity. Nothing should be left to the judgement of the individual when it comes to filling in the data sheet.

(vi) A small pilot study should be carried out, to test the data sheet and the instructions. At this point you may get very useful criticism and suggestions from people who are actually working in the services and who know very well what they can and cannot get. If possible, explain to them in person what it is you are trying to do and why. They will be more likely to take trouble if they see that you are prepared to do the same.

(vii) The results of the pilot study will show you how to improve and sharpen your statistical tool even farther. You are then ready to start the main study. Although data are usually collected for calendar years there is nothing sacrosanct about January 1 as a day for starting data collection.

(viii) Checking coding and data processing : all data sheets should be checked immediately on receipt for missing items or entries which are obviously incorrect; and referred back to the senders as quickly as possible. The aim should be to get an even flow of data through the various processes, so that tabulations can be produced as quickly as possible after the last data sheets have been received.

(ix) Tabulations : since these are destined to answer the questions you raised in the first place and since the data collected are for that purpose, the tabulations can be designed while data collection is going on. Remembering the need referred to at the outset of this paper for economy in statistics, we remind ourselves that as few tables as possible consistent with answering the questions posed should be prepared. Few people can assimilate more than three variables in one table, and for those who do not grasp the meaning of numbers easily, graphs may be much more useful, especially when not cluttered up with too many lines.

(x) Speed is a primary essential in producing statistics of medical or social services; they need to come 'hot from the press' and not two or three years after the event. Here it is indeed a case of 'better half a loaf of bread to-day, than a whole loaf to-morrow'! If the statistics show you that some services are not doing what was intended, the sooner this fact is known and changes are made, the better; otherwise they tend to become 'traditional' and therefore unalterable.

(xi) Feedback of data to those who have produced the raw material is essential if good relationships are to be maintained. A short commentary for each region, for example, can earn a great deal of goodwill from those who are going to be asked to repeat the same work year after year.

7. What to do with the data

We have now dealt with two out of the three points listed in last year's meeting at Addis Ababa, namely :

(1) what to count

(2) how to count.

There remains the third one, which is :

'what to do with the count ?'

In the first place it may be desirable to turn the counts into rates, by dividing by a suitable denominator. This reinforces the need for economy in statistics, since it is hard to compare data when there is no demographic information on which rates can be based. If 'caravan sites' is included in the list of places from which people are admitted to hospital, then a count of zero admissions may either mean no caravan sites, or plenty of caravan sites used by people who are free of the disease in question.

When the data for (a) patients, (b) services and (c) personnel are available, it will be time to subject them to an intense scrutiny in terms of the questions asked. If hospital admissions or residents are decreasing, is this because more people are being treated in out-patient clinics or as day-patients ? We look at the Service Statistics. In order to discharge the patients as quickly as possible back to the community, is it better to have a high proportion of psychiatrists to patients or a high proportion of nurses ? We look at the personnel statistics. In this way we may slowly but surely approach the ultimate goal of evaluation of the services. But it is important to remember that we are only dealing with statistics of those patients who come to the services, and that as long as there is a significant number of people being treated in private facilities or not being treated at all, we cannot make any generalizations from the statistics we have collected. Other papers at this meeting will show how to deal with the general aspects of case-finding and of community surveys.

8. Where to start

With what some may regard as so many counsels of perfection it may be found difficult to know where to start. Since hospitals are still the chief dispensers of services, we might conveniently choose them as the starting point. It would be difficult to find a hospital that did not have some count of admissions, discharges, deaths and numbers of resident patients, or a bed occupancy figure. A routine statistical system may develop in two ways from this beginning. As we have already decided on our questions and the data to be collected to answer them, so we may gradually build into the data collection system of the individual hospital, those items which are not already being collected. And here is another advantage of economy in statistics, in that we shall not immediately submerge the hospital staff with demands with which they have neither the manpower nor the training to cope. At the same time we may proceed in the inverse sense, by dividing up the data already being collected for example by ward, or by specialty or by months and thus within the hospital itself making a statistical mirror

which will show all those working there what is being done. Similarly we may help the hospital staff to break down their working time into its component parts and see what tasks are being done and in what time, and thus cutting out those tasks which are 'traditional' but not always necessary.

Whereas the gradual building up of routine statistics within a hospital, then within all services in the way referred to above in para ( ), and finally by region and nationally, will enable us to arrive at national statistics which are useful for overall evaluation and planning, so the breakdown of data within the hospital will provide the basic elements for a hospital activity analysis. The presentation of data will differ. The number of psychiatric beds per 1 000 population or the number of orthopaedic surgeons per 100 000 population are very crude indicators, apart from the fact that many people will pass through life without wanting the services of either. Nevertheless, they do give us some general comparisons, useful for rather large hospital regions. Statistics in hospital activity analysis will be related to the bed or the personnel as units; how many admissions per bed, how many beds per psychiatrist, how many untrained to trained nurses. Statistics built up in the first way, by accretion, will form the subject of published reports. The discussion of elementary activity statistics and their further development within any given hospital, should be confidential to the hospital.

9. Who is going to do the work ?

What is required at the service level for the task of building up routine statistics is someone who can count and is rather obsessional about getting the right answer. It is astonishing how many such people there are in hospital offices, often frustrated and not working to the highest level of their capacity. The writer once found a clerk who was maintaining a complete in-patient psychiatric register at the provincial level, because he said it enabled him to check the annual counts. He did not realize the value of what he was doing ! Care must be taken to give such people instruction in how to present their data in a way which will interest the staff and stimulate discussion. They will form the category designated 'statistical clerk' by WHO. At the national level are required those workers whom WHO has designated 'statistical technicians'. Their task will be to put into effect instructions given by administrators about the data to be collected, to carry out data collection and analysis and to write simple reports. For persons with secondary education a year's training, including practical work, was proposed. It is surely time to start implementing these proposals.

10. What are the difficulties in the collection and use of routine statistics for Health Planning ?

The first and perhaps most important factor is that planners and statisticians do not see that they have a commitment to a joint task. There may be little direct communication, either because the administrators and statisticians are on different levels of the official hierarchy, or because the statisticians are segregated in special agencies or branches. Until these difficulties in communication can be overcome, each will remain in ignorance of the other's methods, skills and experience. It is to overcome some of these disabilities that the Workshops or Training Courses in the European Region of WHO have been introduced.

The training of statisticians and that of administrators, especially if they are doctors, are quite different. Often poor statistics emanate from a doctor turned statistician (untrained), we should get similar poor results if statisticians tried to treat the patients. Each must stay within his own area, while learning to pull together. On the other hand, planners do not realise how long it takes to produce good statistics, nor the hours of work that may go to the preparation of one small table.

Administrators and planners must clearly enunciate their goals, otherwise the statistician cannot develop a suitable statistical system. He must be given sufficient authority and a high enough rank to be able to insist on receiving the explanations he requires. Too often university-trained statisticians are used as mere data-gatherers, with a role hardly as advanced as that of the 'statistical technician' referred to above.

Finally, there is a need for much greater flexibility in both services and statistical systems, with a willingness to experiment in the provision of services and to constantly seek new and better data and especially more meaningful forms of presentation of statistics.