

Auditing and evaluation in general practice*

KENNETH MOURIN, M.A., M.R.C.G.P., D.R.C.O.G.

General Practitioner, Dereham, Norfolk

After a quarter of a century of a National Health Service in the United Kingdom many doctors are increasingly concerned about the efficiency of its function, about the effects of external pressures modifying medical practice, about ill-informed criticism, and about our own need to adapt and change with advancing knowledge. We need constantly to re-evaluate what we are doing and how we are working in order to make sure that the maximum benefit goes to the maximum number of people—if possible giving the most return for the least cost in terms of manpower, facilities, and money.

Definitions vary, but in this article I have used the term 'audit' to include both the process of enumeration (whether it be the 'medical audit' of mortality and morbidity statistics, and workloads, or the 'clerical audit' which determines and scores the adherence to pre-set criteria) and evaluation, where there are value judgments of A in preference to B, whether these are personal, by groups of peers, or by external 'expert' bodies.

The aims of audit

We are professionally accustomed to gathering information from our patients about the effectiveness of our procedures and therapy; we look with a critical eye on the claims of enthusiastic protagonists of various causes, weighing them in the light of our previous knowledge and experience. Auditing techniques can enable us to gather specific and much more objective facts from our own practice to present in numerical terms for comparison with others. These statements of facts should be capable of external validation and reproduction, and of forming the basis for debate and the judgment of ourselves and others.

Any critical look at the organisation, processes, or outcomes of medical care must lead us to consider whether we can in any way modify the system or the processes to give better results—and these modifications can then themselves be tested and compared.

There are three component parts of audit:

- (1) Setting standards,
- (2) Assessing performance,
- (3) Modifying clinical practice.

They serve the two main purposes of audit: firstly, that of education, and secondly, that of planning.

Auditing

(1) *The practice profile*

This term is used to describe the analysis of the practice and its population. To compare practices, one needs to know the area and the type of practice—whether urban, semi-rural, rural, whether dispensing or not: the numbers of medical, nursing, and non-nursing staff, whether nurses, health visitors, midwives, and social workers are 'attached', whether special clinics are held, whether there are full, partial, or no appointment systems.

*A summary of a report of an Upjohn Travelling Fellowship on auditing and evaluation which was awarded the first Upjohn Prize.

The demography of a practice—the patterns of population, their characteristics in age, sex, race, social class, marital status—define practice populations, making the calculation of prevalence and incidence rates at once simple and useful. Using these ‘profiles’ one can identify variables which may be the cause of marked differences in patterns of workload or morbidity in later studies.

(2) *Practice activity analysis*

Most of the customary methods of analysing workload are fully described by Eimerl and Laidlaw (1969) in the *Handbook for Research in General Practice*. Drug use and interactions (Durno, 1973), drug sensitivities, handicap registers, consultation rates, use of services, and high frequency users (Gruer, 1972) are among other items studied.

(3) *Package programmes*

Data-recording sheets have been devised allowing for individual observation of the doctor’s or his practice activity by himself, giving numerical analyses of exactly how much primary medical care is given according to simple criteria. Standardisation of technique aids objective comparison within local peer-groups, and permits collective analysis by computer programme and the development of models of general-practice activities against which to measure the quality of care (Research Unit, Royal College of General Practitioners, 1973).

Similar studies by Fraser *et al.* (1974) and Patterson *et al.* (1974) have investigated hospital and laboratory referrals, making it possible to indicate aspects for improvement in communications or facilities, with evidence in numerical terms against which to measure local change or comparative studies elsewhere.

(4) *Disease study programmes*

A tentative application of similar data collection sheets is suggested, to collect information in numerically coded form about a series of diseases. For instance, diabetes might be analysed according to initial diagnostic data, years diagnosed, how identified, type, presence of optic, neurological, vascular or renal changes, type of medication, criteria for control, and follow-up interval.

The criteria of quality may be used in different problems, e.g., functioning ability, or other measures of outcome or process.

(5) *Delay pattern analysis*

Hodgkin (1973) has shown how objective evidence of time intervals before patients present with their symptoms, the time taken for the general practitioner to suspect the correct diagnosis, and the total delays to starting treatment, can be transformed into indices for analysis and comparison.

(6) *Morbidity surveys*

Studies of the quality of health care raise such questions as: ‘How many patients are identified with specific diseases, complaints or problems in a particular practice, and how do these rates and numbers show local and regional variation?’

The Second National Study of Morbidity Statistics from General Practice (1974) provides such information from 137 doctors spread over the country. Surveys of screening procedures are necessary to identify the ‘iceberg’ of non-presented and presymptomatic disease (Wilson and Jungner, 1968).

(7) *Accessibility studies*

The availability and accessibility of health care is one factor affecting quality. The changes resulting from a move into a new health centre, providing pathological specimen transport schemes and transport systems for patients, have been studied by the Health Services

Research Unit of the Department of Health and Social Security at the University of Kent at Canterbury (1973). Interestingly, although for small practices health centres were able to provide better staff, facilities, and equipment, in larger practices the practitioners themselves willingly provided such services to an even greater extent.

(8) *Facilities*

The effects of introducing practice nurses, and workers with other skills by employed and attached staff, freeing the doctor for longer and better employment of his own special skills and techniques, has been well demonstrated by Marsh (1969), Weston Smith (1970), and McGuinness (1970).

Irvine's (1972) study of the facilities available in teaching practices gives a most useful index against which one can assess whether one is providing and using equipment which might improve the quality of care given.

The *Design Guide for Medical Group Practice Centres* (1967) attempts to formulate standards of space, lighting, heating, and furniture. The ambience of the practice matters just as much: light, airy, colourful surroundings induce satisfaction in patient and doctor more readily than decrepit, neglected, out-of-date ones.

(9) *Time and motion studies*

Buchan and Richardson (1973) have applied work-study techniques to general practice for the first time, recording time intervals for specified processes going on within the consultation. Further studies are in progress of the activity of the practice nurse, receptionist, and secretary, and the district nurse and health visitor. Judgments as to which time for any particular procedure is correct remain to be determined; and the effects of change or differences between practices need further study.

Evaluation

So far the methods of numerical analysis have been looked at, which may be able to give a good idea of what is going on, and perhaps impressions as to its value. *Measuring* how effective medical intervention is becomes much more difficult, and requires a whole series of new techniques.

(1) *Process and outcome*

Williamson (1971) has drawn the distinction between process and outcome. If, he argues, a diagnosis is correct and adequately stated, then the diagnostic process which leads to that outcome (diagnosis) can be assumed to be adequate. Similarly, if management outcomes match acceptable norms—for example, if the expected proportion of acute coronaries survive—then the therapeutic regimes or processes can be assumed to be adequate.

This concept helps us to identify those areas which most urgently need review or retraining, for we cannot audit every part of every diagnostic or therapeutic process immediately.

Williamson suggests that criteria can be set from pooled experience, from expert advice, or from review of the literature, and limits of confidence for the figures can be set. For instance, if in acute coronary occlusions the criterion of maximum acceptable death rates is set by consensus of experienced physicians at 30 per cent, and 60 per cent of patients die, then a study of the processes of medical care would be indicated. If 31 per cent die, this would be within a 95 per cent confidence limit, and further action would not necessarily be needed.

(2) *Outcome studies*

One of the major difficulties in considering evaluation of medical care is that of deciding what criteria of outcome are to be used; for without a standard of outcome, how can we tell whether our intervention has been effective, ineffective, or positively harmful?

There are a few diseases where accurate numerical figures for pathological investigations can be obtained which describe precisely the desired end result—for instance, the mean corpuscular volume after vitamin B₁₂ therapy. Sometimes one can apply scales of functioning ability, such as the Crichton scale for geriatric assessment.

In others, as Cassel (1974) has said, we must pose the question: "What are the criteria which would lead you to believe, and others to believe, that intervention had resulted in change in the appropriate direction?"

Such standards are difficult to formulate and validate.

(3) *External standards*

There are a few external standards available against which we can judge our activity. The criteria for good antenatal and postnatal care are written into the maternity service regulations of the National Health Service, percentage acceptance of immunisation for infants at appropriate times and intervals can be calculated; the selection and examination of patients requiring contraceptive advice are detailed explicitly in the *Handbook of the Family Planning Association* (1973).

Curtis-Jenkins (1974) and his colleagues are testing and validating a series of paediatric developmental cards, which have great promise for comparing individuals and practices, as well as having enormous potential for planning services required for handicapped children.

(4) *Cohort reviews: thyroid disorders*

The selection of relatively small groups of individuals can enable us to study the effectiveness of our medical care, providing we have adequate measures of outcome.

In my own practice in Norfolk, I have studied 35 patients with present, previous, or potential thyroid dysfunction, in a pilot study (Mourin, 1974). The criterion was simply stated: patients should be maintained in a euthyroid state. The calculated free thyroxine index (F.T.I.), derived from the serum T₃, the uptake test, and the serum T₄ level was used as the criterion.

Many studies have shown that clinical assessments of intermediate levels of hyperthyroidism and hypothyroidism are extremely difficult; in about 50 per cent of untreated patients the diagnosis may be in doubt (Billewicz, 1969). In treated patients I have found clinical assessment even more difficult.

By this criterion of the free thyroxine index, five out of six patients with primary myxoedema, and seven out of 12 with myxoedema secondary to operation or radioactive iodine therapy, were not receiving enough thyroxine, while one patient (with Hashimoto's disease) was being overdosed.

Follow-up within three months of adjustment of dosage has shown that most of the abnormal patients have reached the normal range of free thyroxine index.

Comparison between the free thyroxine index assessment and the hypothyroid diagnostic index devised by Billewicz and others points the difficulty of clinical assessment; the index requires practice to apply correctly and consistently.

Assessment of the functional status of patients is again not rewarding. One grossly myxoedematous, housebound woman was found as a result of undertaking this study; she is now euthyroid and able to come to the practice. On the other hand, the overdosed

Hashimoto's patient regrets some reduction in her energy—she needed every ounce to cope with an autistic child! Assessing intermediate changes was extremely difficult.

(5) *Benefits from the thyroid study*

(a) As an educational tool it has taught me much about the detection, surveillance, management, and complications of thyroid disorders, the importance of family history and of alertness for associated disorders such as vitiligo and pernicious anaemia. The value and use of thyroid function tests (and their limitations) are less of a mystery to me. That clinical judgment is often at variance with biochemical tests has been confirmed.

(b) In management, it is important to identify those patients who have had past thyroid disease or treatment (in one case 40 years previously) and maintaining a register of these, and establishing a recall system capable of working up to a year or more later, and identifying non-attenders. The age-sex register cards have been duplicated and filed by month of birth (or month of next review, if this is earlier) for this purpose.

Thus this outcome study revealed inadequate management; the process of management was therefore examined, modified, and re-tested. Audit has served its two main functions: education, at least for me, and planning.

A broader view of the health services strongly suggests that long-term recall of patients under lifetime supervision for such problems as thyroid follow-up, partial gastrectomy, pernicious anaemia, and possibly many others, could best be initiated by computer. Such programmes are already in being (Hedley *et al.*, 1970); it is only remarkable that they are not more widely used.

(6) *The hospitals activity analysis*

It seems reasonable to forecast an extension of this current programme of analysis of individuals and departments within hospitals, to provide general practitioners with their use rates, say, for inpatient beds, urgency of admission, or use of outpatient or investigating facilities. If one general practitioner can manage to look after his patients by sending only one tenth of the average of his colleagues' numbers into hospital, we should look for reasons and ways in which we can all learn.

(7) *Critical events*

There may be critical events which act as indicators of the quality of process or outcome. Metcalfe (1974) has suggested that these should be sought for each age group, covering the spectrum of prevention, case-finding, acute care, chronic maintenance, and terminal care. If audit shows a practice to be functioning well using these indicator conditions, then there is a reasonable assumption that the rest of the practice is of high quality.

For instance, all positive pathological laboratory investigations are recorded and followed up in a special book devised by Strube (1974). Stevens uses a 'physical/behavioural/social' grid with his trainees to analyse urinary infection, diabetes, hypertension, and other diseases, as well as looking at 'boundary' problems such as terminal care, referrals to social services, etc. Some practices investigate all deaths; we have attempted group evaluation of all electrocardiograms performed.

Marinker (1974) has suggested there might also be 'critical pathways' which should be followed, for example, postmenopausal bleeding provokes the response—refer for curettage; but 'critical path' construction is difficult and time-consuming (Dudley, 1974).

Sanazaro (1974) notes that minimal critical standards or 'essential' criteria for auditing are being introduced, defined as: "Those elements of diagnosis and treatments which are essential to the proper care of every patient with a specified condition."

He lists three main categories:

- (1) Items in the history, physical examination, and laboratory and radiological procedures which confirm the diagnosis, influence the choice, or application of treatment, and help establish prognosis,
- (2) Specific treatment which is known to be efficacious,
- (3) Procedures which are known to be contra-indicated.

However, there must always be the opportunity for 'lateral thinkers' to view problems from different angles.

By correlating therapeutic decisions with 'information collected' rather than 'diagnosis made', Howie (1974) has shown that there is agreement about the 60 per cent of patients with respiratory tract infections who require antibiotics, and the 20 per cent who do not. With the remaining 20 per cent it may be possible to identify more discriminating factors, or to evaluate the risks and benefits of alternative patterns of behaviour—how many penicillin allergies are 'worth' one rheumatic fever? "It may be that high prescribers in this field are only requiring the courage of someone else's conviction to lead to a change in behaviour . . ."

(8) *Random controlled trials*

To eliminate bias and observer variability, Bradford Hill introduced the concept of random controlled trials.

"In this way the characteristics of the patients are randomised between the two groups, and it is possible to test the hypothesis that one treatment is better than another and express the results in the form of the probability of the difference found being due to chance or not" (Cochrane, 1971).

The Ilkeston Geriatric Intervention Study (Skinner, 1974) randomly allocates people approaching age 65 to two groups; one will have the full screening programmes and application of all available services intensively, the other will have the customary level of medical and social care on demand. Comparison after several years may ascertain whether active intervention programmes identify treatable disease, prevent deterioration, and promote health.

(9) *The problem-orientated medical record*

Studying the *process* of medical care depends greatly on the quality of the medical record. Weed (1969) suggests that our records should become 'problem-orientated' rather than 'diagnosis-orientated', as this is much more scientifically accurate; as well as specific diagnoses, problems may include symptoms or abnormal signs or laboratory findings, or social disharmony.

He identifies components of the medical record, each with a specific function: problem list, data base, initial plan, and problem-orientated progress notes and flow charts.

Properly done, such records are capable of 'clerical' audit, that is, a non-expert investigation as to whether the record is complete in content and method; and of 'medical' audit, which examines each record for professional competence and logical thought.

McIntyre (1974), teaching medical students as does Weed, summarises the indices of medical quality under six headings: thoroughness, reliability, analytical sense, efficiency, manipulative skills, and attitudes. How would our present records fare against these criteria?

Studies on care through problem-orientated medical records in general practice are being undertaken by Tait (1973), and by Morrell (1974).

Indeed, how else except through the medical record can performance by the physician be monitored? One-way mirrors and videotape recordings have both been used to analyse component parts of consultations and pattern of patient and doctor behaviour, but these are still experimental tools.

(10) *Are there any alternatives?*

Random examination of patients along the lines of the Dental Estimates Board examination of dentists' patients might be possible, but is fraught with difficulties and likely to be totally unacceptable if imposed from outside.

Teachers in the Norwich Vocational Training Scheme have sat in on each other's surgeries, willingly accept trainees from other practices, and have visiting undergraduates—as do many more. This has great possibilities for learning as well as teaching; what is most lacking is the time to be free to visit other practices, as well as an appropriate auditing format.

(11) *Re-examination*

The examination for membership of the Royal College of General Practitioners, though criticised, is available to every general practitioner voluntarily to test his own knowledge, skills, and attitudes. The Relicensure and the Professional Standards Review organisation in the United States of America will be watched with interest.

Finance

The financial aspects of auditing need entirely different techniques applied by experts from other fields, and though cost awareness has, I hope, been implicit throughout this report, I have confined myself to the medical aspects of auditing and evaluation.

Acknowledgements

I am particularly grateful to Messrs Upjohn Ltd., the Education Awards Sub-Committee of the Royal College of General Practitioners, and to my partners for the opportunity to undertake this Travelling Fellowship. For her continued encouragement, help, and forbearance, I owe an enormous debt of gratitude to my wife. My thanks are due to the many doctors and others with whom I have had invaluable discussions and meetings, to Mrs Batchelor, NANIME Librarian, and Miss M. Hammond, Librarian, Royal College of General Practitioners.

REFERENCES

- Billewicz, W. Z., Chapman, R. S., Crooks, J., Day, M. E., Gossage, J., Wayne, E. & Young, J. A. (1969). *Quarterly Journal of Medicine*, **38**, 255–266.
- Buchan, I. C. & Richardson, I. M. (1973). *Time study of consultations in general practice*. Scottish Health Service Studies No. 27. Scottish Home and Health Department.
- Cassel, J. (1974). Personal communication.
- Cochrane, A. L. (1971). *Effectiveness and Efficiency*. The Rock Carling Fellowship Lecture. London: Nuffield Provincial Hospitals Trust.
- Curtis-Jenkins, G. (1974). Personal communication.
- Dudley, H. (1974). *British Medical Journal*, **1**, 275–277.
- Durno, D. (1973). *Update*, **7**, 279–280.
- Eimerl, T. S. & Laidlaw, A. J. (Edit.) (1969). *A Handbook for Research in General Practice*. Edinburgh & London: Livingstone.
- Family Planning Association (1973). *Clinic Handbook*. London: F.P.A.
- Fraser, R. C., Patterson, H. R. & Peacock, E. (1974). *Journal of the Royal College of General Practitioners*, **24**, 304–319.
- Gruer, K. T. (1972). *Journal of the Royal College of General Practitioners*, **22**, 100–107.
- Health Services Research Unit, University of Kent at Canterbury (1973). First Report. Centre for Research in the Social Sciences.
- Hedley, A. J., Scott, Anne M., Weir, R. D. & Crooks, J. (1970). *British Medical Journal*, **1**, 556–558.
- Hodgkin, G. K. (1973). *Journal of the Royal College of General Practitioners*, **23**, 759–767.
- Howie, J. G. R. (1974). *Update*, **9**, 539–546.
- Irvine, D. (1972). *Teaching practices. Reports from General Practice No. 15*. London: *Journal of the Royal College of General Practitioners*.

- McGuinness, B. W. (1970). *Journal of the Royal College of General Practitioners*, **19**, 111-114.
- McIntyre, N. (1974). Personal communication.
- Marinker, M. (1974). Personal communication.
- Marsh, G. N. (1969). *British Medical Journal*, **4**, 42-44.
- Metcalfe, D. H. (1974). Personal communication.
- Morrell, D. C. (1974). In *Annual Report, Social Medicine and Health Services Research Unit*, p. 88. London: St. Thomas's Hospital Medical School.
- Mourin, K. A. (1974). *Auditing in General Practice—cohort studies of thyroid disease*. Paper presented to the General Practitioner Research Club.
- National Building Agency (1967). *Design Guide for Medical Group Practice Centres*. London: College of General Practitioners.
- Patterson, H. R., Fraser, R. C. & Peacock, E. (1974). *Journal of the Royal College of General Practitioners*, **24**, 237-241.
- Royal College of General Practitioners, Birmingham Research Unit and Department of Engineering Production, University of Birmingham (1973). *Journal of the Royal College of General Practitioners*, **23**, 400-403 & 404-412.
- Royal College of General Practitioners, Office of Population Censuses and Surveys, and the Department of Health and Social Security (1974). *Morbidity Statistics from General Practice. Second National Study 1970-71*. London: H.M.S.O.
- Sanazaro, P. J. (1974). *British Medical Journal*, **1**, 271-274.
- Skinner, J. L. (1974). *The Ilkeston Geriatric Intervention Study*. Communication to the General Practitioner Research Club.
- Smith, J. W. & O'Donovan, J. B. (1970). *British Medical Journal*, **4**, 673-677.
- Stevens, J. (1974). *Journal of the Royal College of General Practitioners*, **24**, 5-52.
- Strube, Gillian (1974). Personal communication.
- Tait, I. & Stevens, J. (1973). *Journal of the Royal College of General Practitioners*, **23**, 311-315.
- Weed, L. L. (1969). *Medical records, medical education and patient care*. Cleveland: Press of Case Western Reserve University.
- Williamson, J. W. (1971). *Journal of the American Medical Association*, **218**, 564-569.
- Wilson, J. M. G. & Jungner, G. (1968). *Principles and practice of screening for disease*. Geneva: World Health Organisation.

THE TREATMENT OF VAGINAL CANDIDOSIS

A comparative trial of miconazole, amphotericin and providone iodine

A study of three antifungal preparations: miconazole (Gyno-Daktarin), amphotericin B (Fungilin cream) and povidone iodine (Detadine vaginal gel) in the treatment of vaginal candidosis is presented.

Of the three preparations miconazole was found to be the most effective.

REFERENCE

- Warren, R. M., Welply, G. A. C. & Elstein, M. (1974). *Clinical Trials Journal*, **11**, 148.

IATROGENIC DISEASE IN GENERAL PRACTICE : ITS INCIDENCE AND EFFECTS

A year's survey of iatrogenic disease in general practice showed that one consultation in every 40 was the result of iatrogenic disease. Iatrogenic disease may affect the doctor-patient relationship, often leading the doctor to feel guilty and the patient to become aggressive.

REFERENCE

- Mulroy, R. (1973). *British Medical Journal*, **2**, 407-410.