Routine data: a resource for clinical audit?

Martin McKee

Accurate and timely information is an essential prerequisite for effective clinical audit. Much time and effort go into its collection, often on specially designed microcomputer systems. This activity takes place in parallel to, but usually quite separately from, the collection of the enormous volume of routine administrative data, covering the entire range of NHS activity. This suggests that routine data are seen as of little use in audit, an impression borne out by the inclusion of only three papers based on routine data, out of a total of over 30, in the *BMJ's* "Audit in Practice" series.[1–3]

There are several possible explanations. The data may be inappropriate to the needs of clinicians, they may be of poor quality, or they may simply be inaccessible. Conversely, the data may be useful but audit programmes may not be in place and able to use them. Yudkin and Redman, for example, comment how "striking and unexpected results" from the Oxford Maternity Information System were fed back to clinicians but often were not discussed and did not lead to changes in management.[1]

This paper examines the appropriateness, quality, and accessibility of routinely collected data for clinical audit. Firstly, the theoretical uses and limitations of routinely collected data in evaluating health care interventions, including the ways in which data are recorded and coded, will be considered, followed by the implications for the use of routine data in audit. In this paper, "routinely collected data" is considered to refer to those data whose primary reason for collection is other than audit. Such data thus include data from hospital and community information systems; cancer registries; systems established to manage specific programmes, such as breast or cervical screeing; and radiology, pathology, pharmacy, and accident and emergency systems.

Routinely collected data: cheap, comprehensive, and consistent

Routinely collected data have several strengths. Firstly, and by definition, their collection is part of the routine management of the service so very little additional cost should be involved in obtaining them for audit. Secondly, they are relatively comprehensive. A minimum data set is collected on every patient admitted to an NHS hospital (box)[4] or a private hospital, where this is part of an NHS contract.[5] Other systems provide information on entire sections of the population being offered certain services, such as breast or cervical screening or child health surveillance.

Health Services Research Unit, London School of Hygiene and Tropical Medicine, London WC1E 7HT
Martin McKee, senior lecturer

Körner minimum data set: inpatients

Patient	*Consultant episode*	*Contract*	*Pregnancy and delivery*
NHS number*	Local patient	Contract identifier	GP code (antenatal care)
Name	identifier		1st antenatal assessment
Address	Episode number	*Psychiatric census*	date
Postcode	Start date	Date detained	Parity
DHA of residence	End date	Status of patient	Delivery place type
Sex	Overseas visitor status	Age	Delivery date
Marital status	Specialty	Duration of	Initial intended delivery
Date of birth	Specialty (shared	care/detention	place/type
Registered GP	care)	Legal status	Change from intention:
Ethnic origin*	Consultant code	Mental category	Reason for change
	Diagnosis (up to 6)	Date detention	Length of gestation
Provider spell	Procedure (up to 4):	commenced	Labour onset method
Referring GP	Date	Ward type	Delivery method
Administrative category	Surgeon code	Diagnosis (up to 6)	Status of person
Decided to admit date	Record type		conducting delivery
Start date			Anaesthetic –
Admission method	*Ward stay*		labour/delivery
Duration of elective	Ward code		Anaesthetic – post delivery
wait	Ward classification		Number of babies
Intended management	(up to 4)		1st to 6th baby:
Source of admission	Ward episode start		Sex
Discharge date	date		Birth order
Discharge method	Ward episode end		Live/still birth
Discharge destination	date		Birth weight
Patient classification	Days in ward		Method of resuscitation
Sick/disabled indicator	classification		Date of birth

*From 1994. DHA=district health authority

Thirdly, data sets are the same throughout the NHS, enabling practitioners to compare their work with others elsewhere. Although the Lothian[6] and North West Thames[7] surgical audit programmes made use of non-routine information systems, they illustrate the potential benefits that can be achieved when clinicians in different hospitals collaborate to compare results.

Despite these advantages certain questions remain. To what extent are limited numbers of variables contained in routine data appropriate to answer the questions being asked? What do we know about their quality? How accessible are they? These questions will be answered in turn.

Information needed for audit
Questions that information systems might be asked typically take the form of either process audit: "What percentage of a defined population is receiving an intervention compared with an agreed standard?" or outcome: "What is the change of health status of patients undergoing a specific intervention?" Answers to process questions require data that can define the intervention and the population being studied, such as all patients in a certain age group or with a particular diagnosis. Questions of outcome are more complex and the items needed are a description of the intervention and measures of the condition of the patient before and after the intervention, with the post-intervention measurement being recorded after an appropriate interval. The data items required can be identified from the conceptual model shown in figure 1, which represents the least complex case of a single, clearly defined intervention, such as a surgical operation. The extent to which routine data can provide each of these items, and the issues that must be taken into account when interpreting them, will be considered in turn.

DESCRIBING INTERVENTIONS
Data on interventions are needed for answering questions on both process and outcome of care. Interventions can take many forms, such as counselling, clinical examination, administration of drugs, and surgery. Information systems established to manage specific programmes, such as screening examinations or child health services, will contain details of those interventions that are part of the programme, usually indicating whether or not they were performed, and when. Hospital information systems and cancer registries are faced with a more complex problem due to the range of possible interventions. Surgical procedures are described in most detail and are classified according to the Tabular List of the Classification of Surgical Operations and Procedures published by the Office of Population Censuses and Surveys, currently in its fourth version (OPCS4).[8] The OPCS4 classification was designed to be comprehensive, with sufficient detail for most administrative and epidemiological purposes. In most cases a procedure can be described by a single code, but for a few types of procedures, such as those in orthopaedics, a code is required for the type of procedure and another one for its site. The rapidly increasing complexity of modern surgery and the need, in some cases, to have much more detailed information have prompted several groups to develop their own classifications for local use.[9] OPCS4, for example, provides no measure of the complexity of a procedure so some surgeons have used the British United Provident Association (BUPA) classification[10] as this reflects surgical workload,[11] with some hospitals adding this variable to their core data sets. For most uses OPCS4 is sufficiently specific if codes are allocated precisely. It is regularly updated to take account of new procedures such as endoscopic surgery.

A limited amount of information on some other interventions is also provided by hospital information systems and cancer registries, using the "V" codes in the ninth revision of the International Classification of Diseases (ICD-9).[12] These identify admissions in which treatment is being given for an established condition, such as chemotherapy or radiotherapy for patients with cancer. Other interventions, such as the administration of specific drugs, cannot normally be identified.

DESCRIBING PATIENTS
In theory, it should be possible to know something about the condition of a patient before the intervention from the code given for the principal diagnosis. This is defined in the United Kingdom (UK) as "the main reason why the patient was admitted to hospital".[13] For audit, a major problem is the absence of a clear indication of the severity of the condition since this is an important determinant of outcome.[14] It can, however, be calculated from routine data in various ways. In some situations, for example, it may be possible to bring together cases with a range of principal diagnoses and rank each diagnosis on the basis of severity, such as simple angina, unstable angina, and myocardial infarction.[15] In other situations information on severity may be inferred from the presence or absence of secondary diagnoses, which may be either complications of the principal diagnosis or comorbidities. Patients undergoing cholecystectomy, for example, may be categorised according to
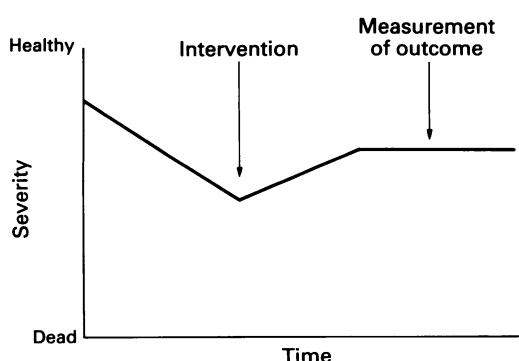

Fig 1 Information required to describe outcome of an intervention

whether or not they are otherwise healthy or have another condition such as diabetes, chronic airways disease, or angina. Roos *et al* described a comorbidity index designed to predict mortality in which these and other diseases are assigned a weighting,[16] and Deyo *et al* described a model incorporating complications and comorbidities.[17] A note of caution is required when examining comorbidities. Jencks *et al* showed how some combordities, such as diabetes and hypertension, may fail to be coded for in cases where there is very severe illness as they are displaced by complications of the principal diagnosis from the limited number of spaces on the record for diagnostic codes.[18] A further problem with data in the UK is the inability to distinguish secondary diagnoses present on admission from complications of treatment. This has been overcome elsewhere by identifying those diagnoses present on admission, as in the New York Statewide Planning and Research Cooperative System.

ICD-9 codes diagnoses as a four digit number, but the tenth revision (ICD-10),[19] to be introduced in 1995, will use an alpha-numeric format. A new feature of ICD-10 is the presence of codes for a series of body-system specific "post-procedural" conditions, such as postoperative renal failure or intestinal obstruction. As indicators of specific adverse outcomes, these may be of use in audit. Some hospitals in the UK use the clinically modified version of ICD-9 (ICD-9CM),[20] which was developed in the United States (US). It contains an additional fifth digit to provide extra information of clinical significance that is not contained in the ICD-9 codes. Some clinically important examples include the differentiation of patients with uncomplicated asthma and status asthmaticus, chest pain suspected to be of cardiac or other origin, and presence or absence of obstruction with cholelithiasis. In general, ICD-9CM codes can be converted to ICD-9 codes by simply removing the fifth digit, but there are a few exceptions. Translating from ICD-9 to ICD-9CM is more difficult. The National Case-mix Office has produced software that adds a fifth digit on the basis of probability but it is recognised that this can lead to loss of information. The main importance of the ICD-9CM system in the UK is that it is required for many of the commercially available case mix systems that are produced in the US.

READ CODES
Some hospitals have adopted the Read coding classification[21] in an attempt to increase the clinical meaningfulness of information. As well as codes for diagnoses and procedures, it contains a range of other variables including symptoms and signs, non-operative procedures, drugs, and social information. The NHS Management Executive has commissioned the Centre for Coding and Classification, under the direction of Dr James Read who designed the system, to develop it further, including codes that describe the work of nurses and professions allied to medicine. Read codes are designed to map to ICD-9CM and OPCS4 codes. With its more detailed coverage of signs and symptoms, Read coding should facilitate more detailed severity adjustment than is possible with ICD-9 but this seems to have received little attention so far. Conversely, the greater detail may increase the scope for interobserver variation.

AUTOMATED SEVERITY SYSTEMS
The examples noted above, ranking principal diagnoses or using secondary diagnoses to introduce a measure of severity, have been developed more formally in a range of "iso-outcome" case mix systems. These seek to adjust for risk of death or other adverse outcomes. Of the available systems, Disease Staging has received most attention in the UK.[22] This system identifies one or more underlying disease processes for each patient, each at a differing level of severity.[23] Disease Staging software combines all listed diagnoses to generate one or more disease categories, with the category generated from the first listed diagnosis being designated as the principal disease category. Within each category, a numeric stage, reflecting the degree of severity, is allocated on the basis of the diagnostic codes making up that category. The disease categories and stages can be combined with other routine data items to generate an overall score that has some ability to predict mortality.[24] Other systems have been developed to identify hospitals with higher than expected patient mortality after adjustment for risk, which are subsequently targeted for intensive review.[25] [26]

ROUTINE VERSUS AD HOC DATA
The less detailed information contained in routine data compared with case notes or questionnaires might suggest a lower level of power to predict outcome, but this is not automatically true and depends on the actual sources of data. Alemi *et al* found that Disease Staging, using routine data, was as good at predicting mortality from myocardial infarction as other severity systems based on data abstracted from case notes.[27] Roos *et al* showed that a model using routinely collected Canadian data has higher power to predict mortality than one using interview data,[28] although routine data in Canada are more detailed than in the UK and can link hospital stays and ambulatory case visits. Other researchers have found non-routine data to be better. Green *et al* found that a severity score based on information contained in case notes[29] added significantly to the predictive power of an early model developed by the Health Care Financing Administration to adjust for severity,[30] although the model subsequently has undergone considerable refinements after initial criticism.[31] Hannan *et al* demonstrated that risk adjustment based on detailed clinical information had higher predictive power for mortality than that based on routinely collected data.[32] Finally, though routine data may be able to suggest that a problem exists,

they will rarely indicate what action should be taken[33] and this will often have to be deducted from information in case notes.

MEASURING OUTCOME
The only unambiguous measure of outcome available is whether the patient died in hospital or was discharged alive, although even this is not always recorded acurately.[31] However, the interpretation of hospital mortality rates is extremely complex. An unexpectedly high mortality rate, even after apparently adjusting for severity, should not automatically be assumed to be due to poor quality of care. The effects of random variation and inadequate adjustment for risk must be considered. Park *et al* demonstrated that random variation accounts for the largest component of observed variation in death rates at hospital level.[34] This will clearly be a greater problem for comparisons at the level of individual consultants. Currently available systems to adjust for risk of adverse outcome, whether they use routine data or information from case notes, can explain no more than about 25% of observed mortality, and many of those working in this field have concluded that it is not yet possible to make valid inferences from risk adjusted outcomes.[35] But observed differences in mortality cannot be explained away entirely by statistical artefact and subtle differences in severity. Dubois *et al*, using case note review with implicit criteria, suggested that those hospitals with unexpectedly high mortality rates after risk adjustment provided worse care, although it must noted that no difference was detected using explicit criteria.[36]

Death rates in hospital vary with length of stay and must be treated with caution when used as a measure of mortality. Jencks showed how in hospital mortality may seem lower for hospitals with short lengths of patient stay than for hospitals that keep their patients longer. Mortality at thirty days is, however, the same.[18] This work was done in the US where lengths of stay are much shorter than in the UK, so its applicability in the UK is uncertain, but the findings should be considered when interpreting apparent differences in death rates in hospital. Other measures affecting mortality rates are differences in social support in the local community, leading to variation in the proportion of people taken home by relatives to die, and differences in hospital "do not resuscitate" policies.

OTHER OUTCOMES
It is not possible to obtain any information about important aspects of outcome such as disability, discomfort, or dissatisfaction, although it may be possible to imply an adverse outcome from the presence of certain secondary diagnoses. Iezzoni suggested that secondary diagnoses in the ICD-9 range 996-999 (complications of surgical and medical care) might be used when screening for cases requiring further investigation,[37] and Roos *et al* have focused on codes for comorbidities, such as pulmonary embolism,

in patients undergoing elective surgical procedures.[38] This approach is likely to be more successful in the US than the UK because of the incentive to code many more secondary diagnoses under the prospective payment system.[39] More information on the consistency with which secondary diagnoses are coded in the UK is required before using them here. The use of measures other than death also requires that information systems link episodes of care. In the UK, apart from locally developed systems of record linkage, such as that in Oxford region, cancer registries provide the only opportunity to follow up patients. Their potential value in audit has been suggested in several evaluative studies.[40-43] They are greatly underused and recent efforts have made them much more accessible, accurate, and complete.

Data quality and accessibility
Traditionally, one of the main reasons for failure to use routine NHS data has been concern about its quality. Data quality covers three measures: completeness, accuracy, and precision.

COMPLETENESS
The completeness of coding is commonly defined as the percentage of records with at least one diagnostic code recorded. This reflects the priority placed by hospital management on the coding process. Although there have been considerable improvements in recent years, with many hospitals having diagnostic codes recorded for over 98% of episodes within one month of discharge, some have made little progress. The percentage achieved and the time interval are both important, and for information to be of use in audit it must be available to clinicians as soon as possible.[16] Data quality has implications beyond audit. Given the demands of the internal market, there must be serious questions about how it is possible to manage a modern hospital that fails to code 60% of cases. A caveat is required. Some hospitals with apparently complete coding may be achieving this with the automatic insertion of default codes. Most commonly this occurs when the system is set up to insert the code 799·9 (no diagnosis) in all records if no other code is present.

ACCURACY
Inaccuracy can arise at any stage in the passage of information from the bedside to the computer. The first step is to make a diagnosis. Assuming that it is correct, the limitations of the ICD classification create the next difficulty. With a few exceptions, such as psychiatric disorders, ICD-9 and ICD-10 contain no explicit definitions of diseases. Thus it is a matter of clinical judgement whether a disease is present or absent.[37] Although additional criteria for diagnosing some diseases, such as diabetes mellitus and AIDS, have been developed by the World Health Organisation or by other bodies, they are used inconsistently. Uncertainty about

whether the patient is normal or has a disease and, if so, what the disease is, arises because clinicians, pathologists, radiologists, and others may disagree about the point at which a particular set of findings should be considered to represent a disease, especially where the findings of different methods of assessment are poorly correlated. Well recognised examples include osteoarthritis of the hip[44] and benign prostatic hypertrophy.[45] Inaccuracy may also arise when codes are allocated if the case notes do not contain sufficient information to identify all relevant diagnoses. This stage may also be subject to differences in interpretation by individual coders. There is evidence of systematic variation in the use of certain diagnoses such as "coronary atherosclerosis" and "angina,"[46] and "chronic bronchitis" and "emphysema."[47] This becomes important when specifying search terms to extract cases. Further errors may occur as codes are transcribed and entered into computers. Many coding departments have quality control mechanisms, but experience suggests that the best results are achieved where there is a close working relationship between coders and clinicians, preferably with consultants "signing off" each completed record.

PRECISION

Variation also arises in the precision with which diagnostic codes are allocated. Some ICD-9 codes describe symptoms and signs, such as malaise and fatigue (780·7) or hepatomegaly (789·1), rather than diagnoses. Although their use may be entirely appropriate (as it would be unjustifiable to investigate patients simply to be able to label them) more often they seem to reflect poor record keeping, such as the frequent use of the code for retention of urine (788·2) as a principal diagnosis in patients undergoing prostatectomy. Imprecision also results from failure to make full use of the fourth digit in ICD-9 with resulting overuse of codes where the fourth digit is "·9", indicating a disorder "unspecified" or "not otherwise specified." This is a greater problem in some areas than in others. In a recent study using a large UK database, for example, almost all patients undergoing peripheral vascular surgery had a principal diagnosis of "peripheral vascular disease, unspecified" despite the availability of much more precise codes.[48]

SCALE OF PROBLEM

How much of a problem is data quality? There have been relatively few published studies addressing all of the tissues described above, and the intrinsic limitations of ICD indicate the methodological difficulties in producing specific figures. Furthermore, genuine clinical uncertainty about whether a particular condition is present or not is greater with chronic medical conditions than with surgical conditions, yet most studies of accuracy have focused on surgical conditions. Using various criteria, some authors have described error rates (variously defined) of between 20% and 40%.[49] [50] Others have found much lower rates and argue that the poor general impression about the accuracy of hospital data may be affected by submission bias or publishing bias, authors being more likely to report their experience when the accuracy of data is low.[51] There is growing evidence from several regions that completeness of data has improved considerably over the past two years. Concern about the quality of routinely collected data has lead to arguments in favour of the establishment of separate information systems, often based on microcomputers, as their ownership by clinicians will lead to improved data quality. It is not clear that this necessarily follows since several groups using these systems have noted levels of completeness similar to those seen with routine systems.[52] [53] What is in no doubt is that data quality is unlikely to improve if the data are never used.

ACCESSIBILITY

Routine data should be readily accessible from hospital information departments. Previously, most requests for such information were met by providing long and often impenetrable paper printouts. Much wider availability of personal computers has meant that it is now often more appropriate to receive data on disk (box). This facility should have transformed access to routine data. Unfortunately this seems not to be the case, although there is wide variation among and within regions. It is amazing that in 1993 some information systems have such difficulty producing files in the industry standard format.

How can routinely collected data be used in audit?

Given their strengths and weaknesses, what can routine data contribute to audit? This question can be thought of in terms of process and outcome and in terms of the points on the audit cycle where routine data may be expected to help: topic selection and description of activity (fig 2).[55] It must be emphasised again that data collection will be of no value if it is not part of a well organised audit process.[56]

PROCESS AUDIT

Routine data are particularly well suited to answer questions of process. When there is clear agreement about the level of an intervention that is desirable in a designated population routine data can often be used to monitor the extent to which this is being achieved. In some cases this has evolved, and audit is now part of the routine management of the service. Examples include the use of child health surveillance systems to monitor immunisation rates and cervical and breast screening systems to monitor uptake rates. For example, Wilson has described two audit cycles in which general practitioners reviewed the percentage of women on their lists who had had a cervical smear in the preceding five years or who had never had one.[57] Although this used a variety of manual and computer

How to access data

Data should be asked for as a flat ASCII file with one record per line. To reduce the size of the data file only those variables that are required should be asked for, and this will require considering in advance the questions to be asked. In addition, for transfer, data files can be reduced in size considerably by using compression software. When requesting data, it is clearly important to specify the time period being examined. This is influenced by the completion of data entry. Most districts aim at having all records entered within one month of discharge. Information is recorded on the basis of consultant episodes. Thus one record will represent each spell spent under the care of a consultant. This should not be confused with a "case" or a patient being treated, particularly with the apparent increase in the tendency to record consultant episodes when patients are transferred for a second opinion or investigation. This is further complicated by multiple admissions in which either a series of investigations (which would have been undertaken during a single admission in the past) or the more widespread use of treatment requiring multiple admissions, such as chemotherapy, have increased.[54]

The scope for coding multiple diagnoses and procedures creates certain problems when linking them with patients. Although there are rules to guide decisions about ordering diagnoses, these may be difficult to implement from the information available in the case notes. Consequently, when searching for a particular diagnosis or procedure it is important to look at all available diagnostic codes.

Certain tools and skills are necessary for meaningful analysis of data. Some microcomputer audit packages can import and analyse routine data, although they tend to have limited flexibility. Generic databases and statistical packages, such as SPSS-PC and DBase, are more powerful and flexible, but require special training. Obviously, a knowledge of health service research methodology is necessary for interpreting results.

based general practice systems, it is now part of the routine cell and recall system.

The essential prerequisites for routine data to be used in process audit are the ability to identify both a specific intervention and a defined population, such as all women giving birth or all patients undergoing non-urgent surgery. Further examples are provided by Yudkin and Redman, in a review of the use of the Oxford Maternity Information System.[2]
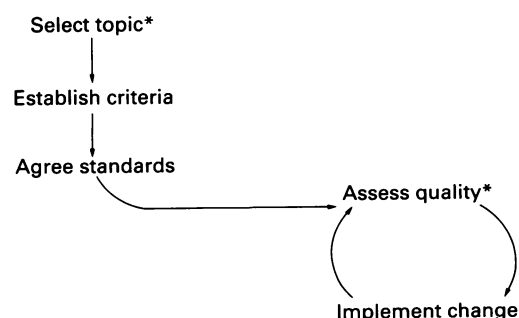

*Stages in which data may be useful

Fig 2 *Potential contribution of data to audit*

They reported how some audits required only data collected on their system, although they also collect non-standard information on interventions and their indications not available elsewhere. More commonly, routine data served to inform topic selection, with the audit process using additional information from other sources. The value of the additional items on the Oxford system demonstrates the potential for using the user definable research variables on some patient administration systems.

The success of these examples reflects the relative simplicity of their design and the level of agreement about the standard to be achieved. This model can be extended to other interventions where there is currently lack of agreement about the desired standard. Examples include the monitoring of the rate of preoperative chest x ray exclusions,[58] laboratory investigations,[59] or caesarean sections.[60] The first two examples require access to radiology and laboratory systems. More complex examples have been described using cancer registry data to compare the types of treatment received by patients of different consultants.[35]

OUTCOME AUDIT

For questions of outcome, the constraints of routine data described above limit their value. Routine data are of most use when specific measures of outcome are recorded, such as child health systems or in the national breast screening programme. Colver described the use of feedback to primary health care teams of routinely collected data on child health surveillance.[2] Children were included in the study population on the basis of the eventual discovery of deafness or physical handicap. The presence of either condition was the measure of their condition before the intervention and the start of treatment was the outcome measure. The interventions were the screening examinations for these conditions and the interval between intervention and measurement of outcome, in this case the follow up period, was four years. The system was able to identify reductions in the ages at which deafness was detected and physiotherapy started for children with physical handicap. Reductions were not seen in adjoining districts where feedback did not take place.

Although mortality in hospital must be treated with caution for the reasons discussed, this does not invalidate it as a measure of outcome.[9] It may be of use, after adjustment for severity, to identify cases for more detailed study. This approach has been described by Hannan *et al,* who examined a variety of criteria for selecting patients for case note review derived from routinely collected data.[61] They included the presence of specific secondary diagnoses occurring in surgical patients, such as infections, wound infections, renal failure, and cardiopulmonary arrest, and deaths among patients with primary surgical procedures or in diagnosis related groups (DRGs),[62] where the expected mortality is

under 0·5%. For many of these criteria the selected case notes were significantly more likely to reveal evidence of poor quality care than unselected notes. Case notes of patients dying in low mortality DRGs were five times more likely to contain evidence of care that departed from professionally recognised standards or care which contributed to death.[61]

CASE FINDING

Databases may also be used to identify cases for more detailed study. This may be on the basis of diagnosis or intervention. Milne, for example, used Hospital Activity Analysis (HAA) data, in conjunction with specific registers, to identify patients with sickle cell disease,[63] although he did not comment on the value of HAA data compared with the registers. A possible extension of this concept would be to programme a computer to send patients, identified as cases for more detailed study, a questionnaire on the outcome of their treatment at a fixed period after discharge.

Conclusions

To return to the questions posed at the beginning of this article, what can be concluded about the appropriateness, quality, and accessibility of routine data for audit? Although routine data have many limitations, many can be overcome, at least in part, by using techniques such as risk adjustment, as long as caution is exercised when interpreting findings. Routine data will never provide a definitive answer to a question about quality of care, but they have an important, and largely unrecognised, role as a screening tool to identify those areas where there seems to be cause for concern, needing further study.[64-66] This might arise where death rates from some condition or after some intervention are unexpectedly high or patients are treated differently from elsewhere.

The ability to compare practice across districts is perhaps the greatest strength of routine data, but it requires complex organisational processes to create an effective audit structure. The caveats about interpretation are important. Temple applauded the conservatism of those most active in promoting the use of routine databases, but he expressed the caution that others will have neither the interpretive skills nor the recognition of the limitations of routine data.[67] The dangers of drawing unjustified conclusions from inadequately tested models have been seen in the response to the publication of the early results of the comparisons of mortality rates in different hospitals compiled by the US Health Care Financing Administration[68] and a similar British report.[69]

The quality of NHS data has improved considerably in the past two years and is continuing to improve. In the hospital sector most hospitals are achieving diagnostic coding rates of over 95% at one month, but a few are performing much less well. Published information on coding quality is lacking, but results from quality control activities suggest that it is

generally high where it is taken seriously by management. There is, however, much that can be done by health care professionals to improve it further, by working closely with coders, agreeing criteria for diagnosis of common conditions, and recording complications and comorbidities. The next few years will see further improvements in quality and coverage. Detailed information will also be available for outpatients, and the use of the NHS number as a unique patient identifier from 1994 will facilitate linkage of episodes of care in different settings. The quality and accessibility of routine data could easily be improved further. However, for this to happen clinicians, managers, and everyone involved in the health service must not only need to use the data but continue to insist on high quality data. The feedback loop must be closed.

1 Yudkin PL, Redman CW. Obstetric audit using routinely collected computerised data. *BMJ* 1990;**301**:1371–3.
2 Colver AF. Health surveillance of preschool children: four years' experience. *BMJ* 1990;**300**:1246–8.
3 Chambers M, Clarke A. Measuring readmission rates. *BMJ* 1990;**301**:1134–6.
4 Steering Group on Health Service Information. *First report to the Secretary of State.* London: HMSO, 1984. (Körner report.)
5 McKee M, Clarke A, Tennison B. Meeting local needs. *BMJ* (in press).
6 Gruer R, Gordon DS, Gunn AA. Audit of surgical audit. *Lancet* 1986;i:23–6.
7 Black NA. A regional computerised surgical audit project. *Quality Assurance in Health Care* 1990;**2**:263–70.
8 Offices of Population Censuses and Surveys. *Classification of surgical operations and procedures. Fourth revision.* London: OPCS, 1987.
9 Nixon SJ. Defining essential hospital data. *BMJ* 1990;**300**:380–1.
10 British United Provident Association. *Schedule of procedures 1989.* London: BUPA, 1989.
11 Jones SM, Collins CD. Caseload of workload? Scoring complexity of operative procedures as a means of analysing workload. *BMJ* 1990;**301**:324–5.
12 World Health Organisation. *International classification of diseases 1975. Revision.* Geneva: WHO, 1977.
13 NHS Management Executive. *Clinical coding.* (Computer based training package.) London: NHSME, 1990.
14 Dubois R, Rogers WH, Moxley JH, Draper D, Brook RH. Hospital inpatient mortality. Is it a predictor of quality? *N Engl J Med* 1987;**317**:1674–80.
15 Ayanian JZ, Epstein AM. Differences in the use of procedures between women and men hospitalized for coronary artery disease. *N Engl J Med* 1991;**325**:221–5.
16 Roos LL, Sharp SM, Cohen MM, Wajda A. Risk adjustment in claims-based research: the search for efficient approaches. *J Clin Epidemiol* 1989;**42**:1193–1206.
17 Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9CM administrative databases. *J Clin Epidemiol* 1992;**45**:613–9.
18 Jencks SF, Williams DK, Kay T. Assessing hospital associated deaths from discharge data: the role of length of stay and comorbidities. *JAMA* 1988;**260**:2240–6.
19 World Health Organisation. *International statistical classification of diseases and related health problems. 10th revision.* Volume 1. Geneva: WHO, 1992.
20 *International classification of diseases.* 9th Revision. Clinical modification. Ann Arbor, Michigan: Commission on Professional and Hospital Activities, 1980.
21 Buckland R. The language of health. *BMJ* 1993;**306**:287–8.
22 McKee M, Petticrew M. Disease staging: A case mix system for purchasers? *J Public Health Med* 1993;**15**:25–36.
23 Gonnella JS, Hornbrook MC, Louis DZ. Staging of disease: a case mix measurement. *JAMA* 1984;**251**:637–44.
24 Christoffersson J, Conklin JE, Gonella G. The impact of severity of illness on hospital utilization and outcomes. *Infoplus* 1991;**1**:1–8.
25 Lohr KN. Outcome measurement: concept and questions. *Inquiry* 1988;**25**:37–50.
26 DesHarnais SI, Chesney JD, Wroblewski RT, Flemming ST, McMahom LF. The risk-adjusted mortality index: a new measure of hospital performance. *Med Care* 1988;**26**:1129–48.

27 Alemi F, Rice J, Hankins R. Predicting in-hospital survival of myocardial infarction: a comparative study of severity issues. *Med Care* 1990;**28**:762–75.
28 Roos NP, Roos LL, Mossey J, Havens B. Using administrative data to predict important health outcomes. *Med Care* 1988;**26**:221–39.
29 Green J, Wintfield N, Sharkey P, Passman LJ. The importance of severity of illness in assessing hospital mortality. *JAMA* 1990;**263**:241–6.
30 Horn SD, Sharkey PD, Bertram D. Measuring severity of illness: homogeneous casemix groups. *Med Care* 1983;**21**:14–31.
31 Blumberg MS. Comments on the HCFA hospital death rate statistical outliers. *Health Serv Res* 1987;**21**:715–40.
32 Hannan EL, Kilburn H, O'Donnell JF, Lukacik G, Shields EP. Adult open heart surgery in New York State. *JAMA* 1990;**264**:2768–74.
33 Brook RH, Lohr KN. Monitoring quality of care in the Medicare programme: two proposed systems. *JAMA* 1987;**258**:3138–41.
34 Park RE, Brook RH, Kosecoff J, Keesey J, Rubenstein L, Keeler E, et al. Explaining variations in hospital death rates: randomness, severity of illness, quality of care. *JAMA* 1990;**264**:484–90.
35 Brook RH, Iezzoni LI, Jencks SF, Kuaus WA, Krakaver H, Lohr KN, et al. Symposium: case mix measurement and assessing quality of hospital care. *Health Care Financing Review* 1987(suppl):39–48.
36 Dubois RW, Rogers WH, Moxley JH, Draper D, Brook RH. Hospital inpatient mortality: is it a predictor of quality? *N Engl J Med* 1987;**317**:1674–80.
37 Iezzoni LI. Using administrative diagnostic data to assess the quality of hospital care: pitfalls and potential of ICD-9CM. *Int J Technol Assess Health Care* 1990;**6**:272–81.
38 Roos LL, Cageorge SM, Austen E, Lohr KN. Using computers to identify complications after surgery. *Am J Public Health* 1985;**75**:1288–95.
39 Hsia DC, Krushat WM, Fagan AB, Terbbluth JA, Kusserprow RP. Accuracy of diagnostic coding for Medicare patients under the prospective payment system. *N Engl J Med* 1988;**318**:352–5.
40 Cox NH. Basal cell carcinoma in young adults. *Br J Dermatol* 1992;**127**:26–9.
41 Finn CB, Luesley DM, Buxton EJ, Blackledge GR, Kelly K, Dunn JA, et al. Is stage I epithelial ovarian cancer overtreated both surgically and systematically? Results of a five year cancer registry review. *Br J Obstet Gynaecol* 1992;**99**:54–8.
42 Thornhill JA, Walsh A, Conroy RM, Fennelly JJ, Kelly GD, Fitzpatrick JM. Physician dependent prognostic variables in the management of testicular cancer. *Br J Urol* 1988;**61**:244–9.
43 Bull A, Mountney L, Sanderson H. Stage distribution of breast cancer: a basis for the evaluation of breast screening programmes. *Br J Radiol* 1991;**64**:516–9.
44 Office of Health Economics. *Total hip replacement and the NHS.* London: OHE, 1982, 19.
45 Garraway WN, Lee RJ. High prevalence of benign and prostate hypertrophy in the community. *Lancet* 1991;**338**:469–71.
46 Iezzoni LI, Burnside S, Sickles L, Moskowitz MA, Sawitz E, Levine PA. Coding of acute myocardial infarction: clinical and policy implications. *Ann Intern Med* 1988;**109**:745–51.
47 Ginsberg PT, Hammons GT. Competition and the quality of care: The importance of information. *Inquiry* 1988;**25**:108–15.
48 McKee M, Petticrew M. *An evaluation of Disease Staging in the United Kingdom. Report to the National Case-mix Office,* London: Department of Health, 1991.
49 Mukherjee AK, Lack I, Langley FA, Ashcroft C. The completeness and accuracy of health authority and cancer registry records according to a study of ovarian neoplasms. *Public Health* 1991;**105**:69–78.
50 Smith SH, Kershaw C, Thomas IH, Botha JL. PIs and DRGs: coding inaccuracies and their consequences for resource management. *J Public Health Med* 1991;**13**:40–1.
51 Sellar C, Goldacre MJ, Hawton K. Reliability of routine hospital data on poisoning as measures of deliberate self-poisoning in adolescents. *J Epidemiol Community Health* 1990;**44**:313–5.
52 Barrie JL, Marsh DR. Quality of data in the Manchester orthopaedic database. *BMJ* 1992;**304**:159–62.
53 Lyons C, Gumpert R. Medical audit data: counting is not enough. *BMJ* 1990;**300**:563–6.
54 Clarke A, McKee M. Consultant episodes: an unhelpful measure. *BMJ* 1992;**305**:1307–8.
55 Crombie IK, Davies HTO. Computers in audit: servants or sirens? *BMJ* 1991;**303**:403–4.
56 Guthrie KA, Kelly M, Lilford RJ. Information systems and audit in antenatal care. *Baillieres Clin Obstet Gynaecol* 1990;**4**:207–27.
57 Wilson A. Cervical cytology in the Vale of Trent faculty of the Royal College of General Practitioners, 1985–8. *BMJ* 1990;**300**:376–8.
58 Fowkes FG, Davies ER, Evans KT, Green G, Hartley G, Hugh AE, et al. Multicentre trial of four strategies to reduce use of a radiological test. *Lancet* 1986;**i**:367–70.
59 Fowkes FG, Hall R, Jones JH, Scanlon MF, Elder GH, Hobbs DR, et al. Trial of strategy for reducing the use of laboratory tests. *BMJ* 1986;**292**:833–5.
60 Myers SA, Gleicher N. A successful program to lower cesarean section rates. *N Engl J Med* 1987;**319**:1511–6.
61 Hannan EL, Bernard HR, O'Donnell JF, Kilburn H. A methodology for targeting hospital cases for quality of care record reviews. *Am J Public Health* 1989;**79**:430–6.
62 Jenkins L, McKee M, Sanderson H. *DRGs: a guide to grouping and interpretation.* London: CASPE Publications, 1990.
63 Milne RIG. Assessment of care of children with sickle cell disease: implications for neonatal screening programmes. *BMJ* 1990;**300**:371–4.
64 Lohr KN. Use of insurance claims data in measuring quality of care. *Int J Technol Assess Health Care* 1990;**6**:263–71.
65 DesHarnais S. Current use of large data sets to assess the quality of providers: construction of risk adjusted indexes of hospital performance. *Int J Technol Assess Health Care* 1990;**6**:229–38.
66 Jencks SF, Daley J, Draper D, Thomas N, Lenhart S, Walker J. Interpreting hospital mortality data. The role of clinical risk adjustment. *JAMA* 1988;**260**:3611–6.
67 Temple R. Problems in the use of large data sets to assess effectiveness. *Int J Technol Assess Health Care* 1990;**6**:211–9.
68 Berwick DM, Wald DL. Hospital leaders' opinions of the HCFA mortality data. *JAMA* 1990;**263**:247–9.
69 Kind P. *Hospital deaths – the missing link.* York: Centre for Health Economics, 1988. (Discussion paper 44.)