National Centre for Clinical Audit

Information for better healthcare

Good Practice in Clinical Audit

— a summary of selected literature

to support criteria for clinical audit

Good Practice in Clinical Audit

— a summary of selected literature

to support criteria for clinical audit

Nancy Dixon

National Centre for Clinical Audit

Information for better healthcare

For additional copies of this review, or for information about clinical audit, contact the National Centre for Clinical Audit.

The National Centre for Clinical Audit
BMA House, Tavistock Square
London WC1H 9JP
Tel: 0171 383 6451
Fax: 0171 383 6373
Email: NCCA@kcl.ac.uk

ISBN: 0 9528511 05

CONTENTS

ACKNOWLEDGEMENTS

Literature search support for this review was provided by Jane Smith and Alison Salmon of the National Centre for Clinical Audit and Christine Allam of Healthcare Quality Quest.

The author gratefully acknowledges the valuable contributions of the following reviewers: Gifford Batstone, King's Fund; John Cape, British Psychological Society; Fleur Fisher, British Medical Association; Elizabeth Fradd, NHS Executive West Midlands; J A Muir Gray, NHS Executive Anglia and Oxford; Debra Humphris, Health Care Evaluation Unit, St George's Hospital, London; Gabby Fennessy, Royal College of Midwives and Royal College of Nursing Audit Information Service; Alison Kitson, Royal College of Nursing; Ann G Lawthers, Centre for Quality of Care Research and Education, Harvard University; and Chris Squire, Royal College of Radiologists.

PURPOSE

The principle of doing clinical audit is widely accepted within the National Health Service. However, the processes of doing clinical audit successfully may be less widely understood.

The Department of Health has defined clinical audit as follows:

> "Clinical audit is widely recognised as the systematic critical analysis of the quality of clinical care, including the procedures used for diagnosis, treatment and care, the associated use of resources and the resulting outcome and quality of life for the patient."[1]

This broad concept of clinical audit has encouraged a wide range of activities to be undertaken in the NHS, such as:

- Conduct of observational studies to identify current patient care practices.
- Provision of feedback on present performance to practitioners, in educational contexts.
- Assessment of patient or customer satisfaction with present services.
- Establishment of standards for more clinically effective patient care.
- Comparisons of performance among individual practitioners or organisations.

These activities, and the variety of objectives which they represent, have been extremely valuable for the NHS because they have engaged very large numbers of practitioners in systematically measuring the care and services provided for members of the public. Nonetheless, recent evaluations of audit[2-12] and editorials[13-14] about the need to "audit audit" suggest that it is an appropriate time to bring clarification to the overall intent of clinical audit and to begin to refine the processes used to carry out clinical audit.

A major purpose of the National Centre for Clinical Audit (NCCA) is to improve the quality of clinical audit activities being carried out within the NHS. To that end, the Centre is involved in examining the goals and processes or methods of clinical audit. The Centre has emphasised that clinical audit is a tool to facilitate improvement of the quality of care or service provided to patients and other users. Thus, the Centre is

extending the focus of clinical audit beyond measurement of current practice: The clinical audit process itself must include strategies for action which result in the achievement of major improvements in the appropriateness and effectiveness of clinical practice.

To contribute to improving the quality of clinical audit activities, the Centre has established criteria for good practice in clinical audit. The criteria are based on a synthesis of expert opinion about essential attributes of the audit process, practitioners' views about how audit works successfully in practice, and a review of literature pertinent to the audit process. This report provides a summary of the literature reviewed for this purpose.

This summary is not intended to present the findings of a systematic review of research evidence pertinent to clinical audit. Rather, it focuses only on selected literature of relevance to criteria for good practice in clinical audit. Also, the review is intended to be descriptive rather than analytic of the audit literature.

It is hoped that the review will serve as a reference for individuals who hold responsibility and accountability for the implementation of audit in their own organisations. Clinical audit convenors, leads, facilitators, or committee chairs and members are sometimes asked by their colleagues about the background of audit and the evidence supporting clinical audit. This review identifies and describes some of the literature, including selected research findings, pertaining to how to do appropriate and effective clinical audit.

The review also is intended to stimulate research on the clinical audit process itself so that, in the climate of promoting evidence-based practice, the NCCA and others interested in clinical audit can "continue the search for the best way to perform it."[13]

The findings in this review are organised in accordance with major stages of the audit process. Each stage is described briefly, pertinent references are summarised, and the potential implications for good audit practice are noted. In addition, the review includes a brief description of origins of criteria for audit and observations about the healthcare system context in which audit is being implemented.

Throughout this review, the meanings of key terms used are consistent with those provided in the Glossary.

SELECTED FINDINGS ON CRITERIA FOR CLINICAL AUDIT

The concept of audit of clinical practice has been established through the work of individual pioneers over the last century. The key principles which tend to guide current audit practice in healthcare were first defined in the 1950s through the 1970s.

Since the widespread introduction in the NHS of the audit of medical, nursing, and therapy practices, and the subsequent introduction of clinical audit, principles of audit have been reviewed and examined for their relevance to the NHS. Several authors have supported previously stated principles and some have offered tools for the evaluation of individual audits based upon agreed principles.

Research on the audit process has tended to concentrate on two stages: The development of criteria or standards and mechanisms for achieving changes in clinical behaviour. Randomised controlled trials (RCTs) on audit tend to focus only on aspects of care provided by doctors. The majority of the published RCTs on audit were carried out in the US or Canada.

Those who have analysed the clinical audit process for application to the NHS have suggested criteria for good practice in clinical audit. In addition, the findings of randomised controlled trials and other studies suggest or confirm criteria for audit. Finally, some criteria for good practice in clinical audit are self-evident, such as having objectives for audit which can be subject to measurement and using methods for defining measures of quality and for data collection which are likely to provide valid and reliable data.

Findings of selected literature reported in this review suggest the following key points about carrying out the process of audit:

- It appears to be desirable that those whose work will be covered by an audit participate in selecting the subject and setting objectives for the audit, using a systematic method which focuses on important subjects for audit as perceived by the participants.

- Measures of practice can be used effectively in audit to contribute to improvements in practice when they are developed by or otherwise made acceptable to the practitioners whose performance is to be reviewed.

- Implicit measures of quality of care appear to be less reliable in audit than explicit measures.

- Measures of processes and outcomes of care or process measures alone are preferable to outcome measures alone.

- The patient medical record may be the most readily available and least costly data source for audit in many organisations; however, data from other sources may be more complete.

- A two-phased strategy of (1) using explicit criteria to measure quality of care across several cases and (2) using structured implicit criteria to review the individual cases which do not meet explicit criteria can enable practitioners to focus on potential reasons why there are cases which do not meet either explicit or implicit criteria.

- Audit leaders can motivate colleagues to improve by focusing on what can be done about obvious shortcomings in care provided to patients.

- Changes in practice may be achieved more effectively if a strategy involving several different types of action is used.

APPROACH TO THE REVIEW

Methods

The following computer bibliographic databases were searched for the years 1975-1996: Allied and Alternative Medicine, DHSS - Data, Health Planning and Administration, Medline, and Nursing and Allied Health Literature. The keywords used were: clinical audit, medical audit, nursing audit, management audit, quality assurance health care, quality of health care, quality of patient care, methods, standards, and organisation and administration. The following words were used in truncated form: assess, appraise, criteria, evaluate, measure, observe, standard.

In addition, author searches of the databases were undertaken using names of known researchers in this field. Manual searches were carried out of references lists provided in research reports or reviews. Additional references were obtained via manual searches of journals and publications not included in the computer bibliographic databases and via personal contacts.

The following criteria were established for selection of references for inclusion in this review:

- Description of key stages or functions in the audit process or characteristics or criteria referring to good practice in clinical audit.

- Research on the audit process itself, ie, research on one or more specific stages or steps which are believed to be a critical part of how to do audit, such as establishing or adopting criteria. Studies included were not limited to randomised controlled trials.

- Interpretation of the importance or significance of a key stage in audit or of criteria referring to good practice in audit.

- Any of the above if they referred to medical care evaluation studies, medical or nursing quality assurance, or peer review methods if the processes described in the reference were identical or highly similar to the clinical audit process.

The following types of material were excluded from this review:

- Published and unpublished reports describing how an individual audit has worked to support improvements in the quality of care provided to patients, unless such reports were illustrative of key aspects of the audit process itself.

- Reports of quality assurance or quality improvement activities if the process referred to was not the same as or similar to clinical audit.

- Research studies on the development or implementation of guidelines unless the method for developing measurable criteria based on the guidelines was defined or the research also included audit.

- Material describing the development of audit or other quality improvement activities.

Limitations

Research studies on clinical audit do not appear to have tested thoroughly all the principles of audit which have developed over the years. Individual studies have investigated aspects of the process of audit, and a selection of such studies and their main findings appear in this review.

There are few published large-scale randomised controlled trials which examine the entire audit process. Many studies on audit use a before-after design in which measurement of actual practice by a group is carried out followed by an intervention with practitioners in the group which is aimed at modifying practice followed by repeat measurement of actual practice. Practitioners in the group may be randomly assigned to subjects being audited and/or to the types of interventions being investigated. Before-after research designs are practical to implement in clinical work settings; however, there is considerable potential for bias.

Research about audit must be interpreted with care for the following additional reasons:

- A tradition of audit exists in countries in which audit is now done routinely. However, the widespread implementation of formal medical or clinical audit in all healthcare organisations has been enforced by the governments in those countries. Thus, audit should be viewed as a mandated innovation in these countries' health-care systems.

  The years which characterise the widespread implementation of audit are 1972-78 in the US and slightly later in Canada, 1979-84 in Australia, and 1989-94 in the UK. The nature of the cultural changes involved in implementing audit means that methods in current use in one country may not be acceptable in another, but also that research carried out several years ago in one country may have current relevance in another.

- Studies on audit often are limited to a small number of organisations as research sites and/or a small number of practitioners who may have volunteered to participate in the research. In the absence of large-scale studies, the findings of some of the research must be seen as limited to the specific types of settings or practitioners in which the studies were carried out.

- Most research on audit applies to single professions, eg, doctors or nurses. There are many published examples of audit involving different professions and even patient or user representatives. However, research studies on audit do not appear to have examined the process when audit involves different professional groups or patients or their representatives.

- Research studies on audit do not uniformly report findings using confidence intervals.

THE NATURE OF THE LITERATURE

Iterations of the Principles and Criteria

A number of publications on audit have served the purpose of examining, interpreting, and disseminating the principles of audit or criteria for good practice in audit to doctors, nurses, and other healthcare professionals.[15-27] In general, the principles and criteria promulgated can be summarised as follows:

- Why audit is done, eg, the basis for selecting a subject or topic, the attributes of quality involved, the rationale or motivation for an audit, specific aims and/or objectives for an audit, and who is concerned or involved in the audit.

- How audit is done, eg, development and dissemination of criteria or standards and precise methods for selecting cases for the audit and for collecting and analysing valid and reliable data.

- What happens with the audit findings, eg, comparison of actual practice with criteria or standards, feedback of the comparison to those involved, and evaluation of practice represented by the comparison.

- How improvements in practice are achieved and maintained, eg, action to implement improvements which benefit patients and remeasurement of practice at a later time to confirm the efficacy of the improvements made.

Russell and Wilson[28] reviewed the premises of explanatory and pragmatic research and suggested that clinical audit should be recognised as "the third clinical science." The scientific basis for audit proposed by Russell and Wilson consisted of a cycle made up of the following steps: choose a general topic for audit and a specific hypothesis to be tested; develop a standard and disseminate and implement it; design unbiased and precise methods for sampling patients; collect valid and reliable data on performance; compare these performance data with the standard by careful statistical analysis; feed a clear summary of this comparison back to participants; and ensure that this process generates beneficial change. The steps related to methods for sampling, data collection, and data analysis and interpretation are consistent with established methods of ensuring scientific rigour. Thus, Russell and Wilson suggest that the remaining steps are those for which research on the audit process may need to be focused.

Tools for Evaluation of Audit

Bhopal and Thomson[29] and Lough *et al*[30] reported on experiences in which doctors were involved in making judgements about attributes or criteria identifying good audit. Bhopal and Thomson asked 25 clinicians who were attending a workshop to list the main features of a publication on an audit project. The features named were related to why the topic was selected, aims, standards, measurements, benefits and outcomes, and implementation of the audit and of change.

Lough *et al* developed a list of 14 criteria to identify a good audit project carried out by doctors at registrar level. They had 135 general practitioner trainers in the west of Scotland evaluate the list with each criterion marked as essential or desirable. The elements of audit which were considered essential or desirable by more than half of the GP trainees included the following: reason for choice of the project; potential for change; preparation and planning of the project; time scale to complete the project; negotiation with relevant team members; staff involvement; relevance of the criteria chosen; standard setting; data presented relevant to the criteria; interpretation of presented data; system for change described; further change proposed where appropriate; and second collection of data compared. Money required, the remaining element, was not thought important by at least half of the trainers.

Frameworks for evaluating audit have been published,[31-32] and Bhopal and Thomson,[29] Lough *et al*,[30] and others[33-35] have published checklists for the evaluation of individual audits which include criteria for good practice in clinical audit.

Reviews

Two reviews of medical audit have been published in the UK. In 1982, Fowkes set out a cycle of audit consisting of (1) observing practice, (2) comparing practice with a standard, and (3) implementing change.[36] He outlined some considerations for choosing what to audit and reviewed literature of relevance to the stages in the cycle, where available. At the time of publication, Fowkes recognised the newness of widespread implementation of audit and noted that definitive methods had not yet been shown to be repeatedly effective or ineffective in different situations.

In 1994, Robinson reviewed literature published in 1985-92 in order to identify current knowledge of the effectiveness of medical audit programmes as a whole and of specific interventions of medical audit as a means of changing clinical behaviour.[8] He noted the lack of clear and consistent objectives for the audit process and that published reports of the achievements of medical audit tended to be based on surveys of those involved in the process and thus were potentially self-serving. Robinson reported that evaluation of entire programmes of medical audit was unusual and that published studies tended to be concerned with specific interventions or aspects of audit. He concluded that knowledge about effective methods of bringing about specific changes in clinical behaviour is rudimentary, and that impact is highly dependent on local factors.

In 1982, Palmer and Nesson reviewed methods for medical care evaluation, ie, medical audit, in use in the US in ambulatory care settings.[37] In the US, ambulatory care is the term used to describe care provided to patients who are capable of "walking" into a clinic to be examined or treated. Ambulatory care includes care provided by medical, nursing, or therapy specialists or family (general) practitioners in freestanding or hospital outpatient clinics.

The review covered identification of problems for evaluation; use of process versus outcome data; choice of data source, topics, and criteria; indices for case identification; units of care to be studied; and feedback interventions. The authors summarised the considerable experience of audit in ambulatory care in the US but concluded that the research did not provide an empirical answer as to how best to conduct internal quality assurance (audit). They emphasised that insufficient evidence exists to prove one method better than another and also that there is "inherent improbability that a single best method exists for all facilities and all problems in patient care."

A number of reviews of the literature on nursing audit have been published.[38-42] The authors describe various nursing audit approaches which have been implemented, but avoid drawing conclusions about the audit methods which are likely to be most effective in particular settings.

Randomised Controlled Trials

A summary of selected randomised controlled trials pertinent to the audit process appears in Table 1. Studies are included in the table only if the audit activity described is consistent with the principles or criteria previously described, ie, measurement of actual practice based on criteria or standards followed by interventions aimed at improving practice followed by remeasurement to confirm improvements.

The following types of research studies are not included in the table: Those which use a before-after or cross-over research design without a randomised control element in the design; those which use analysis of large-scale databases or of normative patterns of practice, as opposed to accepted standards of practice, as a basis for change interventions; those which examine if dissemination of a guideline itself changed practice if no further intervention was attempted to improve the implementation of the guideline; or those which only used the audit process in order to evaluate the effectiveness of a change intervention.

Trials tended to focus on parts of the audit process as opposed to the entire process. Details of procedures or methods used in the research are not provided uniformly.

Table 1. Selected Randomised Controlled Trials on the Audit Process

Author and Year of Publication	Part of Audit Process Studied	No and Type of Practitioners and Setting	Research Question(s)	Research Design	Intervention	Main Findings	P Value
Putnam and Curry, 1985[43]	Topic selection Criteria development	16 family physicians in 2 Canadian provinces	(1) Will physicians' performance in their office practices improve if they participate in patient care appraisal and specific activities designed to correct identified deficits? (2) Will their performance improve more if they participate in the selection of the medical conditions to be audited? (3) Will their performance also improve more if they participate in the generation of the criteria?	Doctors assigned randomly to experimental and control groups	Experimental group and research team selected 5 conditions in total and experimental group developed criteria for 2 (one selected, one not). One topic was concealed. Control group kept logs of patients with named conditions. Data collection of actual records was carried out in both groups on all conditions. Experimental group only received individual feedback and offer of assistance. Data collection was repeated.	(1) Improvement in performance was greater in experimental group but not significant (2) No benefit of involvement in selecting conditions for audit (3) Improvement in performance was greater with participation in criteria development, especially for essential criteria (4) No improvement on "concealed" audit topic	p = .07 p = .02 p = .001 ——
Palmer et al, 1986, 1985, 1984, 1984[44-47]	Agreement with criteria Improvement in practice following feedback	711 practitioners, mostly doctors, working in 16 medical (8) and paediatric (8) group practices in Boston; 2 of the adult medical and 2 of the paediatric practices were in 4 teaching hospitals and the remaining 12 were departments of six teaching hospital-related health centres	Will patient care improve when practitioners participate in cycles of quality assurance (audit) facilitated by free technical assistance?	Crossover randomised controlled trial of cycles of quality assurance (audit)	8 tasks (4 for each setting) were selected to include some for which practitioners' motivation might vary and some for which changes in the system for delivering care might be needed to improve performance. In each of the 16 practices, there were two study cycles, one six months after the other. In each cycle, one of the designated tasks was subjected to experimental observation, ie, an audit cycle, while another paired task was a control. The experimental intervention consisted of revealing the name of the task under study to the practitioners, presenting the criteria to be used in the audit, and providing feedback on actual performance compared to the criteria. Practitioners could record their degree of agreement with the criteria. Cases which varied from criteria were subject to peer review and peer review findings were submitted to practitioners. Data collection was repeated following the intervention period.	(1) Quality of performance was: • Significantly improved in two tasks • Marginally improved in one task (2) Unimproved tasks were associated with the perceived need for system changes beyond the immediate control of the individual practitioner (3) Care deteriorated in paediatric practices in 3 of 4 tasks in the control state although the tendency was reversed when the tasks were included in the experimental state	p <.0001 p = .06

Author and Year of Publication	Part of Audit Process Studied	No and Type of Practitioners and Setting	Research Question(s)	Research Design	Intervention	Main Findings	P Value
Anderson et al, 1994[48]	Improvement in practice following feedback (also continuing medical education, CME, intervention)	15 short-stay hospitals in central Massachusetts; 3158 patients with multiple risk factors for venous thromboembolism	Will continuing education with or without a quality assurance (audit) component affect physician practices in the prevention of venous thromboembolism?	Hospitals were assigned to one of 3 groups each consisting of 5 hospitals, the groups being CME, CME + QA, or control	CME intervention was a multifaceted standardised educational programme which included feedback of data collected on actual practice (QA). QA intervention consisted of periodic reports to the medical staff and individual physicians on their compliance with criteria on indications for prophylaxis. In control hospitals, data were collected for a "retrospective observational survey" but were not made available during the trial.	(1) The proportion of patients at high risk for venous thromboembolism who received effective methods of prophylaxis increased significantly in all hospitals: • CME • CME + QA • Control • Increase in use of prophylaxis was significantly greater in CME + QA hospitals (2) There was no significant difference in the use of prophylaxis in CME and CME + QA hospitals (3) Traditional QA interventions appeared to provide no additional benefit to improve the use of prophylaxis	 p <.0001 p <.0001 p <.001 p <.001 — —
Lomas et al, 1991[49]	Improvement in practice following audit data collection and feedback versus following the use of opinion leaders	76 doctors in 16 community hospitals in Ontario, Canada	Which strategy will encourage implementation of a surgical practice guideline on trial of labour and vaginal birth after previous Caesarean section – audit with feedback or education using doctors nominated by their colleagues as opinion leaders?	Hospitals were assigned to one of 3 groups, audit + feedback (4 hospitals), opinion leader (4 hospitals), or control (8 hospitals)	In audit + feedback intervention, doctors locally developed criteria for the use of Caesarean section, data on actual practice were collected from medical records, and meetings were held every 3 months for feedback and discussion of the findings. In opinion leader intervention, local doctors nominated a colleague as educationally influential opinion leader and the leader made successive mailings of a practice guideline and supporting information, hosted a meeting with an expert speaker, and maintained a log of educational contacts with colleagues. In the control hospitals, the practice guideline was mailed to pertinent doctors with an exhortatory letter.	Doctors with opinion leaders who completed the trial (1 resigned) had labour and vaginal birth following Caesarean section rates higher than the other groups	p <.0001

Author and Year of Publication	Part of Audit Process Studied	No and Type of Practitioners and Setting	Research Question(s)	Research Design	Intervention	Main Findings	P Value
Mayefsky and Foye, 1993[50]	Improvement in practice with ongoing audit and feedback	18 paediatric residents (registrars) working in 2 hospitals in Rochester, NY	Will chart audit be useful to teach paediatric residents the components of a comprehensive well child visit?	Residents at one hospital composed the experimental group and those at the other hospital, the control group	Criteria were established for well child care and data were collected on compliance with the criteria by both groups. Doctors in the experimental group were given individual written feedback on his/her performance; doctors in the control group received no feedback. Data collection was repeated following feedback and one year later.	(1) Performance increased significantly in the experimental group, following feedback (2) Performance in the experimental group was lower one year later but still higher than in the control group	p <.05 p <.05
Winickoff *et al*, 1984[51]	Improvement in practice following different types of action	16 physicians specialising in internal medicine in a health centre in the Harvard Community Health Plan, Boston	Which intervention strategy will improve compliance with a standard established for colorectal cancer screening – educational meeting, retrospective feedback of group compliance rate, or retrospective feedback of individual compliance rate compared with that of peers?	Pre-test – post-test design for the first two strategies, then randomised controlled trial using crossover design of third strategy	Computer-based records were used to measure actual practice with the standard before-and-after (1), followed by (1) an educational meeting at which the standard was discussed and agreed upon and (2) another meeting to present before-and-after group performance data. Individual doctors were then presented with their own practice findings in comparison to peers.	(1) Compliance with the standard did not improve significantly following the educational meetings (2) In the first six months, physicians receiving individual feedback and those in the control group both improved (3) In the second six months, physicians receiving individual feedback (crossover group) continued to improve	—— p <.001 p <.001 p <.001
Winickoff, *et al*, 1985[52]	Improvement in practice	Physician/nurse practitioner teams working in a health centre in the Harvard Community Health Plan, Boston	Will a quality assurance (audit) project with either concurrent or peer comparison feedback directed at providers bring about improvements in the management of hypertension?	Teams were stratified by baseline compliance and assigned to experimental and control groups to test the interventions	Criteria were established for the management of hypertension and data were collected on actual compliance and fed back to the experimental group only.	There were no significant differences between the experimental and control groups	——

ORIGINS OF CRITERIA FOR AUDIT

Historical Development of Principles of Audit

Reviews of the history of audit often acknowledge Florence Nightingale for her contribution of the idea that observing actual practice and comparing outcomes of different practices can identify areas where improvement is needed. Her tabulations of the death rates from leg amputation of English soldiers in the Crimean war, accounting for whether the amputation was above or below the knee, indicated that soldiers operated on in large hospitals were more likely to die than those operated on in small hospitals. She attributed this finding to poor sanitation practices which led to rapid spread of infection among patients. She showed that the action of a doctor washing his hands between seeing patients reduced death rates.[53-54]

Variations between surgical techniques used for hernia repair and surgical outcomes of the procedure appeared in publications in the 1890s in Europe.[55] The need for a systematic collection of records about surgical operations to enable the evaluation of the results of such operations was raised in 1908 by Ernest Hey Groves at a meeting of the British Medical Association.[56] Groves circulated a letter to 50 large hospitals asking about the practice of keeping and publishing operation statistics in relation to mortality. Of the 50 hospitals, 27 replied, with only 17 claiming to have details showing operative mortality. Within the following ten years in the US, Ernest Amory Codman described and carried out systematic audit of the "end results" of surgical procedures.[57-58] Thus, the early principles were set out of a process which involved examining actual clinical practice and the outcomes of current practice as a basis for making improvements.

There was an early intention in the US to audit professional care in American hospitals through the Hospital Standardisation Programme created in 1917 by the newly established American College of Surgeons. However, the private funding which the College was able to raise for the Programme was insufficient for the intensive work that would be necessary to develop the methods and carry out the medical audits in all American hospitals.[59] The College therefore concentrated on five manageable standards for hospitals, one of which required doctors to audit their clinical experience through regular meetings in their own specialties, using patient medical records as a basis for such review.[60]

Lembcke's Scientific Principles

Lembcke, an American surgeon, was the first doctor to write about audit in scientific terms. He referred to confusion in the literature about the meaning of medical audit, stating that some authors referred to Frederick Taylor's methods for increasing industrial efficiency while others referred to methods for reviewing clinical practice through the use of statistics and individual case review.[61]

Lembcke recommended that medical audit make use of three basic methods which were: (1) analysis and evaluation of individual cases based on objective, written criteria, (2) use of epidemiologic data particularly to establish incidence rates to be used in the analysis of cases method, and (3) comparative analyses to determine the accuracy of diagnostic examinations, eg, laboratory, histopathological, or radiological investigations. Lembcke's "scientific method" for auditing cases was as follows:[59]

- Select a disease or operation or [subject for] medical audit based on meaningful classifications of data, eg, reliable and valid diagnosis or procedure indexes or other means to identify a homogeneous group of patients.

- Verify statements made in the clinical record, eg, confirm the diagnosis or current status by reviewing various data sources contained in the record for internal consistency.

- Establish the accuracy of the results of investigations by submitting the specimens or films for independent interpretation.

- Compare the verifiable facts with criteria established as being necessary or important for the care of patients included in the audit.

- Compare the degree of compliance with criteria with a standard degree of compliance found to be characteristic of hospitals "of acknowledged merit".

- Feed back the information periodically to those involved and responsible. (Lembcke provided reports to a committee which included hospital trustees, managers, and doctors, in addition to the specialists involved.)

Using the "scientific method" for medical audit as described, Lembcke was able to demonstrate a significant decrease in the number of unjustified surgical operations over a two-year period. Thus, Lembcke's work served to establish the following "criteria" for the audit process:

- Careful selection of the subject of the audit and cases to be included.

- Criteria and percentage standards as the basis for measuring actual performance.

- Precise description of the interpretation of the criteria so as to enable reliable measurement of performance using designated data sources.

- Feedback to practitioners and to those responsible of actual performance against the criteria and standards.

A similar nursing audit approach using valid, objective, reliable, and usable criteria for the appraisal of nursing care was developed by Reiter and Kakosh.[39] These and later efforts to establish audit methods were prompted by the creation in 1951 of the Joint Commission on Accreditation of Hospitals (now Joint Commission on Accreditation of Healthcare Organisations, JCAHO) which took on the work of the American College of Surgeons' Hospital Standardisation Programme. The Joint Commission required hospitals to supply evidence that audit of patient care was being carried out.

The Joint Commission's Criteria for Audit

The principles of audit established by Lembcke and Reiter remained until the 1970s when medical audit, then later clinical audit, became mandated in the US. A federal law which created Professional Standards Review Organisations (PSROs) required audit as a mechanism for monitoring medical care on behalf of the federal government. The Joint Commission on Accreditation of Healthcare Organisations required systematic audit in all healthcare facilities as a condition of accreditation.

In the face of the new PSRO law and its potential threat to the Joint Commission's role in accreditation and in medical audit, the Commission assumed a proactive role in promoting the implementation of audit through publication of a Performance Evaluation Procedure for Auditing and Improving Patient Care.[62] Influenced by attempts to link medical audit and medical education into a "bi-cycle"[63] and by knowledge from industry about various sources of performance failures, the Joint Commission's approach extended the audit process to include the following:

- Peer review of cases which do not comply with the audit criteria, ie, implicit review by peers of cases which are "screened out" by the application of explicit criteria.

- Formal identification of problems represented by cases which do not meet either explicit or implicit criteria and their underlying causes.

- Formal planning and implementation of action to remedy the causes of problems accounting for cases not meeting either explicit or implicit review criteria.

- Follow-up or reaudit to ensure that the quality of care has improved following the implementation of action.

A substantial contribution of the Joint Commission's work on audit was the systematic approach used to analyse the causes of problems which were shown via audit to be

affecting the quality of care. In some circumstances, feedback to practitioners in an educational context may be sufficient to achieve changes in clinical practice. However, the Joint Commission approach established a key principle that formal multi-faceted action plans may be needed to address all the causes of problems revealed through audit.

Another aspect of the Joint Commission's contribution to audit was unique: The focus was the patient not the professional and audit was of *patient* care not medical care. The Joint Commission was proactive in involving nurses and other professions in criterion-based audit and in evolving a multiprofessional team-based approach to audit.

Applications to Nursing and Ambulatory Care Audit

Some methods for audit of nursing and of ambulatory care used the criterion-based approach to audit but digressed from the model of studying indepth what happened to a group of patients who had a particular medical diagnosis or procedure. The numbers of rapid transactions involving large numbers of patients in nursing or ambulatory care, the numbers of individual professionals involved daily in the transactions, and in ambulatory or primary care, the long time interval between the processes of care and the occurrence of significant clinical outcomes all were factors which led to different approaches to selecting subjects for audit.

Nursing audit methods also attempted to analyse the relationships among the quality of patient care provided by nurses, the available number and skill mix of nursing staff, and the workload based on the number and dependency levels of patients.[38] Thus, in nursing, audit instruments were developed which aimed at providing composite measures of quality of care often with overall performance scores. Nursing criterion-based audit methods which followed the composite model included Phaneuf audit,[64] QualPacs (Quality Patient Care Scale),[65] Rush-Medicus Quality Monitoring Methodology,[66-68] and Monitor.[69]

Early approaches to audit in ambulatory care settings tended to address particular small discrete tasks of care individually and in succession, with the aim of covering many aspects of care over time, improving one then moving on to another.[70] Early examples of this approach have included the micro-sampling of records by a small group of doctors to identify problems to be considered by a quality assurance committee and quality assurance programmes which focused on problem identification and resolution.[71-72]

Further Technical Developments

Through the 1970s and 1980s, the criterion-referenced approach to audit has become widely accepted in the US and other countries and many examples of this

approach have appeared in the international literature.[73-77] Other approaches to defining the key aspects of care to be measured via audit were developed in the US, notably criteria mapping, disease staging, the tracer method, and the trajectory method,[78-81] some of which have been precursors of more advanced systems for large-scale surveillance of the quality of healthcare services or for developing algorithms or guidelines for care.

In 1989, after several years of experience in implementing accreditation requirements pertaining to audit, quality assurance, and quality monitoring, the Joint Commission launched an initiative to identify "indicators" of the overall performance of clinical services in healthcare organisations.[82] The methods used to define, develop, and test indicators as a basis for measuring quality of healthcare services have made substantial technical contributions to the field[83-88] and indicators are now widely accepted internationally by researchers and practitioners as a basis for measuring the quality of healthcare services.[89-98] The Royal College of Radiologists uses an indicator-based approach in its book of suggested audits for radiology services.[99] More recently, indicators have been used by researchers and practitioners as the basis for establishing performance measures for clinical services and healthcare systems.[100-107]

Criteria for Clinical Audit and the NHS

When medical audit was included in the NHS reforms as a formal expectation of doctors,[108] a strong professional tradition of audit already existed. The procedure for carrying out medical or nursing audit was not established uniformly, however. A variation of the audit process as described in the American literature was recognised in which audit consisted of observing current practice unrelated to explicit or implicit criteria with the findings of such observations being expected to lead to setting a standard of practice.[109] This approach recognised the considerable contributions made by a large number of observational studies and audits which were published in the UK in the late 1970s and 1980s[110-127] as well as the lack of valid research-based evidence which might be the basis for specific explicit audit criteria. Audit which observes current practice as a basis for evaluating performance and for identifying opportunities for improvement has been referred to as norm-referenced as opposed to criterion-referenced audit.[128]

In the last decade in the UK, nursing audit has developed in several directions. Composite tools for audit of nursing care, such as Monitor, as well as tools for audit of the management of nursing services have been used.[129] The Royal College of Nursing has led the promotion of standard setting in nursing and quality improvement of nursing care.[20, 130-131] The principles followed in this work have built on previously published models. More than medical audit, nursing audit has appeared to emphasise defining and implementing achievable standards prior to measuring actual practice.[132]

Over nearly 150 years of published work which has established principles for audit of healthcare services, the continuing theme is that knowledge of current practice should

be applied in order to bring about improvements which benefit patients. This understanding of audit is reflected in the NHS Executive policies on clinical audit which describe the process as a systematic method of meeting and improving standards in healthcare and of improving the quality of patient care.[1, 133-134] Specifically, the NHS Executive defines an effective clinical audit programme as one which involves balanced topic selection, employs adequate audit processes, secures implementation of audit results, and is comprehensive. Effective clinical audit "leads to improvements in health outcomes and other aspects of quality, through changes in clinical practice and systems of healthcare delivery."[134] Thus, the focus of audit on acting to achieve measurable improvements is fully consistent with NHS policy.

THE AUDIT PROCESS AND SELECTED LITERATURE

Deciding What, Why, and Who to Audit

The Process

The audit process generally begins with someone making a decision to undertake audit. In these early stages, decisions are made about the following:

- The subject — the patient, client, or user group or the events involved and the particular aspects of quality of care for the group or events which the audit will address.

- The rationale — the background to selection of the subject or the reasons for undertaking the audit.

- The group — the individuals or groups involved in or affected by the audit and their roles in the entire audit process.

- The objectives — the specific purposes of the audit which reflect the attributes of quality to be addressed by the audit.

With these decisions made, the process moves on to define specifically the patients, clients, users, or events to be covered in the audit; the time period for the audit; and the strategy for data collection.

These early decisions about an audit can be made by individuals or groups who may voluntarily wish to undertake audit or who may be assigned to lead or to undertake audit.

The Questions

Questions concerning the selection and design of an audit might include the following:

- Does it matter to the effectiveness of the audit if the subject and objectives for the audit are selected by the group undertaking the audit or are imposed from outside the group?

- Does the probability of improvements being made as a result of the audit vary by whether or not the audit subject and objectives are selected by one individual within a professional group or by the group as a whole?

- Is the effectiveness of the audit likely to be improved if there is a reflective group process involved in choosing what and why to audit?

The Literature

Little research evidence specific to audit can be identified which provides guidance for making the decisions about what and why to audit and who should be involved in an audit. Those emphasising audit as an educational process for clinicians discuss the values of individual self-assessment, personal development, and teamwork applied to medical audit,[135-136] and the implications of these professional values are that individuals and teams should be actively involved in this important stage of audit.

Williamson described the use of a modification of a structured group decision-making technique, nominal group process, with medical and with multiprofessional groups to select topics for audit.[137] Williamson advocated the guiding principle of identifying areas of "achievable benefit not achieved", ie, areas where clinical effectiveness could be improved. The technique involved a systematic process in which practitioners generated audit topics, reviewed available information and literature about the suggested topics, then weighted the topics in terms of potential impact on patient care. The technique was subject to both reliability and validity testing in studies involving practitioner groups in up to eight healthcare settings.[138-139]

A structured approach to team decision-making on audit topics, called Quality Impact Analysis, used a variation of the Delphi process in which the group identifies frequent, risk-associated, and problem-associated aspects of their current services, along with the potential for enhancing effectiveness or cost if desired, and then rated the ideas generated, selecting topics for audit which are scored highly across all factors.[95] Hirschhorn *et al* advocated that organisations should first set their business objectives and then systematically select audit topics which would address the objectives and any problems identified.[140]

Key Point

It appears to be desirable that those whose work will be covered by an audit participate in selecting the subject and setting objectives for the audit, using a systematic method which focuses on important subjects for audit as perceived by the participants.

Defining and Measuring Current Practice

The Process

Once a subject and objectives for an audit are chosen, decisions have to be made about the nature of what is to be measured in the audit and how measurement will be carried out. There is considerable confusion in the literature concerning the terms used to describe the specific measures used in audit. The words used most commonly are criterion, indicator, standard, and performance measure. Early explications of the terminology were provided by Donabedian.[141-143] Meanings used in this review are provided in the Glossary of Terms.

Whatever the terms used, the functions to be achieved in measuring clinical practice appear to include specification of the following:

- The care or service which is to be measured and the specific attributes of that care, eg, timeliness of CT scan of patients referred appropriately for the procedure, or effectiveness of care by a multiprofessional team for people who have had a stroke.
- The specific aspects of care or service which are to be measured, eg, indications for CT scan in the patient group under audit.
- The percentage of patients, clients, or users who are expected to receive the care described.
- Clear and complete operational definitions of all terms used and the data to be collected.
- Data sources and methods to be used for data collection.
- The procedure to be used to select cases to be included in the audit.

Some sources recommend specifying in advance any exceptions, ie, cases or circumstances for which the care described may not be given and which would still be acceptable to clinicians.[62,144]

The Questions

Questions concerning measurement of current practice might include the following:

- Is the use of explicit criteria better than implicit criteria?

- Who should develop measures for audit purposes — are individual practitioners more likely to improve their practice if they set criteria for audit themselves, either individually or in a group, as opposed to using criteria developed by others?

- Are some methods for developing criteria in a group more likely to produce consensus among professionals than others?

- What are appropriate external sources for standards – expert opinion, normative practice, research-based evidence, or best achievable practice elsewhere?

- Is it better to use outcome measures only or outcome and process measures or process measures only?

- Are some sources or methods for data collection likely to produce more valid or more reliable data than others?

- Does direct participation in data collection affect practitioners' readiness to change practice?

The Literature

Some studies are highly focused on one research question; others address several questions in the same design. The findings presented here are organised under the following headings: explicit versus implicit criteria; use of self-developed or other-developed criteria; methods for achieving consensus on criteria; outcome versus process measures; data sources and collection methods; and direct participation in data collection.

Explicit vs Implicit Criteria. In studies on explicit versus implicit review, explicit review means that individual cases were judged in comparison to objective criteria which were specified in advance. Implicit review in studies means that individual cases were judged by experts, using their clinical judgement. In studies involving implicit review, typically, experts were given a form structured with five-point rating scales, eg, very good to very poor, to record their judgements on the quality of care. In recent approaches to using implicit review methods, experts discussed the bases for their judgements among themselves before they made judgements about individual cases. (See section on methods for achieving consensus on criteria.) The studies on explicit versus implicit review tended not to provide details on how the explicit criteria were developed nor on how the validity of the explicit criteria was established.

Three early US studies examined the use of implicit versus explicit criteria and the findings of two of the studies appeared to favour explicit review. Richardson described a series of studies undertaken in the late 1960s in New York state in which 81 qualified medical experts reviewed and made judgements about the same 1523 cases, using

medical records. He observed that the extent of discrepancies in intra-disciplinary judgements of the quality of care "cast doubt on the validity and reliability of uncontrolled peer judgements of this quality."[145] Subsequently, Richardson developed an audit method based on explicit criteria and, through a number of tests, demonstrated the improved validity and reliability of the explicit approach.[146] Richardson also appears to be the first researcher to demonstrate that non-doctors, with training, can be used to abstract information for audit with a high degree of inter-rater reliability with doctors.

Brook and Appel raised two fundamental questions about audit methods in a study published in 1973: how judgements should be made on the quality of care provided, ie, explicit versus implicit review, and whether data on process or outcome or both should be collected and analysed.[147] They assessed five peer review methods in the setting of the Baltimore City Hospitals. Care provided to 296 patients with urinary tract infection, hypertension, or an ulcerated lesion in the stomach or duodenum was evaluated by each of five methods. The ten doctors who made the judgements in each method were faculty members of Johns Hopkins University School of Medicine. Abstracts of the cases were used as the basis for judgement, not the actual medical records, in order to avoid bias attributable to the legibility or organisation of actual records. Explicit criteria were developed by a group of 21 faculty members not participating in the review of cases. Doctors involved in the peer review methods made implicit-process, implicit-outcome, implicit quality-of-care (both process and outcome), explicit-process, and explicit-outcome judgements.

The findings illustrated the variation in peer judgements even in controlled circumstances. With two of three judges constituting agreement among peers using implicit review, 23.3 percent of the patients had an adequate medical care process, and 63.2 percent of cases were judged to have unimprovable outcomes. Only 1.4 to 2.0 percent of the cases met all the explicit process criteria, depending on the condition. Cases met on average between 35 and 38 percent of the criteria. The researchers reported that it was difficult to draw conclusions about explicit outcomes due to small numbers.

In a research report published in 1979, Hulka and colleagues obtained significant positive correlations between ratings based on implicit and explicit review.[148] Implicit review was based on actual patient records. Most of the reviewers had previously participated in an earlier phase of the study which included developing explicit criteria for the same conditions used in the implicit review. As with other studies, the implicit method produced higher ratings on cases which had low or medium levels of compliance with explicit criteria.

More recently, Hayward, McMahon, and Bernard evaluated the ability of implicit review to measure reliably various aspects of care on a general medicine inpatient service.[149] They undertook retrospective review of 675 medical records of a stratified random sample of patients admitted consecutively to a general medicine ward in a university teaching hospital, using 12 faculty members or fellows who were specialists in internal medicine. Approximately 20 percent of the records were selected randomly for

multiple independent reviews to allow for reliability testing. The doctors were trained in structured implicit review and used a structured implicit review instrument. Inter-rater reliabilities were adequate (Kappa = 0.5) for aggregate comparisons on measures relating to overall quality of care and preventable deaths, but were inadequate for reliable evaluations of individual patients. Reviewers did not agree on whether errors in following doctors' orders, dispensing medications, or notifying doctors had occurred and there was a poor level of agreement on judgements related to the timeliness and appropriateness of diagnostic evaluation and therapeutic interventions.

On the other hand, researchers from the Health Programme of Rand Corporation have asserted that quality of care can be judged via either implicit or explicit review.[150-151] Care provided to a sample of 1,366 patients with congestive heart failure, acute myocardial infarction, pneumonia, cerebrovascular accident, or hip fracture was judged via implicit review by 25 doctor reviewers. Findings of implicit review were compared with explicit process reviews to study whether the two methods of process review produced similar results and to examine the predictive validity of implicit process review.

The scores between explicit and implicit reviews varied: In 81 percent of case comparisons between implicit review and explicit criteria, patients whose care was rated good or very good on implicit review had higher explicit review scores than patients whose care was rated poor or very poor on implicit review (whose scores were consistently lower on explicit criteria). For 13 percent of comparisons, patients had equivalent scores and for the remaining comparisons, patients whose care was rated good or very good on implicit review had lower explicit scores than patients rated poor or very poor on implicit review. Mean inter-rater reliability (Kappa) scores on the use of the implicit review instrument were mostly between 0.4 and 0.7. The researchers concluded that implicit review is "as reliable as many other clinical measurements which physicians perform," although a single review may not be reliable enough to judge accurately whether an individual patient received poor care. In a related study, the researchers observed that there was remarkable agreement between implicit and explicit process measures and sickness-adjusted outcomes at different types of hospitals.[152]

It has been asserted that implicit review of medical records by doctors has been the community gold standard for judging medical care.[153] In analyses on the question of explicit versus implicit review, Donabedian has cited the advantage of implicit criteria being that they are adaptable to the precise characteristics of each case individually and the disadvantages being the costliness of use of implicit review, as it relies on experts as reviewers, and the relative imprecision of implicit criteria.[154-155]

Donabedian attributed the following advantages to the use of explicit criteria: simplicity; lower cost of application even though their development can be time consuming; the ability to establish the validity of the criteria; a means of expressing practice expectations to involved practitioners; enabling personal review of everyday practice; influencing actual practice; making possible the opportunity of representing consumer views; and supporting equity in service provision. The disadvantages of explicit criteria,

according to Donabedian, are the need for criteria to be specific to conditions or services under study and the possibility of inappropriate application.[154-155] Donabedian concluded that taking into account the strengths and limitations of explicit and implicit criteria, it may be best to use both in sequence or in combination. This is the approach taken in some approaches to audit.[62, 144]

Self-Developed or Other-Developed Criteria. Studies have explored the degree of agreement on criteria between experts and practitioners and whether or not practitioners change their practice to conform to audit criteria developed either by the practitioners themselves or by others. Key variables which could affect the results of the studies include the inherent motivation of practitioners to study and improve their care and the processes used to develop criteria, including the quality of the facilitation of individuals or groups involved. These variables tend not to be explored in the published studies.

In 1971 in the US, a Joint Committee on Quality Assurance (JCQA) was formed to include representatives of major national organisations whose members provided primary care for children for the purpose of providing empirically-based criteria and a tested audit methodology.[156] The JCQA involved 129 doctors in developing and refining criteria by self-audit and then used 396 experts to validate the criteria using the Delphi method. Four clinical subjects were used: child health surveillance (in four age bands), tonsillopharyngitis, urinary tract infection, and asthma. The resulting criteria were rated by a national representative sample of 1,329 doctors providing primary care to children. Finally, the criteria were used by 166 doctors in audit in actual care settings.

In the initial self-audit phase, practitioners generated a total of 199 criteria for all the clinical subjects. The experts added 12 additional criteria and validated 73.5 percent of the practitioners' criteria as relevant. Overall, the national representative sample of doctors agreed on the relevance of 94.9 percent of the criteria circulated. There were statistically significant ratings by type of doctor: fewer criteria pertinent to child health surveillance of young children were deemed to be relevant by non-paediatricians. The doctors involved in the actual audit volunteered to do so; it is unknown how many, if any, participated in previous stages in the study. Only 16.7 percent of the criteria were recorded in 75 percent or more of the 10,500 records in the study. In total, 45.8 percent of the criteria were recorded in more than 50 percent of the records. In a separate finding, doctors made estimates of the amount of relevant information recorded in records which were higher than percentages confirmed by the audit. The study illustrated that criteria generated by a small group of practitioners can be refined and validated by experts and by large numbers of representative practitioners. However, audit of actual documentation may not confirm practitioners' conformance with the criteria.

In the Hulka *et al* study involving 31 specialists in internal medicine in four population centres in North Carolina, three research questions about criteria development were addressed: Is there less than perfect adherence to self-selected criteria among practising

specialists in internal medicine; is nonadherence to the criteria related to methodologic problems of the research itself; and how valid is the explicit criteria method as a measure of quality of care?[148] Four medical circumstances were used: the follow-up care of patients known to have diabetes; the initial evaluation of patients with hypertension; the general examination; and acute dysuria in the female. The study covered the care provided to 1,334 patients with one of the conditions. Doctors developed their own criteria and the Delphi process was used to develop doctors' consensus on criteria applicable to the four conditions.

The findings included that evidence pertinent to the consensus criteria were more likely to be found in the doctors' records (61 percent of records met consensus criteria versus 22 percent of records met non-consensus criteria). Of the 99 criteria which were present in over 50 percent of records, 86 percent were in the consensus criteria list. The researchers tested whether doctors were more likely to follow their own individual criteria or their consensus criteria: adherence to consensus criteria was greater than to individual criteria for three of four conditions ($p < .001$ for diabetes and hypertension; $p < .05$ for dysuria). Also, adherence to criteria generated by the group as opposed to criteria available from elsewhere was greater for three of four conditions ($p < .01$). In the study, doctors' conformance with the criteria was shown to relate to doctors' workload as measured by the number of patients seen per hour; for busier doctors, there was less evidence in the records of conformance to criteria particularly of information pertinent to history taking or physical examination.

As part of a randomised controlled trial of medical audit in 16 primary care group practices in Boston, Palmer and colleagues measured the degree of doctors' agreement with criteria developed by the researchers for use in the trial and the percent compliance with the criteria by the same doctors, based on review of their records. Only four of 203 paediatricians' responses indicated disagreement with the criteria pertaining to conditions treated by paediatricians; however, from one percent to 47 percent of cases involved in the audit did not conform with the criteria.[45, 47] Specialists in internal medicine involved in the same trial varied more than the paediatricians in their degree of agreement with criteria used in the trial. Among 316 doctor responses, agreement varied from 57 percent to 100 percent. When the doctors' records were compared with the criteria, from 6 percent to 42 percent of cases did not conform with the criteria.[46]

In a small study, Norton and Dempsey used a crossover design involving six family practitioners to assess if doctors' involvement in creating their own criteria and receiving immediate feedback on their performance produced an improvement in performance.[157] Data were collected on the doctors' care of patients with vaginitis and cystitis for two months prior to the doctors' agreement to produce their own criteria lists on one of the conditions. Doctors in two different groups created personal criteria for one of two conditions, cystitis or vaginitis, then were given immediate feedback on their own performance on the condition for which they wrote criteria, and were allowed to change their criteria. Six and 14 months later, data were collected again on all doctors

on their treatment of both conditions. The findings illustrated improvement in doctors' performance at both six- and 14-month intervals for the experimental condition.

Putman and Curry[13] found that participation in the generation of criteria was followed by a significant improvement in doctors' behaviour; interestingly, in the same group of doctors, doctors' involvement in the selection of the condition to be audited was not associated with changes in practice.

UK researchers have demonstrated that active involvement in setting standards for clinical care by general practitioners was associated with improvements in the doctors' prescribing and follow-up practices. Anderson *et al* [158] reported that one year following development and discussion of a protocol for monitoring treatment with digoxin in general practice, monitoring practices had improved significantly among principals who had carried out the audit but not among other principals in the same practices.

In the North of England Study of Standards and Performance in General Practice,[159-160] ten groups of general practice (GP) trainees in the North of England participated in setting clinical standards for two childhood conditions, one for training purposes and one for purposes of the study. In addition, a standard for each condition was drawn up by five mixed groups of specialists, which included paediatricians and general practitioners, the standards were exchanged with one of the GP trainer groups, and the two groups met, after which the trainer group could finalise its standard. Also, the trainer groups each experienced different types of medical audit for three of the four other study conditions, the other medical audit interventions being receiving a standard set by another group; receiving data on their performance as a group compared with the performance of all other groups; receiving data only on their performance as a group; and receiving neither a standard nor data, ie, the control.

Data on the GPs' actual clinical practices were collected from patient records and separate enhancement records, through the use of survey questionnaires with parents, and via activity analysis within practices. Data were collected for up to two years after the completion of standard setting. Only two types of information in records showed a significant change associated with the setting of clinical standards, those on drug management and follow-up decisions. In this study, there was no evidence that the other types of medical audit, ie, receiving a standard, receiving comparative data, and receiving descriptive data, had any effect. Also, the effects of standard setting were no greater for groups who received specialist input than for those who did not.

Several standardised nursing audit instruments[64-69] rely on the use of criteria which are defined within the audit instrument and which have been subject to reliability and validity testing, and thus are not developed locally by nurses. Several of the nursing audit instruments have been subject to evaluation in the UK; however, the evaluations do not specifically address which approach – self-developed or other-developed criteria – enables more significant improvements in the performance of nursing care. In an

analysis of the issues involved in formulating and applying criteria, Kitson *et al* noted that the research questions raised by Hulka et al[148] still demand answers from nursing.[161]

Methods for Achieving Consensus on Criteria. In addition to the studies already cited which have developed and tested methods for establishing criteria, others have made relevant contributions to the question of how criteria can be developed by professional groups. Elements of a method for developing consensus on reliable and valid criteria across practitioners have been described as including the following: development of criteria from state-of-the-art clinical literature; refinement of the criteria by expert panels using the Delphi process; assignment of differential item weights to the criteria which remain following the successive rounds of the Delphi process and identification of criteria which yield little information; development of a scoring system; and training and inter-rater reliability testing of data abstractors.[162]

This overall approach was followed in a study carried out in Scotland.[163] A postal questionnaire survey involving 121 consultant gynaecologists was useful in validating 19 of 20 criteria on induced abortion to be used subsequently in a national audit project.

The extent of agreement on proposed criteria may vary considerably, however. Improvement in the degree of agreement may be achieved through having the practitioners meet and discuss indicators which they had previously rated by post. However, in a US study involving doctors' views on indications for six medical and surgical procedures, agreement following discussion was still only on 42 to 56 percent of indicators, excluding one outlying procedure.[164] In another US study, one group of surgeons was more likely to favour surgical treatment for carotid disease than a multispecialty panel which included neurologists and other specialists; however, the two groups reached consensus on inappropriate indications for such treatment.[165]

In the UK, some groups of general practitioners meeting to set standards for audit were observed to function to achieve the task, whereas in others, the behaviour of individual group members was observed to have had an effect on the meeting outcomes. The authors of the observations concluded, "Doctors taking part in collaborative audit projects should therefore have some appreciation of the skills needed to facilitate meetings and groups .... In groups with a facilitator ... this task is much simplified; otherwise groups have to cultivate a critical attitude towards their own activities."[166]

Kitson and colleagues studied whether and to what extent criteria sets derived from an expert group of nurses differed from criteria generated by local practitioner groups of nurses on the subject of postoperative pain management.[167] The nominal group process and the Delphi technique were used by the researchers working with the expert group to devise a set of criteria. The nominal group process and brainstorming were used by facilitators working with five local groups of nurses from five surgical wards in different hospitals. The expert group generated more criteria and the criteria were more specific. Convergence on criteria across any two ward-based groups varied from 36 to

55 percent agreement. The findings thus illustrated "the significant role the facilitator plays in ensuring an effective process."

External Sources of Criteria. Previously described studies have suggested the circumstances under which expert opinion or expert-developed criteria or standards may or may not be accepted. Research which compares the effectiveness of norms with other sources of criteria or standards could not be located. Studies on the use of antibiotics and laboratory investigations have illustrated difficulties in achieving changes in practice based on patterns of normative usage.[168-172] Examples of the use of benchmarking to best known practice have been published,[173-175] although research studies on the effectiveness of benchmarking compared to other sources of criteria could not be located.

Evidence-Based Measurement. Two issues pervade most of the published work pertinent to measuring clinical practice: the scientific base for the measures used and the adequacy of available data sources to supply reliable evidence of actual practice in relation to the measures. Some of the difficulties reported in developing consensus on valid criteria by which to assess care may be attributable to the lack of a body of current scientifically-sound knowledge of the processes of care which are most likely to contribute to desired outcomes and to prevent undesired ones. Donabedian summarised this point: "It cannot be emphasised too strongly that our ability to assess the quality of technical care is bounded by the strengths and weaknesses of our clinical science."[176] Many of the studies concluded that evidence of actual practice in patients' medical records did not support doctors' perceptions of their own practices. The studies tended to note that audit might be limited by the quality of recordkeeping and to recommend consideration of more structured and comprehensive medical records.

Outcome vs Process Measures. A number of studies have attempted to assess the correlation between processes of care and outcomes or between compliance with audit process criteria and patient outcomes. The findings have illustrated the complexity of the relationships. Hulka and colleagues found no statistically significant associations between compliance to process criteria and outcomes.[148]

The North of England Study assessed functional, psychological, educational, and condition-specific clinical outcomes for the childhood conditions for which GPs developed standards, via interviews and a postal survey with parents of the children whose care was included in the standard-setting part of the study.[160] Only one of the 15 analyses of the data which were undertaken showed a change that could be attributed to standard setting: children who had been seen by GPs who had set standards for recurrent wheezy chest showed a much greater improvement than other wheezy children. These findings were consistent with evidence gathered from medical records that the GPs' practice had changed following standard setting.

Recent argument favours studying the processes of care along with or in preference to outcomes of care.[151] What happens to patients, ie, outcomes, may be the most

meaningful measures of quality of care; however, the following advantages to process-outcome or process measures have been asserted:[37,148,151,176]

- Outcomes do not directly assess quality of performance; they only permit inference about the quality of the process (and structure) of care. The degree of confidence in that inference depends on the strength of the predetermined causal relationship between process and outcome (and structure and process).

- Outcome assessment must include correction for factors such as patient compliance with recommended treatment or case mix, for which sufficiently complete methods of standardisation may not yet be available.

- Some meaningful outcomes are delayed in appearance and it is difficult to obtain reliable information about their occurrence.

- There is insufficient scientific evidence about the contribution of some healthcare professionals' practices to patient outcomes.

- Not all patients who experience a poor process of care suffer a poor outcome.

- The achievement of outcomes depends on the integration of several potentially unrelated inputs working within a system, and it may be difficult to isolate with certainty the specific factors which have contributed to outcomes.

- Practitioners are unable to develop clinical methods to improve outcomes unless there is understanding of how outcomes and processes are related.

- Process measures are less expensive to use and are valid when based upon aspects of care which have been proven by research to be effective in improving patient outcome.

- Outcome-only data can be subject to misrepresentation and misunderstanding by the public.

Data Sources and Collection Methods. The studies which have examined the criteria and data collection phases of audit have tended to recognise the problems with reliance on patient medical records as reliable sources of data for audit purposes. Alternative strategies for assessing the performance of doctors have been investigated, using a criteria map for the outpatient management of chronic obstructive pulmonary disease as the basis for the assessment.[177] The methods used were doctor interview, patient interview, videotaped observation, and abstracting of information from the patient record. A total of 30 patient-doctor encounters were included, with 13 specialists in chest medicine participating. For the 24 patients with whom all methods were applied, the two interview methods were found to supply more information about compliance with the criteria. On criteria pertaining to medication management, investigations,

and identifying complications, the videotaped observation provided data comparable to the interviews. The doctors reported, however, that 35 percent of the time the interview process tended to change the nature of the visit. The record-based audit was the least costly per patient. The researchers concluded that no method excels in all ways as a strategy for evaluating doctor performance in the ambulatory care setting and that such evaluation should consider care over a number of visits and match methods with specific objectives of evaluation.

Lawthers et al[103-104] have pointed out that a crucial and often-missed step in performance measure development is evaluation of the measures. They have recommended that reliability testing, both of alternative data sources and of data abstracting, be carried out, as well as validity testing.

Direct Participation in Data Collection. A small US study carried out in 1986 addressed questions about participation in audit by doctors in training (residents): do results generated by other residents improve trainees' compliance with standards; do residents who spend time doing actual audit data collection improve their compliance more than trainees who simply receive the results of others' audits; and if there is improved compliance by "self-audit", what is the cost of the added improvement.[178] The study used a combination of research designs: a before-after design to assess the efficacy of passive receipt of audit findings, and a randomised controlled design to assess if residents who collected data from records themselves performed better than residents who just received results. The two subjects for study, ordering influenza vaccination and ordering screening mammography, were selected because well-defined criteria for the procedures were accepted, several possible data sources were available for both subjects, and comparisons could be made with studies in other locations. Residents were stratified by training year and assigned at random to one of three groups, one assigned to audit their own records for vaccination and one to audit their records for mammography, and the third had no assignment.

Prior to the intervention, data on the doctors' performance against the criteria had been collected independent of the residents. Following the intervention, all doctors were provided with the results of both of the self-audits as part of educational sessions on the subjects. Subsequently, unknown to the residents, an independent audit of practice related to both procedures was carried out. The subsequent audit of all residents' performance showed that decision making about influenza vaccination was significantly improved for all groups and there was no difference among the groups. Decision-making about screening mammography also was significantly improved and the residents who audited their own records improved significantly more than the others. In terms of cost-effectiveness of the time spent in audit, the authors found that self-audit achieved better compliance with standards but at roughly twice the cost of the control group who received information about colleagues' performance and education only. The researchers suggested explanations for their findings: improvements in vaccination rates did not occur because the overall rates were reasonable before the audit or patients may accept vaccination more readily than mammography and residents needed to be motivated to explain the mammography procedure to patients.

Key Points

Measures of practice can be effectively used in audit to contribute to improvements in practice when they are developed by or otherwise made acceptable to the practitioners whose performance is to be reviewed.

Implicit measures of quality of care appear to be less reliable in audit than explicit measures.

Measures of processes and outcomes of care or process measures alone are preferable to outcome measures alone.

The patient medical record may be the most readily available and least costly data source for audit in many organisations; however, data from other sources may be more complete.

Evaluating Practice Based on Measurement

The Process

Following collection of data on current practice in comparison to measures, the evidence of actual practice is evaluated formally. In audit, evaluation may involve four tasks:

- Review of cases found to not meet the audit criteria in order to verify whether or not any variation from the criteria is clinically acceptable.
- Comparison of the findings of measurement to the criteria, with adjustment for review of any acceptable variations, in order to determine if actual practice meets expectations expressed in the criteria.
- Feedback of the findings of the comparison to those whose work is covered by the audit.
- Analysis, and possibly investigation, of problems represented by instances of criteria not being met, and the causes of the problems.

The evaluation can be carried out with the involvement of those affected by the audit and may include the use of group problem-solving or quality improvement techniques.

The Questions

Questions concerning the evaluation of actual practice might include the following:

- What factors contribute to an effective "peer review" process?
- What are effective ways of presenting the findings of measurement of current practice to professionals, with the aim of having open inquiry about how actual practice could be improved?
- Does feedback of findings alone work to achieve changes in practice?

The Literature

Few published studies have explicitly investigated the evaluation and "peer review" phases in audit in detail. The literature tends to refer to the whole phase of audit as "feedback" and is imprecise and confusing about the nature of the processes involved.

The randomised trial by Palmer *et al*[45] appears to be the only study which described the process in detail. The approach in the study parallels the audit principles established earlier by the Joint Commission and provided for the following:

- Review of all cases which did not meet the explicit criteria by one or more members of the "peer group", called "peer reviewers," using a structured implicit review process.
- Presentation of the findings to the group involved, both the percentage of cases meeting the explicit criteria and the percentage accepted by the peer reviewer.
- Discussion of the findings by the peer group, ie, the implications for patients and practitioners.
- Planning of action to improve care, if needed, as indicated by the cases which did not meet the explicit criteria and were not accepted as appropriate care by the peer reviewers. The action plan may include providing feedback to all those involved or concerned with the subject of the audit.

Much of the literature assumes that feedback and education are synonymous, ie, that practitioners change their behaviour after being told about their current performance. A few sources distinguish clearly between feedback and educative processes. The process of providing feedback involves supplying information to others, eg, "telling" practitioners about the audit findings, whereas the process of providing education involves designing and supplying an experience which changes attitudes or behaviour, eg, "teaching" practitioners about the importance of the audit findings *and* enabling

practitioners to learn new behaviours.[62,144] An early US study pointed out the significance of the distinction between identifying knowledge gaps of practitioners and other reasons for their performance patterns. Ashbaugh and McKean[179] surveyed reasons for non-compliance with criteria in 55 individual audits and concluded that only six percent of "deficiencies" revealed by audit were due to lack of knowledge; the remainder were due to systems-related and other causes.

The evaluation process in audit should also include an assessment of practitioners' readiness to change practices: Do practitioners know there is a variation between their practice and agreed criteria for practice? Do they accept that the variation is important? Do they know how to change, assuming acceptance of the need to change? Is the change required within their personal control to achieve? Are the practitioners likely to maintain any changes in attitude or behaviour over time in the absence of strategies to sustain change? The use of feedback alone where implementation of change involves knowledge or skill which the practitioners may not yet have acquired or providing feeding when what is needed is change in systems which is beyond the control of the practitioner to address has been informally called "shouting at the problem."

Evaluation and Peer Review. In the Palmer *et al* RCT of audit in 16 group practices,[45] a very explicit process of evaluation was used: For each group practice represented in the study, cases which varied from one or more of the explicit audit criteria, based on the data collection phase of the audit, were identified and presented to one of the practice doctors who had agreed to serve as peer reviewer. The peer reviewer analysed each case, looking specifically for the reason why the criteria could not or should not have been met. If a specific reason was not found, the variation was called "unjustified", ie, the variation constituted a deficiency in care. The peer reviewer presented the findings, including the percentage of unjustified cases, for discussion at a meeting of the doctors working in the practice, focusing on the causes for deficiencies and on plans to improve care. Individual doctors received separately a report of findings and detailed case summaries of any of their own cases which had been subject to case evaluation by the peer reviewer.

Findings of evaluation presented by the peer reviewer were not disputed in any of the practice sites. Reviewers varied widely in the leniency of their judgements, some accepting broad interpretations of the criteria or implied or intended though not documented compliance. Lenient peer reviewers tended to justify obvious noncompliance with criteria if they judged that the patient was unlikely to be harmed. Leniency of the peer reviewer was unrelated to the decision to take corrective action. Peer reviewers' judgements were occasionally overruled at meetings by group consensus, in both directions of more leniency or more harshness. Peer reviewers tended to develop strategies for motivating colleagues to improve by avoiding dispute with colleagues on marginal cases and focusing on what could be done about obvious deficiencies in care.

In a different model for peer review, radiology review committees were established in six centres participating in a study of the influence of Royal College of Radiologists' guidelines on hospital practice.[180] Among the committees' tasks was to monitor

compliance with the guidelines, having distributed them to clinical users and considered local adaptations of the guidelines. Initially, all committees expressed reservations about intervening if the monitoring identified certain firms of doctors whose practice was at variance with that of colleagues in the same specialty. Four committees agreed to intervene and a fifth agreed in principle but not on how the intervention was to be carried out. Criteria for when intervention would be used included that a firm had referral rates that were twice the average local specialty rate for at least three consecutive quarters; initially high and steadily increasing rates; or three times the specialty rates in other centres. The committees' intervention ranged from informal discussion with colleagues to a formal presentation to all medical staff.

Feedback. Because the literature on audit and feedback tends to refer to feedback as a strategy for achieving change, a summary of relevant studies is included in the next section.

Key Points

A two-phased strategy of (1) using explicit criteria to measure quality of care across several cases and (2) using structured implicit criteria to review the individual cases which do not meet explicit criteria can enable practitioners to focus on potential reasons why there are cases which do not meet either explicit or implicit criteria.

Audit leaders can motivate colleagues to improve by focusing on what can be done about obvious shortcomings in care provided to patients.

Acting to Improve Practice and Maintaining Improvement

The Process

If evaluation of the audit findings suggests the need for improvements in actual practice, the group responsible for the audit plans action needed to achieve the improvements. At the simplest level, an action plan may consist of implied agreement among practitioners to modify their personal behaviours, eg, document important aspects of patient care more consistently or change current prescribing patterns. Many improvements are achieved through this approach when the behaviours involved in the audit are under the sole and direct control of individual practitioners and when practitioners can accept the value of change.

For several audit findings, however, substantial problems impede the provision of care which would be consistent with agreed audit measures. In these situations, investigative

and analytic processes may be required to uncover the true causes of the problems and action may need to consist of a substantial operational programme to address directly such causes. Historic patterns of resource allocation, lack of systems, persistent breakdowns or failures in traditional systems, redundant paperwork, or lack of communication across professional or service lines are examples of causes of problems which simply can not be addressed by individual practitioners or even a single group of practitioners. In such situations, corporate management commitment and additional problem-solving skills may be needed to achieve any improvements in patient care.

The Questions

Questions concerning taking action as part of the audit process might include the following:

- What analytic methods work to facilitate identification of causes of problems which are demonstrated to be impeding practitioner performance?
- What strategies are effective in achieving and maintaining changes in practitioner behaviour when such changes are under the sole and direct control of individual practitioners? Do the same strategies work equally with all healthcare professionals?
- Are single strategies more likely to be effective than combinations?
- What proportion of all needed improvements in patient care as identified by audit can be made solely by involved practitioners?
- What organisational structures and processes are effective in supporting achievements in improvements in patient care which are indicated by audit findings?

The Literature

As stated earlier, the literature tends to feature "feedback" related to audit and does not reflect the true complexity of analysis of the behavioural, attitudinal, or systems changes required to complete the audit cycle.

Pragmatic advice to practitioner groups to select subjects for audit on which the group itself can act to improve practice tends to lead to avoidance of audit of perceived problem areas for which individuals or specific practitioner groups working in isolation cannot achieve improvements. There may be considerable traditional barriers to attempting to work together with other practitioner or management groups on audit of such problems.

Payne *et al* reported that improvements in the process of medical care were related directly to the intensity and duration of planned interventions by the group involved and followed organisational changes in the research sites involved, primarily managerial and support services initiated by policy decisions which were made in the sites.[181]

The literature tends to address strategies to achieve changes in individual practitioner behaviour. The strategies include the following: feedback; continuing education; implementation of guidelines; and the use of quality improvement techniques. Systematic reviews tend to group all or some of these categories under feedback or continuing education without discriminating among factors which may have affected the research findings, ie, nature of the design of informational or educational interventions, use of a single strategy or a combination of strategies on a one-time or continuous basis, or perceived authority of change mechanism or agent.

Feedback. A review of 36 studies of the role of feedback of statistical information in changing clinical practice included other interventions designed to influence clinical care as well.[182] The authors concluded that "feedback of information most probably influences clinical practice if it is part of an overall strategy which targets decision makers who have already agreed to review their practice and if it is presented close to the time of decision making."

Continuing Education. Comprehensive systematic reviews of 99 RCTs on continuing medical education have used a very broad definition of education to include trials involving opinion leaders, patient-mediated strategies, feedback, and reminders as strategies designed to change the performance of doctors and healthcare outcomes.[183-184] Others have distinguished among these interventions using categories such as (1) education, (2) feedback, (3) administrative changes, (4) group processes, and (5) incentives.[168] The findings of a recent systematic review on CME included the following:[184]

- In the RCTs reviewed, 81 single intervention strategies were used, of which 60 percent demonstrated change in at least one major outcome measure, 37 percent failed to demonstrate change, and 3 percent effected change in doctor performance without altering healthcare outcomes.
- Relatively short formal continuing medical education events such as conferences generally effected no change.
- Of the 39 interventions which used two educational methods, 64 percent were positive, 31 percent were negative or inconclusive, and 3 yielded mixed results.
- Of 39 interventions which used three or more educational strategies, 79 percent were positive, 13 percent were negative, and 8 percent produced mixed results.
- Of 28 interventions targeted at specific behaviours as a result of a "gap analysis technique", such as audit, or addressed specific barriers to change, 89 percent resulted in a positive change and 11 did not.

Consistent with RCT findings, some researchers have reported success with "educational marketing" campaign approaches involving a combination of feedback and educational interventions over time.[185-186] The mechanisms accounting for the success of such campaigns may be explained via a small-scale qualitative study involving general practices, which was carried out by Armstrong, Reyburn, and Jones.[187] They identified three models which accounted for GPs' changes in prescribing practices: (1) accumulation, in which the volume and authority of evidence were important; (2) challenge, in which change followed a dramatic or conflictual clinical event; and (3) continuity, in which change took place against a background of willingness to change, influenced by other factors.

Guidelines. Two UK reviews on studies involving the implementation of guidelines have been published; both reviews include audit studies which used criteria as well as studies using guidelines.[188-189] It is unclear, therefore, if conclusions about practitioners' changes in practice differ if they are engaged in "audit" as opposed to following "guidelines." Four major differences between guidelines and criteria have been identified:[107] (1) Guidelines are designed to assist clinicians and patients in making decisions about healthcare to be given whereas criteria assess care decisions that have already been made; (2) Guidelines recommend patient care data that should be obtained as they are needed to make clinical decisions whereas criteria are applied to data that were collected in the course of delivering care; (3) Guidelines may follow a set of branching pathways corresponding to variation in the health states of patients whereas criteria are generally limited to main important pathways; and (4) Guidelines tend not to include all possible clinically acceptable exceptions whereas criteria tend to acknowledge frequently occurring anticipated exceptions.

The main conclusions of the systematic reviews on guidelines have included the following:[188-189]

- The more overtly educational the strategy for disseminating guidelines, the greater the likelihood that the guidelines will be adopted within clinical practice, provided that dissemination of guidelines is reinforced by an appropriate implementation strategy.

- Of the implementation strategies which operate within the doctor-patient consultation, those which focus on the management of individual patients are more likely to lead to changes in practice.

- Internally-generated guidelines are less likely to be scientifically valid because local groups lack the resources, including skills, needed to develop guidelines.

Quality Improvement (QI) Techniques. The contribution of quality improvement techniques as a form of intervention to achieve change has been recognised in recent studies which combined feedback and the explicit use by doctors of quality improvement techniques.[190-191] As the use of these QI techniques gains more acceptance by doctors and

other healthcare professionals, their utility in facilitating changes in practice is likely to be recognised.

Strategies for Achieving Change. Practitioners frequently recognise the need to "complete the audit cycle" but they may not have the skills, time, patience, authority, or resources needed to engineer significant improvements in the quality of patient care. Baker noted that audits fail when the findings do not bring about changes in day-to-day clinical practice, because deficiencies in patient care are seen as unimportant or because the changes required seem so considerable at a local level so as to be impracticable.[192] O'Hagan provided one example of why "the effector part of the peer review cycle can be the most difficult and time consuming," explaining how it took two years in his hospital to achieve improvements in the management of patients with asthma.[193]

The findings of studies on attempts to change practitioner behaviour should not be surprising and "it is unrealistic to expect a single, 'magic bullet' solution . . . given the number and diversity of barriers to the utilisation of health care evidence."[194] Haynes has also asserted that traditional means of transmitting information about advances in knowledge about effective medical care, ie, professional journals, can serve to confuse practitioners because journals fill a number of communication functions. They publish rigorous studies which scientists are communicating to practitioners, but they also publish short reports, opinions, and communications among practitioners.[195]

Mittman, Tonesk, and Jacobson have argued that decisions, actions, and behaviours, including of healthcare practitioners, are based less on formal evaluations of cost-benefit and instead on habits and customs; assumptions, beliefs, and values held by peers; and on prevailing practices and social norms that define appropriate behaviour.[196] They advocated the use of models of practitioner behaviour change which consider the "social influence" of peers' judgements and beliefs. The implications of this theoretical approach include that transferring information via feedback and educational interventions is unlikely to be as effective as transferring norms and values. The specific strategies suggested for modifying group norms include the use of single individual opinion leaders to influence all members of a group or the facilitated introduction to a group of quality improvement techniques or other participatory decision-making activities, such as the development of clinical practice guidelines, through which new norms can be established within a group.

Maintaining Improvements. In addition to the trials included in reviews on feedback and education, the particular problem has been recognised of maintaining improvements in practice when doctors in training, whose behaviour is being modified via audit, rotate frequently to new jobs as in the NHS.[197-199] In these circumstances, the important role of senior consultant staff's training and supervision of doctors trainees has been noted.

Key Point

Changes in practice may be achieved more effectively if a strategy involving several different types of action is used.

PUTTING THE EVIDENCE IN CONTEXT

Audit as an Innovation

Clinical audit is a mandated innovation in national healthcare systems, and it may be useful to consider audit in this context. Whereas the principles of audit of medical care have developed over a century and individual practitioners voluntarily carried out audit activities over many years, the large-scale implementation of audit in all healthcare organisations has represented an unprecedented requirement for "quality control" of healthcare services by the governments of certain countries including the UK.

Rogers[200] has defined an innovation as an idea, practice, or object that is *perceived as new* by an individual or organisation. "Newness" involves not only knowledge of the innovation but also the attitude toward it which may affect the decision to adopt. Rogers identified five characteristics of innovations which help to explain their rate of adoption: relative advantage, ie, the degree to which an innovation is perceived as better then the idea it supersedes; compatibility, ie, the degree to which an innovation is perceived as being consistent with existing values, past experiences, and needs of potential adopters; complexity, ie, the degree to which an innovation is perceived as difficult to understand and use; trialability, ie, the degree to which an innovation may be experimented with on a limited basis; and observability, ie, the degree to which the results of an innovation are visible to others.

Rogers' characteristics provide a framework for positioning the evidence on clinical audit as follows:

- Relative advantage — For some groups, the advantages of audit have been obvious from the start of using the process. Through clinical audit, groups have been enabled to achieve changes which in turn benefit patients. However, in the NHS, it is unclear if clinical audit supersedes another formal process or activity with the same objective, ie, to systematically measure and improve the quality of patient care. There may be little basis for practitioners to recognise that audit offers advantages over more traditional methods such as self-assessment of one's own clinical practice, grand rounds, educational sessions, standard-setting, or other activities which audit is perceived as replacing. The efficacy of previous peer-accepted methods in systematically improving patient care has not been tested in controlled research contexts.

Clinical audit is an example of what Rogers called a "preventive innovation," ie, one in which the sought-after consequences are distant in time and so the realisation of benefits is delayed. An analogous example in medicine is health promotion or preventive medicine practices. It is sometimes difficult to motivate people to prevent unwanted events which may occur naturally at low incidence rates. Similarly, it may not motivate practitioners to systematically look for circumstances or situations which suggest that the quality of patient care can be improved.

Relative advantages of adoption of the clinical audit process could include the potential for decreasing waste or costs, inconvenience or discomfort to patients, or discordance with professionals' values when services are believed to be inconsistent with patients' needs. However, these potential benefits require commitment to the audit process before they can be realised.

- Compatibility — In the NHS context, there are two potential incompatibilities between the implementation of the clinical audit process and practitioners' values and beliefs: clinicians' autonomy,[201] and practitioners' beliefs and past experiences that the implementation of improvements in patient care will be frustrated by the lack of financial resources or by other constraints such as employees' working practices.

  The clinical audit process introduces new forms of behaviour in some professional or specialty groups, ie, improving *group* as opposed to *individual* performance. Unless a professional or specialty group already has the experience of working as a team or is facilitated to learn how, the audit process may remain incompatible with the group's norms. Also, over the more than forty years during which the NHS was managed as a monolithic bureaucracy, a large number of individuals who may have attempted to achieve patient care improvements based on valid evidence were thwarted by "the system". Thus the objective of the audit process, ie, to improve the quality of patient care, is still met with cynicism by some practitioners.

  Negative experiences with one innovation can affect the adoption of future related innovations. A negativity toward clinical audit may be predisposed in the NHS by previous approaches to "quality improvement," eg, quality assurance or total quality management.

- Complexity — The principles of clinical audit are logical and self-evident. The research-based evidence on audit, however, illustrates the true complexity of the process. After decades of research, some key points about measuring the quality of patient care are understood in a scientific sense. Strategies and techniques for changing practitioner behaviour are now being subjected to intensive inquiry. The clinical audit process has appeared deceptive to some: incredibly easy and straightforward on the face of it yet incredibly difficult and complex to get to "work" each and every time.

- Trialability — The clinical audit process can be tried out very easily by individuals or groups. It is easy for practitioners to find out how audit works in their own circumstances. The value of audits carried out on a "trial" basis has helped considerably to overcome practitioners' uncertainties about the process and to promote the continuation of the activity.

- Observability — Practitioners have been willing to share their positive experiences with audit which has enabled the benefits of the process to be visible to others.

In summary, the clinical audit process is easy to try out and it is observable to others. The process is logical and easily understood, although there are a number of complex technical issues in implementation. There are defined advantages of the process but these are not always apparent to clinicians. Implementation of the process may not be compatible with the social system of healthcare practitioners, as illustrated by the research evidence on changing practitioners' behaviour. This analysis highlights potential barriers to clinical audit and related activities which were classified by Luke and Boss as follows: lack of clarity about what quality in healthcare means; lack of certainty about valid and reliable methods for measuring quality; and lack of certainty about methods for improving quality.[202]

Audit as Part of the Paradigm Shift

There has been considerable debate about the nature of quality in healthcare.[203] Progress is now being made in the UK in developing an understanding of what quality in healthcare means, at least from the patient's perspective.[204] Key attributes of service including appropriateness, effectiveness, and accessibility of services as well as acceptability and efficacy of services are being emphasised.

The lessons learned so far about defining, measuring, and improving the quality of patient care are contributing to a "paradigm shift" in medicine. Kuhn observed that science doesn't exist as a unified body of theory, knowledge, and technique but that there are multiple sciences in existence at any one time. At times of intellectual crises, alternatives are brought forward, new disciplines are established, and a shift in paradigm occurs in which new methods and techniques are applied to perceived problems.[205]

In the Western world, healthcare systems are in crisis caused by an inability of nations to match the demand for and the supply of healthcare services. In response, healthcare systems are experiencing a "Kuhnian revolution"[206] consisting of two parts: a systematic examination of the scientific basis of clinical practice,[206] and application of industrial techniques to achieve improvements in the processes and outcomes of clinical practice.[207]

The lack of scientific basis for clinical practice is being addressed internationally through such initiatives as the Cochrane Collaboration on Effective Professional Practice, consensus development of clinical practice guidelines, and technology assessment and research and development programmes.

The second part of the paradigm shift recognises that, in addition to widespread variation in individual practitioner behaviours, day-to-day operational systems used in healthcare settings are considerably outmoded when compared with those used in industry and they tend to be designed to meet practitioners', not patients', needs. Shortcomings in the provision of patient care services at local level are being addressed by use of a variety of quality improvement techniques, including clinical audit.

The Implications for the Professions

Wennberg has commented on the profound implications of the healthcare crisis for the status of the medical profession. "The doctor-patient relationship is based on the notion that it is rational for patients to delegate decision making to physicians ... because of their formal training, continuing education, and extensive experience, [physicians] are assumed to know the scientifically correct way to treat disease. Moreover, they are assumed to understand vicariously the needs and values of patients and thus are qualified to make utility or value judgements for patients .... It is no longer reasonable or feasible to base health policy on ... [this] theory ... [based on] the evidence..."[206]

Berwick has argued that healthcare professionals and purchasers must establish and hold to a shared vision of a healthcare system which is undergoing continuous improvement[208] and use the scientific method to understand how systems work and to achieve improvements.[209]

Berwick, Enthoven, and Bunker have identified a set of new skills which practitioners need to take part in the paradigm shift in healthcare including the following:[210-211]

- The ability to perceive and work effectively in interdependencies.
- The ability to work in teams.
- The ability to understand work as process.
- Skill in collecting, aggregating, analysing, and displaying data on outcomes of care.
- Skills in "designing" healthcare practices.
- Skill in collecting, aggregating, analysing, and displaying data on processes of work.
- Skills in collaborative exchange with patients.
- Skills in working collaboratively with managers.

Active involvement in clinical audit and other quality improvement techniques offer opportunities to practitioners to hone these skills and to make a contribution to improving care to patients at a time when national resources impose constraints upon healthcare systems.

CONCLUSION

The literature pertinent to clinical audit supports the overall efficacy of the process. Sufficient valid evidence is available to identify criteria for good practice in clinical audit, as indicated in this review.

Nonetheless, the clinical audit process continues to be subject to question at local level, because it has appeared to some that the goals of audit have been unclear; that audits which have shown the need for additional facilities have been ignored by managers and purchasers; and that doctors have not been convinced that audit improves quality.[13] Therein lie the challenges, because it is at local level that acceptance of the value of clinical audit must occur.

In a climate in which the clinical effectiveness of so much of what practitioners are trained to do is being subject to question, it is easy to see why clinicians would prefer to defer involvement in clinical audit until the evidence is overwhelmingly conclusive about its value and how to use it. Regrettably, such a view would mean that members of the public could be denied realisation of some benefits of audit, even from imprecise and perhaps ad hoc application of the process as it may be practised in many healthcare organisations. Also, such a position would deny practitioners the opportunities to develop needed new skills and would leave to others in and outside the healthcare system to monitor the quality of patient care.

A more responsible approach may be to acknowledge the improvements which can be made in the quality of care provided to patients in the NHS and to explore systematically the mechanisms which help to achieve continuous improvement. This review and other publications have identified a research agenda for clinical audit and quality improvement. As this agenda is being pursued over time, there are benefits to sharing further information on clinical audit, including the experiences of practitioners who are implementing audit at local level. The National Centre for Clinical Audit offers the facility of gathering and disseminating information about clinical audit in order to contribute to the understanding of how improvements which benefit patients can be achieved.

GLOSSARY OF TERMS

The US Agency for Health Care Policy and Research (AHCPR) and the Institute of Medicine (IOM) have introduced precision in the use of terms pertaining to audit and guidelines. The AHCPR and IOM definitions are provided in the table.[212-213] The definitions have been explained and interpreted for use in the UK.[214]

Term	Meaning
Criteria	Systematically developed statements that can be used to assess specific health care decisions, services, and outcomes.
Explicit criteria	Objective criteria specified in advance as a basis for making judgements of performance.
Implicit criteria	Criteria formed by a respected clinician who uses clinical judgement in evaluating performance; these implicit criteria remain concealed in the mind of the reviewer.
Indicators	Quantitative measures used to measure and improve performance and quality. Indicators can be rate-based or event-based.
Standards	Authoritative statements of (1) minimum levels of acceptable performance or results, (2) excellent levels of performance or results, or (3) the range of acceptable performance or results.
Performance measures	Methods or instruments to estimate or monitor the extent to which the actions of a health care practitioner or provider conform to a clinical practice guideline
Performance rates	Measurements produced by using a performance measure, providing a quantitative evaluation of events related to patient care.
Clinical practice guidelines	Systematically developed statements to assist practitioners' and patients' decisions about health care to be provided for specific clinical circumstances.

The Royal College of Nursing uses the definition given of clinical practice guidelines. The RCN definitions of other key terms are as follows:[215]

Term	RCN Meaning
Criteria	Variables which enable the achievement of a standard and evaluation of whether it has been achieved or not.
Standard	A statement which outlines an objective, with guidance on how to achieve it including required resources, activities, and predicted outcomes.
Indicator	A measurement tool used to evaluate important aspects of healthcare and direct attention to any problems.

REFERENCES

1. *The Evolution of Clinical Audit.* London: Department of Health, 1994.
2. Barton A, Thomson R, Bhopal R. Clinical audit: more research is required. *J Epidemiol Community Health 1995*;45:445-7.
3. Buttery Y, Walshe K, Coles J, Bennett J. *The Development of Audit. Findings of a National Survey of Healthcare Provider Units in England.* London: CASPE Research, 1994.
4. Buttery Y, Walshe K, Rumsey M, Amess M, Bennett J, Coles J. *Provider Audit in England. A Review of Twenty-Nine Programmes.* London: CASPE Research, 1995.
5. McIntyre N. Evaluation in clinical practice: problems, precedents and principles. *Journal of Evaluation in Clinical Practice* 1995;1:5-13.
6. Miles A, Bentley P, Polychronis A, Price N, Grey J. Clinical audit in the National Health Service: fact or fiction? *Journal of Evaluation in Clinical Practice* 1995;2(1):29-35.
7. National Audit Office. *Clinical Audit in England.* London: HMSO, 1995.
8. Robinson MB. Evaluation of medical audit. *J Epidemiol Community Health* 1994;48:435-40.
9. Rumsey M, Walshe K, Bennett J, Coles J. *The Role of the Commissioner in Audit. Findings of a National Survey of Commissioning Authorities in England.* London: CASPE Research, 1994.
10. Thomson R, Barton AG. Is audit running out of steam? *Quality in Health Care* 1994;3:225-9.
11. Walshe K. The traits of success in clinical audit. In Walshe K (ed). *Evaluating Clinical Audit: Past Lessons, Future Directions.* London: Royal Society of Medicine Press, 1995, 13-21.
12. Walshe K, Coles J. *Evaluating Audit: A Review of Initiatives.* London: CASPE Research, 1993.
13. Sellu D. Time to audit audit. *Br Med J* 1996;312:128-9.
14. Fulton RA. Goals and methods of audit should be reappraised (letter). *Br Med J* 1996;312:1103.
15. Baker R, Fraser RC. Development of review criteria: linking guidelines and assessment of quality. *Br Med J* 1995;311:370-3.
16. Collopy BT. Criteria auditing: an outline of the system and a commentary upon its use. *Aust Clin Rev* 1981;1:15-7.
17. Crombie IK, Davies HT. Towards good audit. *Br J Hosp Med* 1992;48(3-4):182-5.
18. de Lacey G. 1. What is audit? Why should we be doing it? *Hosp Update* 1992;18(6):458-66.
19. Dixon N. Practical principles of medical audit. *Postgrad Med J* 1990;66(Suppl 3):517-20.
20. Kitson A, Kendall H. Rest assured. *Nursing Times* 1986;82(35):29-31.
21. Malby R. The process of change in nursing audit. *British Journal of Nursing* 1992;1(4):205-7.
22. Morrell C. Developing criteria for evaluating audit: a brief overview of the literature. London: National Centre for Clinical Audit, 1995. 8pp. Unpublished.
23. Shaw CD. Acceptability of audit. *Br Med J* 1980;290:1443-6.
24. Shaw CD. Looking forward to audit. *Br Med J* 1980;290:1509-11.
25. Shaw CD, Costain DW. Guidelines for medical audit: seven principles. *Br Med J* 1989;299:498-9.
26. Shaw CD. Criterion based audit. *Br Med J* 1990;300:694-51.
27. Whelan J. Troublespots in quality assurance. *Nursing Standard* 1991;5(25):36-8.
28. Russell IT, Wilson BJ. Audit: the third science? *Quality in Health Care* 1992;1:51-5.
29. Bhopal RS, Thomson R. A form to help learn and teach about assessing medical audit papers. *Br Med J* 1991;303:1520-2.
30. Lough JRM, McKay J, Murray TS. Audit and summative assessment: a criterion-referenced marking schedule. *Br J Gen Pract* 1995;45(400):607-9.
31. Walshe K, Coles J. *Evaluating Audit: Developing a Framework.* London: CASPE Research, 1993.

32. Walshe K, Buttery Y. Measuring the impact of audit and quality improvement programmes. *JAQH* 1995;2(4):138-47.
33. Derry J, Lawrence M, Griew KA, Anderson J, Humphreys J, Pandhur KS. Auditing audits: The method of the Oxfordshire Medical Audit Advisory Group. *Br Med J* 1991;303:1247-9.
34. Lawrence M, Griew K, Derry J, Anderson J, Humphreys J. Auditing audits: use and development of the Oxfordshire Medical Audit Advisory Group rating system. *Br Med J* 1994;309:513-6.
35. *The Development of an Audit Review Tool. Research Report No. 5. Report of the Joint Audit Review Group.* Leicester: Eli Lilly National Clinical Audit Centre, 1995.
36. Fowkes FGR. Medical audit cycle. A review of methods and research in clinical practice. *Med Educ* 1982;16:228-38.
37. Palmer RH, Nesson HR. A review of methods for ambulatory medical care evaluation. *Med Care* 1982;20(8):758-81.
38. Balogh R. Audits of nursing care in Britain: a review and a critique of approaches to validating them. *Int J Nurs Stud* 1992;29(2):119-33.
39. Kitson A. Quality assurance in nursing practice. *Proceedings of the Royal Society of Edinburgh* 1993;101B:143-63.
40. Lang NM, Clinton JF. Assessment of quality of nursing care. *Annu Rev Nurs Res* 1984;2:135-63.
41. McGinnis S. History of nursing audit. A survey of the literature. *QRB* 1975;1(5-7):10-2.
42. Wright D. An introduction to the evaluation of nursing care: a review of the literature. *J Adv Nurs* 1984;9:457-67.
43. Putnam RW, Curry L. Impact of patient care appraisal on physician behaviour in the office setting. *Can Med Assoc J* 1985;132:1025-9.
44. Palmer RH. Does quality assurance improve ambulatory care? Implementing a randomized controlled trial in 16 group practices. *J Ambulatory Care Manage* 1986;9(4):1-15.
45. Palmer HR, Louis TA, Hsu LN, Peterson HF, Rothrock JK, Strain R, *et al.* A randomized controlled trial of quality assurance in sixteen ambulatory care practices. *Med Care* 1985;23(6):751-70.
46. Palmer RH, Strain R, Maurer JVW, Rothrock JK, Thompson MS. Quality assurance in eight adult medical group practices. *Med Care* 1984;22:632-43.
47. Palmer RH, Strain R, Maurer JVW, Thompson MS. A method for evaluating performance of ambulatory pediatric tasks. *Pediatrics* 1984;73:269-77.
48. Anderson FA, Wheeler HB, Goldberg RJ, Hosmer DW, Forcier A, Patwardhan NA. Changing clinical practice. Prospective study of the impact of continuing medical education and quality assurance programs on use of prophylaxis for venous thromboembolism. *Arch Intern Med* 1994;154:669-77.
49. Lomas J, Enkin M, Anderson GM, Hannah WJ, Vayda E, Singer J. Opinion leaders vs audit and feedback to implement practice guidelines. *JAMA* 1991;265(17):2202-7.
50. Mayefsky JH, Foye HR. Use of chart audit: teaching well child care to paediatric house officers. *Med Educ* 1993;27:170-4.
51. Winickoff RN, Coltin KL, Morgan MM, Buxbaum RC, Barnett GO. Improving physician performance through peer comparison feedback. *Med Care* 1984;22(6):527-34.
52. Winickoff RN, Wilner S, Neisuler R, Barnett GO. Limitations of provider interventions in hypertension quality assurance. *Am J Public Health* 1985;75(1):43-6.
53. Nightingale F. *Notes on Nursing: Enlarged and For the Most Part Rewritten.* 3rd ed. London: Longman, Green, Longman, Roberts and Green, 1863.
54. Nightingale F. Mortality of the British army at home and abroad and during the Russian war as compared with the mortality of the civil population in England. Reprinted from: *Report of the Commission Appointed to Inquire Into the Regulations Affecting the Sanitary State of the Army.* London: Harrison and Sons, 1858.
55. Devlin HB. Audit and the quality of clinical care. *Ann R Coll Surg Engl Suppl* 1990;72:3-14.
56. Groves EWH. Surgical statistics: a plea for a uniform registration of operation results. *Br Med J* 1908;2:1008-9.
57. Codman EA. The product of a hospital. *Surgery, Gynecology and Obstetrics* 1914;18:491-6.

58. Codman EA. *A Study in Hospital Efficiency as Demonstrated by the Case Reports of the First Five Years in a Private Hospital.* Boston: 1916,137, reproduced in Ann Arbor, Michigan: University Microfilms, 1972.
59. Lembcke PA. Medical auditing by scientific methods. Illustrated by major female pelvic surgery. *JAMA* 1956;162(7):646-55.
60. American College of Surgeons. The Minimum Standard of the American College of Surgeons' Hospital Standardization Program. *Bull Am Coll Surgeons* 1924;8:4.
61. Lembcke PA. Evolution of the medical audit. *JAMA* 1967;199(8):111-8.
62. Jacobs CM, Christoffel TH, Dixon N. *Measuring the Quality of Patient Care: The Rationale for Outcome Audit.* Cambridge MA: Ballinger Publishing Company, 1976.
63. Brown CR, Fleisher DS. The bi-cycle concept – relating continuing education directly to patient care. *N Engl J Med* 1971;284:88-97.
64. Phaneuf MC. *The Nursing Audit: Profile for Excellence.* New York: Appleton-Century-Crofts, 1972.
65. Wandelt MA, Ager JW. *Quality Patient Care Scale.* New York: Appleton-Century-Crofts, 1974.
66. Hegyvary ST, Haussman RKD. Monitoring nursing care quality. *J Nurs Adm* 1975;5(5):17-26.
67. Hegyvary ST, Haussman RKD. Nursing professional review. *J Nurs Adm* 1976;6(9):12-6.
68. Hegyvary ST, Haussman RKD. The relationship of nursing process and patient outcomes. *J Nurs Adm* 1976;6(9):18-21.
69. Goldstone LA, Ball JA, Collier MM. *Monitor: An Index of the Quality of Nursing Care for Acute Medical and Surgical Wards.* Newcastle upon Tyne: Newcastle upon Tyne Polytechnic Products, 1983.
70. Fleishman SJ, Nesson HR. Point of view: on the road to ambulatory quality assurance. *J Amb Care Management* 1978;1:91.
71. Winickoff RN, Barnett GO, Morgan M, Coltin KL. Quality assurance in a prepaid group practice. *J Amb Care Management* 1979;2:19.
72. Batalden PB, O'Connor JP. *Quality Assurance in Ambulatory Care.* Germantown MD: Aspen Systems Corporation, 1980.
73. Collopy BT, Hart JA, Hooper JC, Mills KW, Rush JH, Russell IS, *et al.* An inter-hospital criteria audit of infection in total hip-joint replacement surgery. *Aust Clin Rev* 1984;13:19-21.
74. Collopy B, May J, Morgan B, Torr S. The results of an audit of primary appendicectomy in two Australian States. *Aust Clin Rev* 1983;10:6-10.
75. Collopy BT, Russell I, Scott D, Browne J. Criteria audit of surgery for carcinoma of the large intestine in Melbourne teaching hospitals. *Med J Aust* 1981;2:241-5.
76. Suñol R, Delgado R, Esteban A. Medical audit: the Spanish experience. *Br Med J* 1991;303:1249-51.
77. Tugwell P, Manganelli E. The clinical audit cycle. *Aust Clin Rev* 1986;6(21):100-6.
78. Greenfield S, Lewis CE, Kaplan SH, Davidson MB. Peer review by criteria mapping: criteria for diabetes mellitus. The use of decision-making in chart audit. *Ann Intern Med* 1975;83:761-70.
79. Gonnella JS, Goran MJ. Quality of patient care – a measurement of change: the staging concept. *Med Care* 1975;13(6):467-73.
80. Kessner DM, Kalk CE, Singer J. Assessing health quality – the case for tracers. *N Eng J Med* 1973;288(4):189-94.
81. Rhee KJ, Donabedian A, Burney RE. Assessing the quality of care in a hospital emergency unit: a framework and its application. *QRB* 1987;13:4-16.
82. Joint Commission sets agenda for change. *QRB* 1987;13:148-50.
83. Bernstein SJ, Hilborne LH. Clinical indicators: the road to quality care? *Jt Comm J Qual Improv* 1993;19:501-9.
84. Lehmann R. Forum on clinical indicator development: a discussion of the use and development of indicators. *QRB* 1989;15:223-7.
85. Marder RJ. Relationship of clinical indicators and practice guidelines. *QRB* 1990;16:60.
86. Nadzam DM. Data-driven performance improvement in health care: the Joint Commission's indicator measurement system (IM system). *Jt Comm J Qual Improv* 1993;19:492-500.
87. Nadzam DM. The agenda for change: update on indicator development and possible implications for the nursing profession. *J Nurs Qual Assur* 1991;5(2):18-22.

88. *Primer on Indicator Development and Application. Measuring Quality in Health Care.* Oakbrook Terrace IL: Joint Commission on Accreditation of Healthcare Organizations, 1990.
89. Ansari MZ, Collopy BT. Assessment of postoperative pulmonary embolism as a measure of patient care. *J Qual Clin Pract* 1995;15:75-80.
90. Collopy BT. Audit activities in Australia. *Br Med J* 1991;303:1523-5.
91. Collopy BT. Developing clinical indicators: The ACHS Care Evaluation Program. *Aust Clin Rev* 1990;10:83-5.
92. Collopy BT. Improving appropriate use of surgical services. *Quality Assurance in Health Care* 1991;3(4):221-5.
93. Collopy BT, Balding C. The Australian development of national quality indicators in health care. *Jt Comm J Qual Improv* 1993;19:510-6.
94. Ewing HP, Cade RJ, Cocks JR, Collopy BT, Thompson GA. Developing clinical indicators for cholecystectomy. *Aust N Z J Surg* 1993;63:181-5.
95. *Getting Ahead With Clinical Audit. A Facilitator's Guide.* Bristol: NHS Training Directorate, 1994.
96. Kazandjian VA, Lawthers J, Cernak CM, Pipesh FC. Relating outcomes to processes of care: the Maryland Hospital Association's Quality Indicator Project (QI Project®). *Jt Comm J Qual Improv* 1993;19(11):530-38.
97. Kazandjian VA, Wood P, Lawthers J. Balancing science and practice in indicator development: the Maryland Hospital Association Quality Indicator (QI Project®). *Int J Qual Health Care* 1995;7(1):39-46.
98. Lynch JT, Mattie AS, Shevchenko IP, Reed-Fourquet LL. The "Toward Excellence in Care" Program: a statewide indicator project. *Jt Comm J Qual Improv* 1993;19(11):519-29.
99. Godwin R, de Lacey G, Manhire A, eds. *Clinical Audit in Radiology. 100+ Recipes.* London: The Royal College of Radiologists, 1996.
100. Bloomberg MA, Jordan HS, Angel KO, Bailit MH, Goonan KJ, Straus J. Development of clinical indicators for performance measurement and improvement: an HMO/purchaser collaborative effort. *Jt Comm J Qual Improv* 1993;19(12):586-95.
101. Garnick DW, Lawthers AG, Palmer RH, Moentmann SJR, Fowles J, Weiner JP. A computerized system for reviewing medical records from physicians' offices. *Jt Comm J Qual Improv* 1994;20(12):679-94.
102. Hargraves JL, Palmer RH, Zapka J, Nerenz D, Frazier H, Orav J, *et al.* Using patient reports to measure health care system performance. *Clinical Performance and Quality Health Care* 1993;1(4):208-13.
103. Lawthers AG, Palmer RH, Banks N, Garnick DW, Fowles J, Weiner JP. Designing and using measures of quality based on physician office records. *J Ambulatory Care Manage* 1995;56-72.
104. Lawthers AG, Palmer RH, Edwards JE, Fowles J, Garnick DW, Weiner JP. Developing and evaluating performance measures for ambulatory care quality: a preliminary report of the DEMPAQ Project. *Jt Comm J Qual Improv* 1993;19(12):552-65.
105. Nerenz DR, Zajac BM, Rosman HS. Consortium Research on Indicators of System Performance (CRISP). *Jt Comm J Qual Improv* 1993;19(12):577-85.
106. *The Measurement Mandate. On the Road to Performance Improvement in Health Care.* Oakbrook Terrace IL: Joint Commission on Accreditation of Healthcare Organizations, 1993.
107. *Using Clinical Practice Guidelines to Evaluate Quality of Care. Vol 2. Methods.* Rockville MD: Agency for Health Care Policy and Research, 1995.
108. Secretaries for State for Health, Wales, Northern Ireland, and Scotland. *Working for Patients: Medical Audit.* Working Paper 6. London: HMSO, 1989.
109. Royal College of Physicians Working Party on Medical Audit. *Medical Audit. A First Report.* London: Royal College of Physicians, 1989.
110. Anderson ID, Woodford M, de Dombal FT, Irving M. Retrospective study of 1,000 deaths from injury in England and Wales. *Br Med J* 1988;296:1305-8.
111. Buck N, Devlin HB, Lunn JN. *The Report of a Confidential Enquiry into Perioperative Deaths.* London: Nuffield Provincial Hospitals Trust and the King's Fund, 1987.

112. Dunn DC. Audit of a surgical firm by microcomputer: five years' experience. *Br Med J* 1988;296:687-91.
113. Ellis BW, Michie HR, Esufali ST, Pyper RJD, Dudley HAF. Development of a microcomputer based system for surgical audit and patient administration: a review. *Roy Soc Med* 1987;80:157-61.
114. Fielding LP, Stewart-Brown S, Blesovsky L, Kearney G. Anastomatic integrity after operations for large bowel cancer: a multicentre study. *Br Med J* 1980;281:411-4.
115. Fleming DM, Lawrence MSTA. Impact of audit on preventive measures. *Br Med J* 1983;287:1852-4.
116. Fowkes FGR, 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;1(8477):367-70.
117. Fowkes FGR, Evans RC, Williams LA, Gehlback SH, Cooke BRB, Roberts CJ. Implementation of guidelines for the use of skull radiographs in patients with head injuries. *Lancet* 1984;2(8406):795-6.
118. Gilmore OJA, Griffiths NJ, Connolly JC, Dunlop AW, Hart S, Thomson JP, *et al.* Surgical audit: comparison of the workload and results of two hospitals in the same district. *Br Med J* 1980;281:1050-2.
119. Gough MH, Kettlewell MGW, Marks GC, Holmes SJK, Holderness J. Audit: an annual assessment of the work and performance of a surgical firm in a regional teaching hospital. *Br Med J* 1980;281:913-8.
120. Gruer R, Gordon DS, Gunn AA, Ruckley CV. Audit of surgical audit. *Lancet* 1986;i:23-5.
121. McColl I. Medical audit in British hospital practice. *Br J Hosp Med* 1979;22(5):485-9.
122. Prout WG, Blood PA. The establishment of a microcomputer based diagnosis and operations index in a department of surgery of a district general hospital. *Br J Surg* 1985;72:48-51.
123. Rees AM, Roberts CJ, Bliss AS, Evans KT. Routine pre-operative chest radiography in non-cardiopulmonary surgery. *Br Med J* 1976;i:1333-5.
124. Roberts CJ, Fowkes FGR, Ennis WP, Mitchell M. Possible impact of audit on chest x-ray requests from surgical wards. *Lancet* 1983;2(8347):446-8.
125. Royal College of Radiologists. Pre-operative chest radiology. *Lancet* 1979;ii:83-6.
126. The Swansea Physicians' Audit Group. Audit reviewed. Implementing audit in a division of medicine. *J R Coll Physicians Lond* 1982;16(4):252-4.
127. The Swansea Physicians' Audit Group. Audit of clinical practice: opportunities to rationalise the use of resources. *J R Coll Physicians Lond* 1983;17(4):235-8.
128. Dixon N, ed. *Clinical Audit Manual.* Romsey: Healthcare Quality Quest, 1994.
129. *Standards for Nursing Services in the Wessex Region.* Winchester: Wessex Regional Health Authority, 1989.
130. Kitson A. Raising the standards. *Nursing Times* 1988;84(25):28-32.
131. Kitson AL, Hyndman SJ, Harvey G, Yerrell PH. *Quality Patient Care: The Dynamic Standard Setting System.* Harrow: Scutari, 1990.
132. *Framework of Audit of Nursing Services.* London: NHS Management Executive, 1991.
133. *Meeting and Improving Standards in Healthcare.* A Policy Statement on the Development of Clinical Audit. NHS Executive Letter (93)59, 2 July 1993.
134. *The New Health Authorities and the Clinical Audit Initiative: Outline of Planned Monitoring Arrangements.* NHS Executive Letter(95)103,4 October 1995.
135. Batstone GF. Educational aspects of medical audit. *Br Med J* 1990;301:326-8.
136. Coles C. Self assessment and medical audit: an educational approach. *Br Med J* 1989;299:807-8.
137. Williamson JW. Formulating priorities for quality assurance activity. Description of a method and its application. *JAMA* 1978;239:631-7.
138. Williamson JW, Braswell HR, Horn SD, Lohmeyer S. Priority setting in quality assurance: reliability of staff judgments in medical institutions. *Med Care* 1978;16(11):931-40.
139. Williamson JW, Braswell HR, Horn SD. Validity of medical staff judgments in establishing quality assurance priorities. *Med Care* 1979;17(4):331-46.
140. Hirschhorn N, Lamstein J, Klein S, McCormack J, Goldberg R. *Quality by Objectives: A Practical Method for Quality of Care Assessment and Assurance for Ambulatory Health Centres.* Boston: GK Hall, 1978.

141. Donabedian A. Criteria and standards for quality assessment and monitoring. *QRB* 1986;12(3):99-108.
142. Donabedian A. Criteria, norms and standards of quality: what do they mean? *Am J Public Health* 1981;71(4):409-12.
143. Donabedian A. *Explorations in Quality Assessment and Monitoring. Vol II. The Criteria and Standards of Quality.* Ann Arbor: Health Administration Press, 1982.
144. Dixon N. *Medical Audit Primer.* Romsey: Healthcare Quality Quest, 1991.
145. Richardson FM. Peer review of medical care. *Med Care* 1972;10(1):29-39.
146. Richardson FM. Methodological development of a system of medical audit. *Med Care* 1972;10(6):451-62.
147. Brook RH, Appel FA. Quality-of-care assessment: choosing a method for peer review. *New Engl J Med* 1973;288:1323-9.
148. Hulka BS, Romm FJ, Parkerson GR, Russell IT, Clapp NE, Johnson FS. Peer review in ambulatory care: use of explicit criteria and implicit judgments. *Med Care* 1979;17(3)(Suppl):1-73.
149. Hayward RA, McMahon LF, Bernard AM. Evaluating the care of general medicine in patients: how good is implicit review? *Ann Intern Med* 1993;118:550-6.
150. Rubenstein LV, Kahn KL, Reinisch EJ, Sherwood MJ, Rogers WH, Kamberg C, *et al.* Changes in quality of care for five diseases measured by implicit review, 1991 to 1986. *JAMA* 1990;264(15):1974-9.
151. Kahn KL, Rogers WH, Rubenstein LV, Sherwood MJ, Reinisch EJ, Keeler EB, *et al.* Measuring quality of care with explicit process criteria before and after implementation of the DRG-based prospective payment system. *JAMA* 1990;264(15):1969-73.
152. Keeler EB, Rubenstein LV, Kahn KL, Draper D, Harrison ER, McGinty MJ, *et al.* Hospital characteristics and quality of care. *JAMA* 1990;268(13):1709-14.
153. Dans PE, Weiner JP, Otter SE. Peer review organizations: promises and potential pitfalls. *N Engl J Med* 1985;313:1131-7.
154. Donabedian A. Advantages and limitations of explicit criteria for assessing the quality of health care. *Milbank Q* 1981;59(1):99-106.
155. Donabedian A. The quality of care. How can it be assessed? *JAMA* 1988;260(12):1743-8.
156. Osborne CE, Thompson HC. Criteria for evaluation of ambulatory child health care by chart audit: development and testing of a methodology. Final report of the Joint Committee on Quality Assurance of Ambulatory Health Care for Children and Youth. *Pediatrics* 1975;56(4)(Suppl Pt 2):625-92.
157. Norton PG, Dempsey LJ. Self-audit: its effect on quality of care. *J Fam Pract* 1985;21(4):289-91.
158. Anderson CM, Chambers S, Clamp M, Dunn IA, McGhee MF, Sumner KR, *et al.* Can audit improve patient care? Effects of studying use of digoxin in general practice. *Br Med J* 1988;297:113-4.
159. North of England Study of Standards and Performance in General Practice. Medical audit in general practice. I. Effects on doctors' clinical behaviour for common childhood conditions. *Br Med J* 1992;304:1480-4.
160. North of England Study of Standards and Performance in General Practice. II. Effects on health of patients with common childhood conditions. *Br Med J* 1992;304:1484-8.
161. Kitson A, Harvey G, Hyndman S, Yerrell P. Criteria formulation and application: an evaluative framework. *Int J Nurs Stud* 1994;31(2):155-67.
162. Ashton CM, Kuykendall DH, Johnson ML, Wun CC, Wray NP, Carr NM, et al. A method of developing and weighting explicit process of care criteria for quality assessment. *Med Care* 1994;32(8):755-70.
163. Penney GC, Glasier A, Templeton A. Agreeing criteria for audit of the management of induced abortion: an approach by national consensus survey. *Quality in Health Care* 1993;2:167-9.
164. Park RE, Fink A, Brook RH, Chassin MR, Kahn KL, Merrick NJ, *et al.* Physician ratings of appropriate indicators for six medical and surgical procedures. *Am J Public Health* 1986;76(7):766-72.
165. Leape LL, Park RE, Kahan JP, Brook RH. Group judgments of appropriateness: the effect of panel composition. *Quality Assurance in Health Care* 1992;4(2):151-9.

166. Newton J, Hutchinson A, Steen N, Russell I, Haimes E. Educational potential of medical audit: observations from a study of small groups setting standards. *Quality in Health Care* 1992;1:256-9.
167. Kitson A, Harvey G, Hyndman S, Yerrell P. A comparison of expert- and practitioner-derived criteria for post-operative pain management. *J Adv Nurs* 1993;18:218-32.
168. Berwick DM, Coltin KL. Feedback reduces test use in a health maintenance organization. *JAMA* 1986;255(11):1450-4.
169. Eisenberg JM, Williams SV, Garner L, Viale R, Smits H. Computer-based audit to detect and correct overutilization of laboratory tests. *Med Care* 1977;15(11):915-21.
170. Everett GD, deBlois S, Chang PF, Holets T. Effect of cost education, cost audits, and faculty chart review on the use of laboratory services. *Arch Intern Med* 1983;143:42-4.
171. Parrino TA. The nonvalue of retrospective peer comparison feedback in containing hospital antibiotic costs. *Am J Med* 1989;86:442-8.
172. Wones RG. Failure of low-cost audits with feedback to reduce laboratory test utilization. *Med Care* 1987;25(1):78-82.
173. Kleefield S, Churchill WW, Laffel G. Quality improvement in a hospital pharmacy department. *QRB* 1991;17:138-43.
174. Mosel D, Gift B. Collaborative benchmarking in health care. *Jt Comm J Qual Improv* 1994;20(5):239-49.
175. Lenz S, Myers S, Nordlund S, Sullivan D, Vasista V. Benchmarking: finding ways to improve. *Jt Comm J Qual Improv* 1994;20(5):250-9.
176. Donabedian A. The role of outcomes in quality assessment and assurance. *QRB* 1992;18(11):356-60.
177. Kemeny ME, Hargreaves WA, Gerbert B, Stone GC, Gullion DS. Measuring adequacy of physician performance. A preliminary comparison of four methods in ambulatory care of chronic obstructive pulmonary disease. *Med Care* 1984;22(7):620-31.
178. Brady WJ, Hissa DC, McConnell M, Wones RG. Should physicians perform their own quality assurance audits? *J Gen Intern Med* 1988;3:560-5.
179. Ashbaugh DG, McKean RS. Continuing medical education. The philosophy and use of audit. *JAMA* 1976;236(13):1485-8.
180. Royal College of Radiologists Working Party. Influence of the Royal College of Radiologists' guidelines on hospital practice: a multicentre study. *Br Med J* 1992;304:740-3.
181. Payne BC, Lyons TF, Neuhaus E, Kolton M, Dwarshius L. Method of evaluating and improving ambulatory health care. *Health Serv Res* 1984;19(2):219-45.
182. Mugford M, Banfield P, O'Hanlon M. Effects of feedback on clinical practice: a review. *Br Med J* 1991;303:398-402.
183. Davis DA, Thomson MA, Oxman AD, Haynes RB. Evidence for the effectiveness of CME. A review of 50 randomized controlled trials. *JAMA* 1992;268(9):1111-7.
184. Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing physician performance. A systematic review of the effect of continuing medical education strategies. *JAMA* 1995;274(9):700-5.
185. Landgren FT, Harvey KJ, Mashford ML, Moulds RFW, Guthrie B, Hemming M. Changing antibiotic prescribing by educational marketing. *Med J Aust* 1988;149:595-9.
186. Thompson RS, Kirz HL, Gold RA. Changes in physician behavior and cost savings associated with organizational recommendations on the use of "routine" chest xrays and multichannel blood tests. *Prev Med* 1983;12:385-96.
187. Armstrong D, Reyburn H, Jones R. A study of general practitioners' reasons for changing their prescribing behaviour. *Br Med J* 1996;312:949-52.
188. Grimshaw J, Freemantle N, Wallace S, Russell I, Hurwitz B, Watt I, *et al.* Developing and implementing clinical practice guidelines. *Quality in Health Care* 1995;4:55-64.
189. Grimshaw JM, Russell IT. Achieving health gain through clinical guidelines. II: Ensuring guidelines change medical practice. *Quality in Health Care* 1994;3:45-52.
190. Amadio PC, Naessons JM, Rice RL, Ilstrup DM, Evans RW, Morrey BF. Effect of feedback on resource use and morbidity in hip and knee arthroplasty in an integrated group practice setting. *Mayo Clin Proc* 1996;71:127-33.

191. O'Connor GT, Plume SK, Olmstead EM, Morton JR, Maloney CT, Nugent WC, *et al.* A regional intervention to improve the hospital mortality associated with coronary artery bypass graft surgery. *JAMA* 1996;275(11):841-6.
192. Baker R. Problem solving with audit in general practice. *Br Med J* 1990;300:378-80.
193. O'Hagan JJ. Completing the peer review circle. *Aust Clin Rev* 1986;17:197-200.
194. Lomas J, Haynes RB. A taxonomy and critical review of tested strategies for the application of clinical practice recommendations: from "official" to "individual" clinical policy. *Am J Prev Med* 1987;4(Suppl):77-94.
195. Haynes RB. Loose connections between peer-reviewed clinical journals and clinical practice. *Ann Intern Med* 1990;113:724-8.
196. Mittman BS, Tonesk X, Jacobson PD. Implementing clinical practice guidelines: social influence strategies and practitioner behavior change. *QRB* 1992;18(12):413-22.
197. Bareford D, Hayling A. Inappropriate use of laboratory services: long term combined approach to modify request patterns. *Br Med J* 1990;301:1305-7.
198. Clarke JA, Adams JE. The application of clinical guidelines for skull radiography in the accident and emergency department: theory and practice. *Clin Radiol* 1990;41:152-5.
199. Gama R, Nightingale PG, Ratcliffe JG. Effect of educational feedback on clinicians' requesting of cardiac enzymes. *Ann Clin Biochem* 1992;29:224-5.
200. Rogers EM. *Diffusion of Innovations.* 4th Ed. New York: The Free Press, 1995.
201. Pollitt C. The politics of medical quality: auditing doctors in the UK and the US. *Health Services Management Research* 1993;6(1):24-34.
202. Luke R, Boss R. Barriers limiting the implementation of quality assurance programs. *Health Serv Res* 1981;16:305-14.
203. Dixon N. The Challenge of Defining Quality in Healthcare. Paper presented at Partnership for Quality Meeting, London, 26-27 September 1995.
204. *Improving Quality in Healthcare — Partnership Agenda.* Winchester: Association for Quality in Healthcare, British Medical Association, National Association of Health Authorities and Trusts, 1996.
205. Kuhn T. *The Structure of Scientific Revolutions.* Chicago: University of Chicago Press, 1960.
206. Wennberg JE. Improving the medical decision-making process. *Health Aff* 1988;7(1):99-106.
207. Rawlins R. Paradigm regained: toward a partnership model for quality. *JAQH* 1994;1(2):46-8.
208. Berwick DM. Continuous improvement as an ideal in health care. *N Engl J Med* 1989;320(1):53-6.
209. Berwick DM, Godfrey AB, Roessner J. *Curing Health Care.* San Francisco: Jossey-Bass Publishers, 1990.
210. Berwick DM, Enthoven A, Bunker JP. Quality management in the NHS: the doctor's role - I. *Br Med J* 1992;304:235-9.
211. Berwick DM, Enthoven A, Bunker JP. Quality management in the NHS: the doctor's role - II. *Br Med J* 1992;304:304-8.
212. *Using Clinical Practice Guidelines to Evaluate Quality of Care. Vol 1.* Rockville MD: Agency for Health Care Policy and Research, 1995.
213. Field MJ, Lohr KN, eds. Institute of Medicine (IOM) Committee on Clinical Practice Guidelines. *Clinical Practice Guidelines: Directions for a New Program.* Washington DC: National Academy Press, 1990.
214. Grimshaw J, Russell I. Achieving health gain through clinical guidelines. I. Developing scientifically valid guidelines. *Quality in Health Care* 1993;2:243-8.
215. *Clinical Guidelines. What You Need to Know.* London: Royal College of Nursing, 1995.