Achievements of audit in the NHS

Martin J Buxton

Audit, now usually described as the “systematic, critical analysis of the quality of medical care, including the procedures used for diagnosis and treatment, the use of resources and the resulting outcome for the patient,”1 is seen by many as an inherently “good thing”: an activity that by definition deserves support and encouragement. Certainly, the NHS reforms gave a blanket endorsement to audit, requiring all doctors to be involved on a regular basis. Russell and Wilson added to its stature by describing it as “the third clinical science, with its own theories, techniques, and literature,”2 and argued that, if undertaken appropriately audit “has the potential to deliver substantial benefits to patients and health professionals.”

However, it is appropriate to stand back and view audit as just one more health technology competing for scarce healthcare resources. It is an organisational technology, well within the standard definition of technology,3 and it should be subject to rigorous assessment, in the manner now properly being demanded for all health technologies.4

To use the terminology of the diffusion of a technology,5 audit in the United Kingdom was, up to the end of the 1980s, the preserve of the “innovators” and “early adopters.” With the exception of a small number of top-down “national” initiatives such as the Confidential Enquiries into Maternal Deaths6 originally established in the 1930s and the more recent National Confidential Enquiry into Peri-operative Deaths,7 most reports of audit came from, and related to, the work of well motivated enthusiasts. Theirs was very selective use of the technology of audit: undertaken by choice, in contexts where the findings were likely to fall on fertile ground, and where change in behaviour, by these same enthusiasts, might reasonably be expected to follow. Such case study reports seemed to lend support to the belief that the technology might be efficacious, at least in such atypical settings. Unfortunately there were very few controlled studies undertaken that could show convincingly that audit had a real impact on practice. Thus the NHS reforms, with their simple policy requirement, propelled the emergent technology of medical audit into a new phase of its diffusion. It was precipitously diffused throughout the NHS, embracing not only the sceptical “majority”, but also most of those who would have remained reluctant “laggards” if diffusion had followed its natural course.

This act of faith by the government was shared by many, or at least many chose to suspend any disbelief. Of all the changes proposed in the reforms, audit was among the least contentious. It was not contested on a party political basis nor challenged by the medical profession, particularly once it was clear from further guidance8-10 that a confidential peer review process with professional ownership and control of audit was envisaged.11

... the NHS reforms . . . propelled the emergent technology of medical audit into a new phase of its diffusion.

Publications on audit from the United Kingdom are numerous; a Medline search of UK journals published in 1992 discloses 71 papers indexed on the terms of medical or nursing audit and 104 in which “audit” appeared in the title or abstract. This paper does not attempt a formal systematic review of this literature, nor of the increasing number of activities and initiatives in progress. More modestly, it draws on key papers from this literature and on research with which I have been directly involved, to highlight a growing unease about the lack of evidence of the impact of medical audit as a health technology. It attempts to draw out some common threads relating to the inadequacy of much of the process of audit as practised and to the frequent absence of adequate underpinning research. In so doing, it calls for a much more questioning and sceptical approach to audit.

Uncertain impact of audit

The act of faith that many accepted in encouraging audit had very shaky foundations. Studies had already shown that passive information feedback is not sufficient to produce change.12 13 Nevertheless, many initial descriptive studies, lacking formal evaluation of patient outcomes but focusing on process, were seen as positive towards audit. For example, Gabbay et al came to the preliminary conclusion, based on one year’s evidence, that medical audit apparently resulted in appreciable improvements in clerking and record keeping.14 But later, looking at three years’ data, the study team concluded more sceptically that the extended audit of note keeping had failed to sustain the initial improvement in practice.15

One of the few studies in the United Kingdom to evaluate formally the effectiveness of medical audit was set up before the reforms but reported after the policy changes.16 17 This large study (of 62 practices and 92 general practitioner trainers) used a before and after
design (with a replicated Latin square) during 1984–90 to test the impact of setting clinical standards on general practice behaviour in relation to common childhood conditions. The authors concluded that setting standards did significantly improve aspects of clinical practice, and, for at least one condition, this change in management was reflected in improved health of the affected children. But though the authors’ interpretations were optimistic that the potential for the 1990s is greater than could be achieved in the North of England in the 1980s, it is very clear that many problems were encountered in the process of setting standards and following through the full audit cycle.

Unfortunately, given the magnitudes of expenditures on audit, the lead of this early established project has not been followed in much subsequent research. A review of current or recent evaluation studies relating to audit in the hospital and community health service in England identified 20 studies. None used a randomised control or similar design; four made use of natural experiments, and the bulk were surveys. In part this is a reflection of the inherent difficulties in undertaking research on audit. The very act of evaluation is likely to change the qualitative nature and perhaps even the quantitative substance of the audit under evaluation. For example, the process of evaluation can easily impose a structure on the audit process, the outcome measures chosen for evaluation may influence what are chosen as parameters for audit, and collecting data for the evaluation may change the audit task. Evaluation of audit is open to contamination by the Hawthorne effect; subjects knowing they are being observed behave in the way they believe is expected of them. This phenomenon might lead us to expect that results from studies would be more positive than the results in (unobserved) practice, further weakening the generalisability of the limited evidence of possible impact.

Although there may be many positive reasons to adopt qualitative methods to study research and to accept that strict controls may not always be feasible, there is at least one example that shows that the rigorous methodology of trials can sometimes be applied to audit. A study undertaken in Ontario (Canada) compared audit and feedback with local opinion leader education using a randomised control trial design. The trial measured the success of these alternative methods of encouraging compliance with an evidence based guideline for the use of caesarean section for women who had had one previous section. The trial involved 76 physicians and 16 community hospitals. This methodologically rigorous study concluded that the use of audit and feedback was ineffective, with rates of caesarean section being no different to those in a control group. The performance of physicians encouraged by local opinion leaders was better, but the authors noted that the compliance rate even of that group was no better than 70% and questioned whether it might not be better to use regulatory or economic incentives that had already proved to be effective.

A further problem is that there are a number of alternative, or at least complementary, objectives to audit. Most evaluative studies have assumed that the objective of audit is to achieve better quality of care, through changed practice. But much audit activity has a broader educational role, particularly for junior doctors. Surveys of junior doctors have shown the importance of this role, and, though education may also contribute in the longer term to improving quality of care, the results may not be observable within the institution undertaking the audits, as it contributes to the education of successive intakes of junior doctors, who move on quickly to practise medicine elsewhere. This has major implications for the content, process, and timing of audits – for example, the audit cycle may have to be completed very quickly. It also has major implications for how, and against what criteria, such audit should be evaluated.

The current evidence on the impact of audit is far from perfect or complete: it should, however, serve to instil a degree of caution amongst those who see compulsory audit as a certain route to improved quality of care. It also points to two sets of issues that need particular attention: the inadequate process of audit and inadequate research foundation of audit.

**Inadequate process of audit**

All proponents of audit emphasise that it requires that a proper cyclical process be followed and that the audit loop be completed. Russell and Wilson propose a cycle of nine distinct steps for conducting scientific audit. But nearly all empirical studies point out the practical problems in achieving proper process and the frequency with which the process is inadequate. Kerrison et al studied the implementation of routine audit in hospital general medicine. They identified several major weaknesses in the audit process – for example, lack of explicit and evidence based criteria, use of samples that were too small or in other ways inadequate, and failure to revisit topics and to complete the audit cycle. They also observed the frequency of organisational difficulties in planning and running audit meetings: often small problems in themselves, but each of which might reasonably be expected to diminish the impact of the audit process. They noted particularly that audit activity was still marginal: fitted in as one more task of relatively low import. The same problems have been observed in the context of maternity unit audits: in general, audit meetings “did not
Achievements of audit in the NHS usually measure up to all the recommendations of the Department of Health" and in particular, meetings were likely to be cancelled at short notice and the senior staff were the least likely to attend.

Where outcomes have been measured, inadequacy of process is frequently suggested as a factor limiting impact. Gabbay and Layton suggested that better results might be seen if greater attention was paid to certain interpersonal and managerial aspects of the medical audit process. In the study of management of common childhood conditions previously referred to the difficulties of the process involved in using small groups of general practitioners to set standards were noted and the suggestion was made that better leadership and communication skills are needed. Indeed, the whole concept of audit as a process of peer review can be problematic at a local level where there is only a small group of peers who have to work together on an ongoing basis.

**... scientific audit is a complex and not easily replicable technology.**

Overall, the evidence to date suggests that scientific audit is a complex and not easily replicable technology. It is not a technology embodied in hardware or software or purchaseable "off the shelf" but instead has to be created locally. Audit needs to follow a relatively complex sequence of procedures to be effective, and it entails a difficult set of organisational processes. The skills and commitment necessary to achieve these are not readily available in many of the contexts to which routine audit has now been diffused.

**Inadequate research foundation of audits**

Criterion based audit requires that practice be assessed in relation to agreed criteria or guidelines. But for many aspects of medicine the evidence base for such guidelines is either completely lacking or not readily available to audit practitioners. It has been suggested that audit should be seen as the final step in a good clinical research programme: "good audit is only possible when clinical interventions and innovations are based on good clinical evidence." Most local medical audit has been shown to be criterion based, and the basis for the criteria adopted in local audits may be weak or even arbitrary. This means that those whose behaviour does not meet the criteria may, with justification, be unwilling to change to be consistent with criteria based not on firm evidence but on beliefs they do not share. And even if behavioural change occurs, without a good evidential basis for the criteria there is no way of knowing whether the change can be expected to be beneficial. The question of why change does, or does not, occur is fundamental: it has been suggested that the missing step in the audit cycle is to identify the underlying reasons why standards are not met. One factor may well be the perceived credibility of the standards.

The alternative response to the absence of research underpinning for audit is to turn audit into a research process and to use this research to establish criteria for subsequent audit processes. This has been a particularly common strategy when funds for local "audit" have been relatively easier to obtain than funds for local research. Such "audits" may or may not be good research: a recent paper on cataract surgery randomly allocated patients to day or inpatient surgery, yet called itself a “prospective audit.” Many of the audits undertaken nationally under the auspices of the royal colleges or professional societies do not involve randomisation but do collect case series data from which to generate hypotheses about the appropriateness of the characteristics of the care provided. Thus the value of much local audit is limited by the lack of research underpinning, and much regional and national “audit” in practice focuses on the prior step of trying to provide an evidence base against which to set standards or guidelines.

**Conclusions**

There is no good reason to adopt audit as an act of faith: although it is easy to share the view that audit seems a priori to be a laudable activity, it is no more obviously beneficial than an unproven drug or procedure. We would be appalled if substantial funds had been top sliced from monies for patient care to ensure the rapid diffusion throughout the NHS of a new, and largely unproved, diagnostic technique. But this is the equivalent of what has happened with medical audit. This is not a criticism of those enthusiasts who are trying to make audit work but of the unquestioning policy context in which their work is located.

The limited evidence available does suggest very clearly that the process necessary for good audit is difficult and not easily replicated and maintained over time without appropriate skills and enthusiasm. The relation with research also clearly needs more thought and attention. But rather than linking the diffusion of audit to the generation of the necessary evidence on its effectiveness and on how to maximise that effect, policy has been to earmark funds so that diffusion could proceed without hindrance even from the need to justify the resources that are devoted to it. The new funding arrangements for audit may go some way to improving matters by involving purchasing authorities in determining how audit monies will be spent.

But even with these changes, the increasing demands for a thorough evaluation of this technology called audit remain unmet. The NHS Management Executive pays lip service to creating a research based culture, but the largely unevaluated diffusion of medical audit is quite inimical to that aspiration. Audit, in all its forms, should be treated as any other medical technology and subjected to formal and rigorous assessment.
And, as audit has now been widely diffused in the United Kingdom, this evaluation has to be of the widescale programme that is in place: not just case studies of the work of audit enthusiasts.

This paper is a revised version of that originally presented at the Quality '93 conference in London in November 1993. The research on medical audit undertaken in the Health Economics Research Group has been supported by the King’s Fund and the Department of Health. Helpful comments were made on the earlier draft by my colleagues Tim Packwood and Susan Kerrison, and by the editor. The views expressed, however, remain those of the author and should not be attributed to any of the aforementioned bodies or individuals.