



<https://theses.gla.ac.uk/>

Theses Digitisation:

<https://www.gla.ac.uk/myglasgow/research/enlighten/theses/digitisation/>

This is a digitised version of the original print thesis.

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses

<https://theses.gla.ac.uk/>
research-enlighten@glasgow.ac.uk

**THE DEVELOPMENT OF INTEGRATED AUDIT
FOR THE TRAINING OF GENERAL PRACTICE REGISTRARS**

DR JAMES MURRAY LOUGH FRCGP

MD THESIS

UNIVERSITY OF GLASGOW

SEPTEMBER 2002

ProQuest Number: 10390830

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10390830

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

GLASGOW
UNIVERSITY
LIBRARY:

12928
copy.2.

TABLE OF CONTENTS

<u>Chapter</u>		<u>Page No.</u>
	Summary	12
1	Introduction	16
2	Aims of Thesis	28
3	Trainers' and trainees' attitudes to and experiences of audit	34
4	The introduction and educational impact of an audit project	54
5	Audit and the assessment of competence	69
6	Teaching and learning: issues in knowledge, methods and confidence	103
7	A strategic approach to the integration of audit into training practices	122
8	Facilitating the collection of data for audit	142
9	Completion of the audit cycle – evaluating versus proposing change	150
10	Qualitative audit: assessing an analysis of a significant event	173
11	Introducing audit into SHO vocational training schemes for general practice	189
12	Integrating quantitative and qualitative audit into vocational training for general practice – future implications	201
	References	211

LIST OF TABLES

		<u>Page No.</u>
Chapter 3		
Table 1a	Frequency distribution of areas of care chosen for audit by trainers	48
Table 1b	Frequency distribution of areas of care chosen to audit by trainees	49
Table 2	Attitudes to audit	50
Table 3	Rank correlation (Spearman's) among 'attitude' statements in relation to audit a) Trainers b) Trainees	51
Table 4	Trainer – trainee comparison of responses to attitudinal statements towards audit	52
Table 5	Difficulties with audit	53
Chapter 4		
Table 1	Trainees' ratings of the usefulness of the audit project	65
Table 2	Trainees' estimates of the amount of time spent on the different stages of their audit project	66
Table 3	Areas of the audit project where trainees felt more help was required	67
Table 4	Reasons given by trainees for not continuing the previous year's audit project	68

List of Tables (contd)

Chapter 5		<u>Page No.</u>
Table 1	Trainers' opinions on essential or desirable elements of a registrar's audit project	99
Table 2	Audit assessors' marking exercise (1) on agreement between pairs	100
Table 3	Audit assessors' marking exercise (2) on agreement between pairs	101
Table 4	Factor analysis of assessment of 333 audit projects (1994 – 1997)	102
Chapter 6		
Table 1	Trainers' versus assessors' judgements on identifying substandard criteria	117
Table 2	Trainers' marking behaviour when assessing criteria as substandard	117
Table 3	Trainers' versus assessors' overall judgements for each project	118
Table 4	How confident are you with criterion audit method?	119
Table 5	How confident are you with the teaching of criterion audit method?	119
Table 6(a)	Has your confidence with criterion audit method changed since the submission of a summative assessment audit project?	119
Table 6(b)	Has your confidence with teaching criterion audit method changed since the submission of a summative assessment audit project?	120
Table 7	Is more help needed with the teaching of:	120

List of Tables (Contd)

Chapter 6 (Contd)

Table 8	Did audit teaching involve:	120
Table 9	How much audit teaching input was there from:	121

Chapter 7

Table 1	Practice appointment systems – differences in proportions of training practices achieving programme criteria	138
Table 2	Chronic Disease Management – differences in proportions of training practices achieving programme criteria	139
	• Asthma	139
	• Diabetes	139
	• Epilepsy	139
	• Hypertension	140
	• Rheumatoid Arthritis	140
Table 3	Significant Event Analysis – difference in proportions of training practices achieving criterion	141

Chapter 8

Table 1	Average cost (in pounds sterling) of transferring data for workload audit per 1000 patients	149
Table 2	Average cost (in pounds sterling) of retrieving data for each chronic disease	149

List of Tables (Contd)

Chapter 9

Table 1	Change in trainers' opinions on essential or desirable elements of a registrar's audit project	164
Table 2(a)	Audit assessors' initial marking exercise to identify an incomplete audit cycle	165
Table 2(b)	Summary of statistics of 26 assessors' judgements of 20 audit projects to identify an incomplete audit cycle	166
Table 3(a)	Audit assessors' final marking exercise to identify a 'poor' project	167
Table 3(b)	Summary statistics of 24 assessors' judgements of 11 audit projects	168
Table 4	Sensitivity and specificity of the instrument using one or two assessors	169
Table 5	Factor analysis of judgements of 210 eight criteria projects (1996 – 1998)	170
Table 6	Reliability testing: summary statistics of 20 assessors' judgements on assessments of 100 audit projects (1996 – 1998)	171

Chapter 10

Table 1	Summary statistics of marking exercise for proportion of agreements on 100 SEAs by 20 assessors	188
---------	---	-----

Chapter 11

Table 1	Actual (<i>with expected</i>) audit project submissions per unit per specialty	199
Table 2	Number of audit projects submitted by SHOs over each six-month post	200

LIST OF FIGURES

		<u>Page No.</u>
Chapter 5		
Figure 1	Summative Assessment – Audit Marking Schedule	95
Figure 2	Instructions for audit assessors	96
Figure 3	Development of strategy for ‘referral’ of an audit project	97
Figure 4	Assessment of 102 audit projects: August 1994 – July 1995	98
Chapter 9		
Figure 1	Instructions for audit assessors and audit marking schedule	172
Chapter 10		
Figure 1	The development of the process for assessment a significant event analysis	186
Figure 2	Significant Event Analysis – Assessment Schedule	187

MD by published research

The work for this thesis has taken ten years to complete. There are eight published papers in peer reviewed journals with a further paper in press and the author assisted in the work of three further papers relevant to the thesis, a total of 12 papers.

The research presented in the thesis spans a period between the implementation of a specific criterion for audit in training practices in April 1991 and the subsequent implementation of new criteria covering audit in training practices informed by the academic method and findings presented in this thesis. The templates for assessment of an audit project are now the accepted national standard for all training practices throughout the United Kingdom.

ACKNOWLEDGEMENTS

Professor Stuart Murray has offered support and encouragement throughout this thesis. I also heartily acknowledge the loyalty shown to me by members of the Audit Development Group without whose collaboration for over 10 years I could not have completed this work.

List of Publications (main author)

LOUGH, J.R.M., McKAY, J. and MURRAY, T.S. (1995) Audit : trainers' and trainees' attitudes and experiences, *Medical Education*, 29, pp.85-90.

LOUGH, J.R.M., McKAY, J. and MURRAY, T.S. (1995) Audit and summative assessment : two years' pilot experience, *Medical Education*, 29, pp.101-103.

LOUGH, J.R.M., McKAY, J. and MURRAY, T.S. (1995) Audit and summative assessment : A criterion-referenced marking schedule, *British Journal of General Practice*, 45, pp.607-609.

LOUGH, J.R.M., MCKAY, J. and MURRAY, T.S. (1997) Audit and summative assessment : system development and testing, *Medical Education*, 31, pp.219-224.

LOUGH, J.R.M. and MURRAY, T.S. (1997) Training for audit: lessons still to be learned, *British Journal of General Practice*, 47, pp.290-292.

LOUGH, J.R.M. and MURRAY, T.S. (1997) Teaching audit – lessons from summative assessment, *British Journal of General Practice*, 47, pp.829-830.

LOUGH, J.R.M., WILLMOT, M. and MURRAY, T.S. (1999) Supporting practice-based audit: a price to be paid for collecting data, *British Journal of General Practice*, 49, pp.793-795.

LOUGH, J.R.M. and MURRAY, T.S. (2001) Audit and summative assessment - a completed audit cycle, *Medical Education*, 35, pp.357-363.

LOUGH, J.R.M., KELLY, M.H. and MURRAY, T.S. (In press) SHOs in vocational training schemes for general practice: Audit and the educational contract, *Education for Primary Care*.

List of Publications (Other than first author)

MCKAY, J., LOUGH, J.R. and MURRAY, T.S. (1995) General practitioner trainees' experience of undertaking audit projects: preliminary report from the west of Scotland region, *British Journal of General Practice*, 45, pp.301-303.

BOWIE, P., GARVIE, A., OLIVER, J. and LOUGH, M. (2002) Impact on non-principals in general practice of the summative assessment audit project, *Education for Primary Care*, 13, pp.356-361.

MCKAY, J., MURRAY, T.S. and LOUGH, J.R.M. (2002) Registrar audit and implementation of change: a pilot project, *Education for Primary Care*, 13, pp.336-339.

SUMMARY

In 1991 the Joint Committee on Postgraduate Training for General Practice (JCPTGP) issued a new criterion for training practices stating that they “must provide opportunities for trainees to become familiar with the principles of medical audit and to participate in medical audit; and they must be able to demonstrate their trainees have actually done so.”

It is possible that no training practice in the west of Scotland could have implemented this criterion.

This thesis considers the development of a model appropriate for a training environment which overcame the difficulty in interpreting whether the criterion was being implemented. This required clear learning objectives to be set integrated into a system where competence in achieving these objectives could be assessed.

In 1992, 131 questionnaires were returned from trainers (85%) and 104 from trainees (67%) enquiring into their experience of and attitudes to audit. 52% of trainees stated that they had started collecting data for audit purposes, 23% claimed to have set standards as part of an audit, and 12% claimed to have implemented change as part of an audit. Audits of chronic diseases predominated for both groups. Attitudes to audit were generally positive. Time and resources were the two main difficulties cited by both trainers and trainees, possibly explained by the time-consuming nature of collecting data for chronic diseases.

From August 1992 all trainees in the west of Scotland were required to submit an audit project using a format based on a criterion audit described by Irvine. 99% of the 103 trainees responded to a questionnaire about the perceived usefulness of the

project. 87 trainees (85%) found it to be a useful method for carrying out future audits with the majority of trainees completing their project in less than one working week. Those trainees who chose their own subject were more likely than those who did not to evaluate change having collected two sets of data.

In 1994 104 trainees (89%) responded to a questionnaire about the perceived benefit to them in undertaking such a project. For 82 (79%) this was their first ever experience of audit and consequently 88 of them (85%) felt more confident about the possibility of introducing change in their next practice with none feeling less confident. 70 trainees (87%) stated that the audit project was rarely or never discussed at a practice meeting and only 35 trainees (34%) confirmed that they had actually seen the previous trainee's project.

The construction of a criterion audit project provided a useful format for allowing a trainee to demonstrate their competence in understanding some of the principles of audit method. Five specific criteria were developed from an initial fourteen elements considered to be important in the inclusion of an audit project. The final choice of criteria was informed by responses from 70% of the 155 trainers. A system for assessing the projects with three assessors independently marking a project as part of a screening process achieved an acceptable balance of sensitivity and specificity. Issues of validity, reliability, acceptability, feasibility and educational impact were considered.

Interviews with 22 registrars who had required further training after two attempts at submitting their audit project suggested that the trainer's advice (or lack of it) was a cause for concern.

In order to test this a marking exercise consisting of five submitted audit projects was completed by 114 trainers (72%). Three trainers (3%) correctly identified the

five criteria - one from each project – which had been judged as deficient by the “expert” assessor group. Despite the limitations of the design of this study concerns were raised about the implications for trainers’ ability to teach criterion audit to their registrar.

In order to provide support for the trainers and to maximise the opportunities for trainees to participate in audit a programme for the region was constructed covering organisation of appointment systems, chronic disease management and significant event analysis with educational objectives set out for each area. Progress in implementing the programme was assessed at each reaccreditation visit and evaluations were completed in 1998 and 2001. All areas of the programme showed modest improvements between the evaluation dates although few reached statistical significance.

The time involved in and the cost of collecting data for audit purposes were evaluated by offering ten training practices audit support staff to collect their data for parts of the audit programme. The costs were compared with the hypothetical use of a practice receptionist or the practice nurse. The conclusion was that data collection carries significant costs both in time and expense for a practice and the need to agree on appropriate use of practice staff is vital.

Between 1996 and 1997 an increasing number of registrars was evaluating the change they had proposed in their audit project. There was also a significant increase in the proportion of trainers who felt that a completed audit cycle should now define the audit project. An increase to eight criteria followed with two assessors being used to screen the projects without compromising sensitivity or specificity. Registrars were therefore now expected to demonstrate their competence as defined by evaluating rather than proposing change.

Significant event auditing added a qualitative format for an audit project. The analysis of such an event involved addressing four specific questions with two assessors reviewing each analysis. The integration of quantitative and qualitative methods encouraged training practices to think more broadly about different approaches to teaching the assessment of quality of care.

Between 1998 and 2001 all senior house officers on vocational training schemes were asked to submit a criterion audit cycle or a significant event analysis in each post. Teaching and support were provided. The output ranged from 28% (accident and emergency) to 45% (geriatrics) of the total number of projects expected.

The JCPTGP revised its criterion for audit in training practices in 2000 proposing a model based on the work in this thesis. The lessons learned have implications for the non-training environment of general practice.

CHAPTER 1

INTRODUCTION

The Cohen Committee (1950) recognised the need for appropriate postgraduate training of doctors and the first annual report of the College of General Practitioners (1953) described the College being founded in 1952 to “encourage, foster and maintain the highest possible standards of general medical practice”. In 1965 the College of General Practitioners described the special vocational training for general practice and the need for high quality trainers was emphasised further in the Todd Report (1965) where enhancement of general practice as a career could only be secured through compulsory vocational training. This was confirmed in 1977 and enshrined in law in 1979 (National Health Service Regulations, 1979).

Based on evidence from the College of General Practitioners (1966) the Todd Report further proposed a five-year programme of training which, for reasons of pragmatism and finance, never materialised.

Prior to 1973 the appointment of trainers was the responsibility of Local Medical Committees. Thereafter this responsibility was overseen by regional general practice sub-committees which developed more detailed criteria for appointing trainers (JCPTGP, 1980). Describing audit, Gray (1984) stated that “good training practices almost invariably have several audits running at one time and these may be one of the most influential forms of education for trainees”.

The final version of the work of the Royal Commission on the National Health Service under the Chairmanship of Sir Alex Morrison (1975) was published. In its evidence to the Committee the British Medical Association (1977) said “we are

not convinced of the need for further supervision of a qualified doctor's standard of care." In response the Commission reported "we are not convinced that the profession regards the introduction of medical audit and peer review with a proper sense of urgency."

Referring specifically to the evidence provided by the Royal College of General Practitioners (1977) in paragraph 5.4 the Commission also stated that "Medical education needs radical reshaping to place much greater emphasis on continuing education and medical audit".

In 1979 the chairmen of the General Medical Services Committee (GMSC) and the College made a joint declaration that medical audit should be a professional activity.

Quality in General Practice

A paper by Honigsbaum (1972) on the wide variations in the quality of care in general practice was widely reported in the medical and lay press. Citing 137 references it provoked a vigorous response (Marson *et al*, 1972). Despite describing its "serious shortcomings" - not least "a medical audience may be forgiven for not seeking out all the references quoted" - the authors acknowledged that "it challenges the profession to develop methods for evaluating medical care".

A briefing paper to Council of the College from Donald Irvine (1983), then chairman, stated that "Despite many major improvements unacceptable differences in the quality of general practice still exist today". He described the quality of care in general practice as "Our outstanding problem". The resulting "Quality Initiative" was in effect the Council's policy on quality issues (RCGP, 1983). Its two aims were that:

- each general practitioner should describe his or her work and should be able to say what services his or her practice provides.
- each general practitioner should define specific objectives for the care of his or her patients and should monitor the extent to which these objectives are met.

This initiative was to be implemented in every general practice in the United Kingdom within 10 years. There followed a series of constructive statements and documents from successive College Councils (RCGP, 1984-5; RCGP, 1985a,b,c; RCGP, 1987).

The positive attitude to clinical audit was also taken up by successive conferences of representatives of Local Medical Committees of the BMA (1979 and 1989).

The academic contribution from the College to the work on quality issues and rigorous audit methods started with the work of the Birmingham Research Unit in the 1970s, in particular its Practice Activity Analysis which were some of the earliest examples of individual doctors' performances of variables such as hospital referrals (RCGP, 1978a) and home visiting (RCGP, 1978b) compared to the range of performances as demonstrated by their colleagues.

Further publications in the Occasional Paper series on medical audit included Watkins' PhD Thesis (Watkins, 1981), Sheldon's Gold Medal Essay (Sheldon, 1982) and Baker's work on practice assessment of quality issues (Baker, 1988). In 1986 the College published *In Pursuit of Quality* (Pendleton *et al*, 1986) and in 1987 *The Front Line of the Health Service* (RCGP, 1987) both of which focused firmly on the importance of performance review and the principle of peer review.

With the publication of criteria to allow members of five years' standing to apply for College Fellowship (RCGP, 1990) Gray (1990) described the College as being "at the heart of medical audit in general practice".

In April 1989 the College announced that "Research, education and educational audit are the three top priorities for the College" (RCGP, 1989).

Thus in just under 10 years the College had laid out its audit stall.

Early evidence of audit activity in training practices

In 1970 222 general practitioners in 12 vocational schemes in England, Wales and Northern Ireland were invited to complete a questionnaire which was a modification of that used within the British Medical Association Planning Unit (Irvine & Jeffreys, 1971). 190 doctors replied (86% response rate). This survey was followed up by personal visits to 50 practices chosen at random to confirm data in the questionnaire (RCGP, 1972). In response to the question: "*Does the practice have any form of medical audit?*", 29% of teachers' practices undertook some form of "routine audit" consisting of case conferences (70%), monitoring of workload (67%) and death analyses (20%).

The conclusion was that "if routine audit does not happen in the teaching practice it is unlikely that the young doctor will accord it a high priority when he becomes a principal".

In 1980 the fourth National Trainee Conference was held in Exeter. The theme of the conference was based on the presentation and analysis of 1457 questionnaires from trainees throughout the United Kingdom concerning all aspects of training. One of the questions asked: "*Have you as a trainee been shown in your training practice any form of clinical audit?*" A median of 31% replied positively with a

range across the regions of the United Kingdom from 16% to 54%. A positive response was received from 32% in the west of Scotland. The conclusion was that there appeared to be a gap in vocational training which required attention. The regional variation was significant indicating that regional factors were more important than characteristics of individual trainees (Ronalds *et al*, 1981).

This survey was repeated in 1988 after the 12th National Trainee Conference (Crawley & Levin, 1990). From 1581 replies to the same question as above 64% replied positively. The authors speculated that one possible reason for the perceived increase in certain activities such as audit in training practices was the introduction of tighter guidelines for the approval of trainers which were to be enforced through visits to the regions (JCPTGP, 1985).

The Joint Committee on Postgraduate Training for General Practice

The Royal College of General Practitioners was responsible through its Vocational Training Sub-Committee of the Regional Postgraduate Medical Education Committee for initiating and developing ideas of training for general practice as well as for approval and re-approval of trainers. The Sub-Committees set their own criteria for approval guided by the statutory regulations.

The formation of the Joint Committee on Postgraduate Training for General Practice was conceived, it is said by Lawson (1992), by Ekke Kuenssberg on the top of a London bus. The idea, however, had its origins in 1974 in a partnership between the College's Vocational Training Committee and the General Medical Services Committee with additional members from the Postgraduate Deans, the Councils for Postgraduate Medical Education in Scotland and England and the Conference of Regional Advisers. It assumed responsibility in law for the approval of training posts for general practice training (JCPTGP, 1976).

In 1980 and 1982 the Joint Committee published further guidelines and criteria for the appointment of trainers (JCPTGP, 1980; JCPTGP, 1982a). Appropriate areas in relation to audit were:

- trainer assessment - a trainer “subjects his work to peer review”. The trainer’s competence in this area was assessed by: “to what extent does the practice undertake audit procedures and how much is the trainee himself involved in patient care evaluation?”
- hospital training for general practice - “particular attention should be paid to departmental audit in management review”.
- the training practice - “there should be an effective appointment system so that the trainee can learn how to run one well” and “there should be arrangements to enable the trainee to audit his own work critically”.

Visits to the regions were carried out to ensure that the guidelines were being observed which confirmed that wide regional variations were in place. Schofield and Hasler (1984a, b, c) gave a detailed description of the results of a working party set up in their region in 1981 to examine the criteria and methods of the appointment of trainers and training practices. They stated that a “trainee should have the opportunity of seeing audit methods in practice and should audit his own work”. Whether this was assessed is not discussed although the method of assessment was based on that described in *What sort of doctor?* (RCGP, 1985b) which assessed audit as one of the professional values using the Practice Profile as its source of information.

In its first 10 years the Joint Committee averaged 12 regional visits per year. As a result of these visits the guidelines and criteria were reviewed (JCPTGP, 1985). Those relevant to audit were:

- teaching abilities - It is reasonable to expect trainers to understand how self-audit, performance review and research projects are conducted and to ensure that every trainee is shown examples of such activity and also provided with every opportunity of putting these into practice.
- assessment by peers - regions should devise systems which enable doctors to assess each other's and their own performance as GPs and trainers. A willingness to submit to such appraisal by peers is an important requirement as a teacher.
- peer review - trainers should be prepared to show that they regularly review both organisational and clinical aspects of their practice. They should therefore be able to demonstrate appropriate plans for monitoring such activities such as repeat prescribing, immunisation programmes and care of patients with chronic diseases.
- re-approval - regions will ensure that trainees have been offered experience of performance review in clinical audit. Trainees should have taken part in performance review either as part of a practice programme or as a special project.

Between 1976 and 1989 therefore the Joint Committee carried out its responsibilities "based on the observations of and heavily dependent on the inevitably subjective assessments of the visitors of the day". It therefore changed the primary focus of its emphasis from scheme recognition to regional re-

accreditation. With the profession moving towards a recognition that formative and summative assessments would be required to guarantee a national standard of entry to general practice new guidelines were published (JCPTGP, 1992a). Specific guidance had been published on the accreditation and re-accreditation processed at practice visits (JCPTGP, 1982b). Each practice visit in a district or scheme was to provide a profile which would include "the results of practice audits". More specifically:

- the training experience - trainees should have regular experience of determining and reviewing criteria and standards of care and of performance monitoring within the practice, so acquiring the knowledge and skills needed to carry out and implement the results of practice audits.
- trainer selection - the attributes of the trainer as a clinician will include a commitment to audit and peer review. The qualities of the trainer as a teacher will include a commitment to peer review related to teaching. The practice suitable for training will include involvement in quality assurance.
- special features of hospital posts - The extent to which medical audit contributes to the trainee's experience.

Between April 1989 and April 1991 discussions between and within the Government, the Royal Colleges and the Joint Committee formed the basis upon which the teaching of audit was to be defined. For this reason these discussions are described in some detail.

The JCPTGP and the Educational Criterion for Audit, April 1991

Following the publication of the White Paper Working for Patients and the subsequent Working Paper on Medical Audit (Department of Health, 1989a & b) a

letter from Dr J C Rivett on 10th April 1989 confirmed the Government's view that medical audit should be professionally based. The letter invited the Joint Committee to contribute to the discussion on medical audit based on its "great experience in the development of systems and professional review". It was further asked how medical audit might develop in future in training practices. In its reply, the Joint Committee supported the principle of medical audit as an educational tool but there was disagreement about negotiations over contractual commitment to participate in audit. Dr Rivett left the initiative to discuss the matter further with the Joint Committee.

The minutes of the Fifteenth Meeting of the Joint Committee state: "that there should be clarification of the concept of audit in general practice and that a framework for the development of audit should be given as soon as possible". The Joint Committee's 1985 Criteria on the Approval and Re-Approval of Trainers in General Practice were to be reviewed and all visits in 1989 were to concentrate on how regions were implementing audit.

Government circulars (National Health Service Circulars, 1989a & b) on the structure within which audit should be undertaken were considered by the Joint Committee at its meeting in November 1989. Both circulars identified and emphasised the need for education and training in audit.

In its reply to the circulars the Joint Committee stated that "it recognised the selection of trainers based on defined criteria and standards with an emphasis on peer review as a form of medical audit". Its final paragraph stated that "General practitioners in training will need to acquire the knowledge and skills needed to undertake audit in their own practice in the future and that training in audit methodology should be available during the vocational training period, in both its

hospital and general practice setting". It promised to consider how to achieve this "as a matter of high priority".

The implications for vocational training and specifically how trainee general practitioners could learn the skills needed for clinical audit were discussed. Dr Bill Styles, then Chairman of the Education Division, stated that "despite promotion by the College for over 10 years on the use of clinical audit as a desirable characteristic of training practices, visit reports had confirmed the wide regional variations in progress in this area".

All Regional Advisers, the Joint Higher Training Committees and the Royal Colleges were to be asked for a report on progress on the development of medical audit in teaching practices. The five questions to be answered were:

- Is audit being developed?
- Does audit form part of the regional criteria for appointment?
- What percentage of trainers practise audit?
- Are specific areas covered?
- How do trainees learn about audit?

In its letter to all Regional Advisers and lead visitors to training practices the Joint Committee confirmed that it would be implementing a new minimum educational criterion on the use of medical audit in training practices from April 1991 and that the Joint Committee's visitors would be looking at the opportunities being provided by training practices for trainees to learn the skills of medical audit. In particular the Committee was "interested in the different types of audit, the facilities available for it and teaching methods being used in the regions". The replies would form the basis of a report to the Joint Committee on the current status of audit in training practices.

The subsequent report on the replies from Regional Advisers, Joint Higher Training Committees and the Royal Colleges was discussed at its meeting of May 1990. The detail was variable and the questions were not all addressed. Teaching of audit was usually through a variety of half or one day courses. In his response, Dr Ian Bogle stated that it was clear that medical audit would need to become an integral part of every training practice and it was therefore important that trainees were given appropriate training in this area and that this should be in the training practice setting rather than on a day-release course.

The Committee consequently agreed to introduce a new “minimum educational criterion” on medical audit. This would ensure that training practices “should be able to provide their trainees with the opportunity to practise and learn the skills of medical audit”.

The Committee met in August to confirm the wording of its minimum educational criterion and standard on medical audit to take effect from 1st April 1991. It stated that:

“All training practices must provide opportunities for trainees to become familiar with the principles of medical audit and to participate in medical audit; they must be able to demonstrate that trainees have actually done so.”

This was distributed to all Regional Advisers in the United Kingdom in September 1990. Thus the Royal College of General Practitioners, the Joint Committee on Postgraduate Training for General Practice and the Government were all agreed that audit method should be taught, practical audit should be demonstrated and systematic change should be implemented in training practices in the United Kingdom as an example for others to follow. The necessary knowledge and skills

required were assumed to be covered within the training curriculum and there is no evidence that obstacles to these processes were anticipated.

CHAPTER 2

AIMS OF THESIS

The questions underpinning this thesis started with an ambiguity. The messages emanating from the various educational bodies were positive; systems were in place to facilitate the teaching of audit method in the assessment of quality of care. Early indications from the consumers, in this case trainee general practitioners, were more sceptical. Verification in training practices confirmed this. The definitions covering audit at the time were varied and confusing and there was an urgent need for a working definition, i.e. one which took account of those responsible for teaching the subject and a system to check whether the teachers and learners were competent to address this. Although projects had been popular as part of the training curriculum for general practice their formats were variable and often unfocused. There was therefore an urgent need to answer the question of how training practices could ensure that audit method was being taught whereby future principals in general practice could systematically and critically analyse the quality of their medical care.

A formal null hypothesis of no difference between training practices implementing the JCPTGP criterion and those who had not was difficult to construct and made little sense.

The research supporting the thesis therefore considered the two parts of the criterion:

- whether training practices were providing opportunities for trainees to participate in medical audit.
- to construct a system whereby a trainee had to demonstrate that they were familiar with the principles of medical audit. In addition, the research sought to explore reasons for the deficiencies where they were found.

Clarifying the wording of the criterion was a priority. “Provide opportunities for”, “become familiar with”, “principles of medical audit”, “to participate in”, “able to demonstrate that trainees have actually done so” are specific examples where deficiencies in an educational context had to be recognised in order to minimise confusion in the potentially high stakes situation of a re-accreditation visit. In reality, whether the criterion was being implemented or not was virtually impossible to judge.

Specific objectives in achieving the aims were:

- to assess experience with and attitudes to audit in training practices.
- to develop an audit project as a means of defining both quantitative and qualitative audit methods.
- to develop a system for the successful submission of an audit project as a test of competence in analysing the quality of an aspect of medical care.
- to assess the confidence of both teacher and learner with audit method.
- to establish a core syllabus for supporting audit in training practices.

- to assess measuring improvement rather than measuring performance.
- to assess the hospital component for vocational training in the delivery of a satisfactory audit project.
- to develop and test a system for assessing an analysis of a significant event.

The work in this thesis forms part of a slowly evolving picture of the reality of clinical audit as one means of assessing quality of care in training practices in the west of Scotland. Events following the change of legislation of the GP contract in 1990 against rigorous opposition from the profession exposed the audit process to much scrutiny and offered an opportunity to explore the reality as opposed to the rhetoric. The decade began with the establishment of a specific criterion for audit in training practices and ended with a change in that criterion, which took into consideration many of the lessons learned from the work in this thesis.

Although the work began in 1992, published research began in 1995 and has continued to date. As a result, the objectives evolved over a period of four to five years and were not immediately apparent at the beginning of the research.

Scope and limitations of the Thesis

The work in this thesis involves published research which took place in training practices in the west of Scotland between 1992 and 2001. The systems developed from this work to assess the competence of trainee GPs in measuring aspects of the quality of their care are now established in all training practices throughout the United Kingdom. Some of the results therefore are presented on a national basis; the legislation covering vocational training is similar throughout the United Kingdom.

The world-wide literature on audit is vast. As the focus of this thesis is on training practices in the United Kingdom the literature base used in the thesis reflects this. There are, however, implications from the research beyond training and appropriate reference such as to non-training practices is considered where appropriate.

The definitions of audit are many and varied.

"The word audit inspires nobody." With these words Richard Smith (1992a) began his introduction to a book based on a collection of published articles in the British Medical Journal in the early 1990s entitled, paradoxically, "Audit in Action". Perhaps he was reminded of Shaw's combinations of 96 words which he showed (1980) had been, or could be, used to mean a review of health care.

Baker (1990a) described four types of medical audit in use in general practice. One of these - an audit project - he described as being virtually synonymous with medical audit in the minds of general practitioners. In training practices they had been taken as evidence of audit being carried out and, in some cases published, and had led to substantial improvement in care. He pointed out, however, that the method was "loosely based on a familiar audit cycle approach", completion of which was rarely demonstrated. Included in the reasons for this were lack of knowledge of audit method and lack of commitment to act on results.

The work in this thesis builds on the advantage of an audit project being already familiar to the training curriculum in general practice. The application of a more objective and systematic approach to the execution of an audit project is attractive in a teaching and learning environment. Donabedian (1981) described the use of explicit criteria as reducing "to a minimum the use of health care professionals

whose time is exceedingly costly, and whose interest in the review process is less than enthusiastic". Considered to be a classic illustration of criterion audit is illustrated by Lembcke (1956). He emphasised the importance of well-defined criteria based on accepted and verifiable evidence and showed how with simple techniques such as a card system for collecting data a comparison of proportions before and after an intervention could demonstrate significant improvement in the quality of care. He included personal feedback and although this was carried out in the context of teaching hospitals it provided a quantitative method of audit which could add structure to an audit project. This method was described in more detail by Shaw (1990). He emphasised the repeatable nature of obtaining objective quantitative data. Coles (1990) described a "double loop cycle for audit to become more educational" where he applied the structure of criterion audit to the learning cycle after Kolb (1984). The advantages of this are described later in the thesis. A criterion-based audit method was therefore taken as the quantitative example of an audit project for a trainee general practitioner.

Flanagan (1954) described a critical incident technique, lessons from which resulted in the safety reporting system in use today in the aviation industry. The principles of the critical incident technique required a qualified observer to give a clear account of actual events contemporaneously, not retrospectively. More attractive in an educational setting was the reflection on events which were retrospective - either examples of good or not so good practice - but which often has, as a result, strong emotional impact. This form of case-based audit came to be known as a significant event analysis (Pringle *et al*, 1995) and was chosen as the qualitative type of audit project which would be included in the programme of audit method in training practices and in vocational training schemes in hospitals in the west of Scotland.

The JCPTGP minimum educational criterion and standard of April 1991 stated that "all training practices must provide opportunities for trainees to become familiar with the principles of medical audit and to participate in medical audit; and they must be able to demonstrate that trainees have actually done so." Immediately prior to this the Government defined medical audit in "Working for Patients" (Department of Health, 1989a) as:

"the systematic critical analysis of the quality of medical care, including the procedures used for diagnosis and treatment, the use of resources, the resulting outcome and quality of life for the patient."

The key elements of this definition were expanded upon in Working Paper No.6 (Department of Health, 1989b) and a separate Scottish Working Paper 2 (Department of Health, 1989c).

The Shorter Oxford English Dictionary (1985) describes a thesis as "a proposition laid down or stated as a theme to be discussed and proved".

The theme laid down in this thesis is the deconstruction of the JCPTGP criterion, the wording of which made it virtually unworkable in its ability to judge whether or not it was being implemented. As a consequence of this, further research developed the theme in constructing a more sustainable criterion where learning objectives of demonstrating an understanding of quantitative and qualitative audit methodologies could be met.

CHAPTER 3

TRAINERS' AND TRAINEES' ATTITUDES TO AND EXPERIENCES OF AUDIT

At the end of their year's training, a trainee is required to send a confidential report to the Regional Adviser on all aspects of their training experience during the year. The report consists of the educational criteria set out by the JCPTGP, the interpretation and implementation of which are the responsibility of the Regional Adviser. The degree to which each criterion is met is divided in the report into:

- meets all aspects of the criterion.
- meets some aspects of the criterion.
- serious deficiencies with the criterion.

The trainees' responses then formed part of the discussion at the re-accreditation visits to the practices following which the reports were destroyed, militating against quantifying the deficiencies. The criterion in the report relating to audit is worded as follows:

"The practice should be involved in medical audit and the trainee should be encouraged to take an active role in audit."

The evidence required from a practice at a re-accreditation visit to demonstrate that the criterion was being met is that one audit carried out in the practice should be available for inspection. The word "audit" was not defined.

Feedback from the trainees had consistently highlighted serious deficiencies on this criterion suggesting that active practice audit was not taking place. Over a three-month period in early 1992 all audits presented at practice re-accreditation visits were assessed for having collected data (measuring performance) or evaluating change (measuring improvement). A total of 48 audits were evaluated. Six audits (12%) had measured improvement in care and none had repeated this at a later date.

The accumulating evidence, therefore, suggested that audit activity was at best patchy and unfocused relating to the collection of data, with a very small number resulting in the evaluation of change.

Method

A visit was carried out by the author to each of the 12 trainers' groups in the west of Scotland region over a period of three months in early 1992 to explore the helping and hindering forces influencing audit activity. A questionnaire was designed based on these discussions and the issues formed the basis for a series of attitudinal statements which was piloted on a 10% sample of trainers and trainees in the region. No changes were made.

The questionnaire was divided into four parts:

- Demographic details of the doctors and their practices.
- Experience with audit as defined by Fowkes' audit cycle (Fowkes, 1982)
 - (a) Started data collection (observing practice).
 - (b) Set standards.
 - (c) Implemented change.

- Five attitudinal statements were formed from statements made in the training groups relating to the perceived usefulness and appropriateness of audit. Positive and negative statements were alternated and arranged on a likert scale from 1 (strongly agree) to 5 (strong disagree).
- Difficulties in carrying out audit were explored.

The completed questionnaire was sent to all 155 trainers and their trainees in May 1992 with non-responders being followed up two weeks later.

Analysis

Statistical analysis was carried out using SAS version 6.03.

Positive and negative attitude statements were coded in such a way as to achieve conformity in “direction of effect”, i.e. “high is good” convention. This made the interpretation of results, particularly correlation coefficients, more straightforward.

Associations among responses to the five attitudinal statements were analysed by rank correlation, performed separately for trainers and trainees. Trainers’ and trainees’ responses to the five attitudinal statements were compared by Mann-Whitney tests, the null hypothesis being one of no difference in median values between the two groups.

Results

310 questionnaires were sent out and 235 returned (76%) after two mailings - 131 from trainers (85%) and 104 from trainees (67%).

Demographic details

The practices were equally divided between surgeries which were owned by the practice and health centres rented from Health Boards. Undergraduate teaching was undertaken by 65 practices (50%) and 97 practices (74%) received deprivation payments.

129 practices (98%) were computerised, 114 of which (87%) used the GPass relational database (General Practice Administration System for Scotland) which is distributed free to general practices in Scotland and funded by the Scottish Office Home and Health Department. 117 practices (89%) received the Scottish Prescribing Analysis Level 2 data (similar to PACT Level 3 in England and Wales).

Experience with Audit

- a) 118 trainers (90%) and 54 trainees (52%) stated that they had started collecting data with a view to starting an audit.
- b) 73 trainers (56%) and 24 trainees (23%) claimed to have set standards as part of an audit.
- c) 71 trainers (54%) and 13 trainees (12%) claimed to have implemented change as part of an audit.

The topics chosen for data collection, standard setting and completed audit cycles are shown in Tables 1a (trainers) and 1b (trainees).

Attitudes

Attitudes to audit are shown in Table 2a (131 trainers) and 2b (104 trainees). Matrices of the rank correlations (Spearman's) among the five attitude statements

are shown for trainers in Table 3a and for trainees in Table 3b. All correlation coefficients were significant at the 5% level with the exceptions of:

- “audit may be used to assess doctors” with “audit improves patient care” (trainers)
- “audit may be used to assess doctors” with “audit is an appropriate use of resources” (trainees).

All the correlations were positive. There is thus no indication that positive attitudes to one aspect of audit tended to be associated with negative attitudes to any other aspect, or vice versa.

In general the correlation coefficients were poor to moderate, ranging from 0.12 to 0.52 in the trainer group and 0.14 to 0.61 in the trainee group. One interpretation of this is that attitudes are not monolithic, in the sense that respondents do not tend to rate all five statements similarly, such as “strongly agree” with all five or give a neutral response to all five. This would have resulted in higher observed correlations.

The correlations of the use of audit to assess doctors with the other variables are lower than those involving the variables other than the use of audit to assess doctors. This suggests that attitudes to the role of audit in assessing doctors may be rather different from attitudes to the other aspects of audit captured in the remaining four variables. This could merely reflect a broader wariness of assessment than a specific reaction to audit.

Table 4 shows trainer - trainee comparisons of responses to the five attitudinal statements. None of the differences is significant at the 5% level, although the

“may be used to assess doctors” statement is marginally so. There is thus modest evidence that the statement “may be used to assess doctors” is viewed in a systematically different way by trainers and trainees. With a median response of 4 for trainers and 3 for trainees, the latter group may be less convinced of the use of audit in assessing doctors than their more experienced colleagues.

Difficulties with Audit

Time and resources – the latter not being defined – were the two main difficulties cited by both trainers and trainees. The complete list of difficulties for both groups is shown in Table 5.

DISCUSSION

Being a training practice is one of the factors which has been shown to influence participation in audit activity. In 1982, one year after vocational training for general practice acquired legal status, a study was carried out in the Severn Faculty of the RCGP (Baker, 1985) which compared standards in training and non-training practices. A questionnaire covering 69 practice characteristics was returned by 98% of the 153 practices. “Internal audit” was carried out by 44% of training practices against 17% of non-training practices ($p < 0.001$). The overall conclusion of the study was that “differences between training and non-training practices were extensive”.

In a follow-up study (Baker, 1992) to try to explain the variation in standards multiple regression analysis was used to show that being a training practice related to a higher level of practice development. This level of development was partly dictated by a combination of professional factors such as the decision to become a training practice and the fact that a recognised practice inspection may be acting as a strong stimulus to such development. The suggestion from these two studies

was that innovation might be more developed in training practices which, in itself, would act as an external stimulus for introducing changes such as audit and improved record systems.

In order to determine this, a group of practices who had completed questionnaires in both 1982 and 1990 in the previous two studies were compared on 39 features of practice which could have resulted in change by gaining or losing each feature between the two surveys (Baker & Thompson, 1995). Results showed that 12 features were significantly more likely to have been gained or less likely to have been lost by training practices than non-training practices. One of these was audit ($p < 0.05$). A conclusion was that training practices were not only more developed than non-training practices but also more innovative and that the gap between the two types of practice had increased between the two surveys. The importance of innovation as a factor in commitment to teamwork and effective communication has been shown (West & Wallace, 1991). Without the commitment to teamwork audit participation taking into account factors such as list size or other structural issues may be less effective.

Lervy *et al* (1994) also found that being a training practice was associated with a higher uptake of audit activity. 90% of 63 practices in West Glamorgan were visited with practice audit data pre-recorded on semi-structured questionnaires sent in advance of the visits to the practices. As part of the agreed protocol for the visit "discussion of any completed audits or the opportunity of converting data collection into audit" took place. It is unclear from the paper whether actual data were verified. This is important as the authors claim that "all training practices had undertaken audit compared with 63% of non-training practices." For the purpose of the visit audit was defined as "any formal evaluation of performance in any aspect of practice which has resulted in a change in the future performance of that activity" based on a definition by Lyons and Gumpert, (1990). The paper did

acknowledge that although all training practices were confirmed to be conducting audit only one quarter of practices in West Glamorgan were accredited to train which was much lower than the proportions found in other studies such as the 44% previously referred to by Baker.

Nearly all training practices in the west of Scotland are computerised (98%), with nearly 90% of them using a common software application (GPass). Although appearing to be advantageous it was designed as an administration system, as its name suggests, and is mainly used for the three Rs - registration, repeat prescribing and recall (Milne *et al*, 1991). Having a computer was one of the characteristics associated with undertaking audit (Chambers & Bowyer, 1993), possibly by allowing a higher level of practice organisation and access to data useful for audit. This had also been found in Lervy's survey (1994) where visits confirmed that training practices and those which had developed their record systems were more likely to be involved in audit activity. Although not specifically mentioned the computer systems used in England and Wales are known to be more sophisticated in handling clinical data than Gpass.

One half of training practices were also involved in undergraduate teaching. In a survey of undergraduate departments of general practice in the United Kingdom and Eire (Spencer, 1992) only 11 of 27 responders provided formal teaching about audit. Among the reasons given for the shortage of teaching was a lack of expertise and knowledge of the subject among the staff. Opportunities for collaboration between undergraduate students and trainers were therefore unlikely to be generally productive although one study (Campion *et al*, 1992) involving a University department (Liverpool) and general practitioners in the Liverpool Medical Audit Advisory Group showed that joint planning between students and practitioners resulted in a highly motivating experience for both groups with change occurring through audit in a clinical setting and the learning needs of both

undergraduates and established practitioners being met through application of an audit project.

The high numbers of audits claimed to have been started and completed by both trainers and trainees in the west of Scotland are at variance with the evidence presented at practice visits. This is in keeping with many of the surveys previously cited and is probably explained by the reliance of unsubstantiated questionnaires augmented by a confusion in terminology as few of the questionnaires were followed up by a visit to the practices. Of those which were, such as Lervy's (1994), the raw data were not seen and discussion was based on semi-structured questionnaires previously completed by the practice and sent to the visiting group in advance of the visit. The importance of actual verification to determine the quality and quantity of audit activity is therefore emphasised.

A large proportion of the stated audit activity in training practices in the west of Scotland related to chronic diseases such as diabetes and asthma, both of which were being actively encouraged - and paid for - by Government initiatives to increase health promotion activity. By contrast, patient-centred audit of patient satisfaction, appointment systems and waiting times were less popular despite patient groups and the Government preferring to see an increase in this type of activity.

Difford (1990) described data which could be considered essential for the use of audit in general practice. He placed importance on the available sources of data, the use of population denominators for providing accurate data - hence the importance of an accurate age/sex register and organised case records - and the increasing importance of computers for clinical recording of data. The reality, however, was shown by Webb *et al* (1991) in a survey carried out on behalf of Leeds Local Medical Committee in 1990. The results showed that some doctors

had experience in setting up basic information systems but that much data collection was poorly focused. In the responses from 317 general practitioners data collection on chronic diseases did not feature in the top 10 subjects strongly suggesting that the Government contract of 1990 had a significant influence on practices' choices of data collection. Despite this an association between preferred choice of audit and a positive effect on income rather than clinical performance was shown to be less important by Chambers *et al* (1996) with 50% of 601 responding general practitioners disagreeing or strongly disagreeing with that sentiment.

Trainers' and trainees' attitudes to audit are positive. The source of the attitude statements was the discussions held at their trainers' groups consisting of both trainers and trainees and represented their own judgements on what was pertinent to audit for them. Although generally in the moderate range the correlation coefficients were positive for trainers and trainees. The correlation, however, between audit as a method of assessing doctors with other attitude statements was generally lower than those involving the other four statements. This differences may have reflected the imposition of the new Contract with its heavy reliance on medical audit in a culture where the tools for the job could not be assured. The use of audit to assess doctors was the only statement which nearly achieved significance at the 5% level in the difference between trainers and trainees indicating that trainees, with a median score of 3, were rather less convinced of the role of audit in assessing doctors than the trainers, with a median score of 4. One plausible reason for this was the background of a summative assessment working group in the west of Scotland where trainees who were due to be assessed would understandably be more wary than their teachers.

Investigations into the attitudes of general practitioners to audit increased sharply in the early 1990s. Prior to this one of the few published surveys on attitudes to

audit was carried out jointly by the Local Medical Committee and the Royal College of General Practitioners in the Doncaster area in 1981 (Waters *et al*, 1983). 129 local doctors were told that the exercise was to give them "a taste of audit". In order to achieve this they were asked to participate in a practice activity analysis exercise of psychotropic drugs over a designated 14 day period and return their record locally or direct to the Research Unit in Birmingham. A follow-up questionnaire was sent to the doctors after the exercise should have been completed and enquired into whether they had taken part or not and to answer nine attitudes statements pertaining to audit. The survey into attitudes was therefore focused around a piece of work, participation in which was low (28%) with the follow-up questionnaire being answered by 28 of the 36 doctors (78%) who had completed the audit and 25 doctors (38%) of those who had not. The overall conclusion was of strong opposition to outside control of audit with some respondents having doubts about control from within the profession. A response from 8 doctors of the 24 who had been previous local trainees was noted and the authors suggested that this indicated that implications for the future of audit were bleak and at variance with the optimism expressed by Stevens (1977) in his Butterworth Essay.

In 1987 a postal survey (Baker & Green, 1990) enquiring into opinions of quality assurance and the methods used to teach this subject was sent to all course organisers in general practice in Scotland, England and Wales and to all family practice residency directors in the USA. Due to the potential for confusion with the terminology the respondents were told that quality assurance was assumed to include medical audit, performance review, peer review and utilisation review. Despite the difficulties inherent in this type of study involving two different countries with their different systems for quality assurance there was a response from 74% of the 334 course organisers in the UK and 67% of the 381 residency directors in the USA after one mailing only. A good response may have been

enhanced by providing specific examples for most of the terms used in the questionnaire. Although the opinions of course organisers may not necessarily reflect those of their trainers and trainees they are likely to be influential. With only 39% of the opinion that quality assurance ensured high standards of care in general practice it is possible that this might affect attitudes in training practices. This compared with 53% in response from the USA where quality assurance is more of an accepted activity. More positively, however, was the fact that 80% of course organisers thought that the subject of quality assurance was very important or essential in the training of general practice trainees with a virtually identical proportion of respondents from the USA.

In the Leeds survey (Webb *et al*, 1991) attitudes to medical audit as a way of improving the quality of care given to patients were positive although perceived to be expensive in time and opportunity cost.

Participation was found to be more likely if attitudes towards audit were more positive. The importance of a positive attitude to audit was shown in an audit of Vitamin B12 in Leicestershire (Baker *et al*, 1995). The topic was chosen for its simplicity implying that inability to take part in the audit would probably be generalisable for any other type of audit. The decision to participate was associated with discussion on the audit having taken place in the practice and with a less positive attitude to the Medical Audit Advisory Group who were there to help them, possibly a consequence of the confidence with their own discussions within the practice. The main reasons for withdrawal from the audit or failure to take part at all were lack of time or resources similar to those highlighted in training practices in the west of Scotland.

Pringle *et al* (1994) used six pairs of audit statements to assess attitudes to audit. Of 323 GPs in Stockport and Derbyshire in late 1991 and early 1992, 69% of the

66% who responded strongly or moderately agreed about medical audit as an effective means of improving quality of medical care with a similar percentage (65%) moderately or strongly disagreeing that medical audit was a waste of time. Again, doctors who had experience in medical audit had more positive attitudes towards audit.

Chambers *et al* (1996) sent a questionnaire to 870 general practitioners in Staffordshire in 1992 enquiring into the attitudes to medical audit and any associations between their attitudes and their personal characteristics. Of the 69% who responded, 86% felt that audit was time consuming, 71% that ongoing training and education were needed, 68% felt that doctors were being compelled to do audit and 65% felt that extra resources should be provided for audit. Again, the authors found that those doctors who had the most experience of audit were most positive about it. They concluded, however, that general practitioners were far from convinced about the value of audit.

SUMMARY

The numbers of trainers and trainees purporting to have started collecting data, to have been involved in standard setting and to have implemented change are at variance with the evidence presented at training practice re-accreditation visits. The importance of data verification is emphasised. What data are being collected tend to be focused on chronic diseases such as diabetes and asthma which have been promoted by the Government and are time consuming. Trainers' and trainees' attitudes to audit are positive with trainees' attitude to its use in assessment being more wary than their trainers, possibly reflecting a background increase in assessment activity. Time and resources are the two main perceived difficulties by both trainers and trainees with audit. The level of organisation, development and innovation seen in training practices, in part as a result of the

stimulus of a re-accreditation visit, would appear to put them at a significant advantage over non-training practices in their potential ability to deliver audit.

Audit carried out in such training practices is based on a familiarity with the concept of a single audit project which has been part of the training culture for some years. More research on the role of an audit project may offer an opportunity to explore in more depth the teaching and successful implementation of audit in training practices.

Table 1a - Frequency distribution of areas of care chosen for audit by trainers

(n = 131)

Data collection (n = 118)		Standard setting (n = 73)		Implemented change (n = 71)	
Diabetes	(33)	Asthma	(8)	Diabetes	(9)
Asthma	(25)	Diabetes	(8)	Hypertension	(6)
Hypertension	(17)	Hypertension	(7)	Asthma	(5)
Prescribing	(12)	Prescribing	(5)	Appointments	(5)
Epilepsy	(9)	Benzodiazepines	(4)	Cytology	(4)
Benzodiazepines	(9)	Appointments	(3)	Benzodiazepines	(3)
Thyroid	(7)	Thyroid	(3)	Immunisation	(3)
Consultation lengths	(7)	HRT	(2)	Thyroid	(3)
Patient satisfaction	(6)	Antagonists	(1)	Patient satisfaction	(2)
Referrals	(6)	Consultation lengths	(1)	Referrals	(2)
Appointments	(5)	Family planning	(1)	Prescribing	(2)
Disease management	(4)	Immunisation	(1)	Repeat prescriptions	(2)
Immunisation	(4)	Oral contraceptive	(1)	Surgery/On-call visits	(2)
Surgery/On-call visits	(4)	Cholesterol	(1)	Antagonists	(1)
Laboratories	(4)	Cytology	(1)	Consultation rates	(1)
HRT	(3)	Epilepsy	(1)	Drug abuse	(1)
Antagonists	(2)	Repeat prescriptions	(1)	Psychiatric morbidity	(1)
Drug abuse	(2)			Epilepsy	(1)
Digoxin	(2)			Laboratory requests	(1)
Family planning	(2)			Practice management	(1)
Oral contraceptive	(2)				
X-ray referrals	(2)				
Coronary	(2)				
Hypothyroidism	(2)				
Income and expenditure	(2)				
Records	(2)				
Antenatal care	(1)				
Angina	(1)				
Chicken pox	(1)				
Income of activities	(1)				
Psychiatric morbidity	(1)				
B12	(1)				
Cholesterol	(1)				
Diuretics	(1)				
Practice management	(1)				
Rubella	(1)				
Well women clinic	(1)				

Table 1b - Frequency distribution of areas of care chosen to audit by trainees

(n = 104)

Data collection (n = 54)		Standard setting (n = 24)		Implemented change (n = 13)	
Asthma	(6)	Asthma	(3)	Appointments	(3)
Appointments	(4)	Hypertension	(3)	Diabetes	(3)
Antagonists	(3)	HRT	(2)	HRT	(2)
Hypertension	(3)	Appointments	(1)	Contraceptive pill	(2)
Digoxin	(2)			Night calls	(1)
HRT	(2)			Laboratories	(1)
Waiting times	(2)				
Cervical cytology	(2)				
Epilepsy	(2)				
Patient satisfaction	(2)				
Breast and bottle feeding	(1)				
Consultations	(1)				
Cardiac assessment	(1)				
Contraceptive pill	(1)				
Gynaecology	(1)				
Hypothyroidism	(1)				
Out of hours visits	(1)				
Rhesus status	(1)				
Arthritis	(1)				
Benzodiazepines	(1)				
Cholesterol	(1)				
Diuretics	(1)				
Geographical spread	(1)				
Laboratories	(1)				
Thyroid	(1)				
Wart clinic	(1)				

Table 2 - Attitudes to audit (expressed as a percentage)

(a) Trainers (n = 131)	Strongly Agree				Strongly Disagree	
Opinions were asked of the following statements:	1	2	3	4	5	NR
Audit is a useful way of assessing work	43	44	10	1	11	1
Audit is not an appropriate use of resources	3	5	14	47	24	2
Audit may be used to assess doctors	10	40	32	9	5	3
Audit is not an important use of time	2	10	15	49	21	4
Audit improves patient care	27	47	18	4	1	4

(b) Trainees (n = 104)	Strongly Agree				Strongly Disagree	
Opinions were asked of the following statements:	1	2	3	4	5	NR
Audit is a useful way of assessing work	37	57	4	2	0	0
Audit is not an appropriate use of resources	1	10	14	55	20	0
Audit may be used to assess doctors	9	35	28	16	12	0
Audit is not an important use of time	0	10	16	56	18	0
Audit improves patient care	15	59	17	7	1	1

NR = no response

Table 3 - Rank correlation (Spearman's) among 'attitude' statements in relation to audit

a) Trainers

	An appropriate use of resources	May be used to assess doctors	An important use of time	Improves patient care
Useful in assessing work	0.46	0.23	0.34	0.50
An appropriate use of resources		0.18	0.52	0.43
May be used to assess doctors			0.22	0.12
An important use of time				0.41

b) Trainees

	An appropriate use of resources	May be used to assess doctors	An important use of time	Improves patient care
Useful in assessing work	0.38	0.24	0.43	0.45
An appropriate use of resources		0.14	0.61	0.28
May be used to assess doctors			0.22	0.28
An important use of time				0.36

Table 4 - Trainer - trainee comparison of responses to attitudinal statements towards audit

Statement	p-value ¹
A useful way of assessing work	0.80
Appropriate use of resources	0.64
May be used to assess doctors	0.06
Appropriate use of time	0.95
Improves patient care	0.10

¹ Mann - Whitney test of no difference in median values between trainers and trainees.

Table 5 - Difficulties with audit

Trainers		Trainees	
Time	(49)	Time	(37)
Resources	(12)	Resources	(5)
Motivation	(7)	Setting standards	(4)
Money	(7)	Money	(3)
Cooperation from partners	(5)	Motivation	(3)
Used by Government	(4)	Data collection	(3)
Outcomes	(4)	Training/Skill	(2)
Knowledge/Experience	(4)	Implementing change	(2)
Training/Skill	(3)	Attitudes	(2)
Setting standards	(2)	May be used for assessment	(2)
Agreeing standards	(2)	Co-operation from partners	(1)
Completing cycle	(2)	Experience	(1)
Ideas for audit projects	(1)	Completing a cycle	(1)
Data collection	(1)		
Implementing change	(1)		

CHAPTER 4

THE INTRODUCTION AND EDUCATIONAL IMPACT OF AN AUDIT PROJECT

From August 1992 all trainees in the west of Scotland were required to submit an audit project of their choice during their training year.

Two studies were carried out in 1993 and 1994 to examine:

- a) Trainees' experiences in the undertaking of an audit project (1993).
- b) Trainees' perceptions of the educational impact of the audit project both on themselves and the training practice (1994).

a) Trainees' experiences in the undertaking of an audit project

Method

104 trainees in the west of Scotland who finished their training in 1993 were sent a questionnaire exploring:

- whether completing the audit project had been useful as a method for carrying out future audits. In particular they were asked whether it had led to an increased knowledge of disease process, an increase knowledge of therapeutics, an increased knowledge of practice administration and the perceived relevance of the project to patient care.

- estimations of the time taken for background reading and literature searches, discussion on choosing criteria and setting appropriate standards, preparation for and collection of data including searching for and retrieving case records and discussion of changes to be implemented.
- whether the trainee had chosen his/her own topic to audit.

Analysis was carried out to compare the proportions of trainees who had chosen or not chosen their topic with those who had collected two sets of data using the chi squared test.

Results

Responses were received from 103 trainees (99%).

Respondents' ratings of the perceived usefulness of the audit project are shown in Table 1. The audit project was strongly felt by 87 trainees (85%) to be a useful method for carrying out future audits and 70 trainees (68%) felt that the project had been relevant to patient care. Smaller numbers of trainees felt that the project had been useful in enhancing knowledge of practice administration, disease process and therapeutics.

The length of time trainees spent on the various stages of their audit project is shown in Table 2. Data collection took the longest time, with 23% of trainees estimating to have spent more than 24 hours on this activity.

The majority of trainees completed their audit project in less than one working week.

Of the 47 trainees who chose the topic for their audit project, 33 (70%) collected two sets of data. Of the 56 trainees who had their topic chosen for them, 16 (29%) collected two sets of data. (chi squared 17.8, 1df, $p < 0.001$).

Summary

Although based on retrospective estimates of time taken this study showed that an audit project was perfectly feasible in the time limit of one year in general practice. Although lack of time had previously been cited as a difficulty with carrying out audit it might be possible to reduce the time taken by utilising practice staff more effectively in collecting relevant data. The perceived usefulness of a project as a method of presenting an audit was acknowledged as was its perceived relevance to patient care. The ability to choose the topic to audit was more likely to result in a change in practice being evaluated.

b) Educational impact of the introduction of an audit project

Method

In May 1994 all 117 trainees in the west of Scotland were sent a postal questionnaire with a second follow-up to non-responders two weeks later.

The questions being asked were taken from a series of informal discussions earlier in 1994 with 11 trainee groups in the region concerning their opinions (positive and negative) on the submission of a compulsory audit project during their training year.

Three particular issues were raised from the discussions of concern to the trainees: their inexperience and lack of confidence with audit, the level of support available from the practice and their ignorance of their predecessor's audit project.

A series of questions was drawn up with the trainees based on their concerns.

- **Perceived benefit by the trainee in undertaking an audit project.**

Was this your first practical experience of carrying out an audit project?

As part of your training how would you assess your confidence in introducing change in your next practice?

- **Practice support for the audit project.**

From the audit chosen in which areas would you have liked more help and from whom?

Have you attended a formal practice meeting? If so, how often?

Was the audit project discussed? If not, why not?

- **Confirmation of last year's trainee audit project.**

Have you seen last year's trainee audit project?

If so, has it been repeated?

If not, why not?

The questions were piloted on 10 trainees and the format adjusted slightly on their advice with more emphasis being placed on the frequency of practice meetings.

Results

104 replies were received giving a response rate of 89%.

- **Perceived benefit by the trainee in undertaking an audit project.**

For 82 trainees (79%) this was their first ever practical experience of audit.

As a result of carrying out the audit 88 trainees (85%) felt more confident about the possibility of introducing change in their next practice after completing their audit project with none feeling less confident.

- **Practice support for the audit project.**

More than half of the trainees would have liked more protected time to carry out their audit project (Table 3). This was more than twice any other area where more help was required.

More help from the trainer was deemed appropriate by 35 trainees (34%) and from the partners in the practice by 24 trainees (23%). There was much more satisfaction with the amount of help given by reception staff, with only 5 trainees (5%) feeling they required more.

91 trainees (88%) attended a formal practice meeting at some time in their training year. 70 trainees (87%) stated that the audit project was rarely or never discussed. 30 trainees (29%) attended four or fewer practice meetings during the training year. The 11 trainees who attended no practice meetings stated that they were "not encouraged to attend" .

- **Confirmation of last year's trainee audit project.**

All the trainees should have had access to the previous trainee's audit project. However, only 35 trainees (34%) confirmed that they had actually seen it. Of these, 12 claimed to be continuing the audit cycle by repeating the project to evaluate the changes implemented. The reasons given for not continuing with the previous project are given in Table 4.

DISCUSSION

Most courses at all levels of study in Colleges and Universities have incorporated some form of project work. Project work has been championed earlier last century by Dewey (1916) and Kilpatrick (1918) in an attempt to overcome the "passivity

through rote learning". It is increasingly common in open-learning.

Project-based learning can be defined as an activity in which the learner develops an understanding of a topic or issue through involvement in an actual (or simulated) real-life problem (Morgan, 1982). Inherent in this is the degree of responsibility for designing one's own learning activities. Dressel and Thompson (1973) argued that to develop a student's ability to carry out independent study alone or with peers should be an important goal within education. Projects are usually designed to test discipline - specific skills, possibly in addition to transferable academic skills (Henry, 1994).

In 1978 Sir George Pickering expressed the view that doctors would benefit their education more from the chance to study and reflect on their work in contrast to the "tragic over-emphasis on the recall of factual knowledge". He concluded that "Young doctors need to think more critically about their clinical practice".

Fraser's definition of a project, adapted from Chamber's dictionary, is "a scheme of something to be done arising from a speculative imagination" (Fraser, 1982). This has a certain resonance with Warnock's description of education as "the development of the imagination, not only to see what is but to guess what might be" (Warnock, 1973). Projects as an organised component of vocational training began in 1976 with the creation of the Syntex Award Scheme (RCGP, 1985d), which was piloted through the Department of General Practice at the University of Exeter. Despite such incentives, however, Mukherji (1979) reported that while 20% of trainees showed initial interest only 10% completed a project. An editorial (1979) on the subject in the same journal described the conditions required to encourage more participation by trainees in projects

- the project should be useful and relevant to practice work.
- the exercise must be valued by the trainee.
- there is a need for trainer and peer support.
- it must be possible to complete in the time available.

Suggestions for justifying the encouragement of projects was set out by Gray (1985). They were:

- an alternative to the term "research" because of the negative connotations associated with the latter by many trainees.
- the need for protected time.
- encouraging improved record and information systems in training practices in order to support project activities.
- active trainer support.

Fraser (1982) attempted to clarify understanding between projects, audit and research. He described projects as "essentially educational exercises designed to familiarise participants with a systematic approach to the collection and interpretation of data". Audit was described as "the process of looking critically at our professional activities with a view to improving doctor performance and patient care" while research was "concerned with the acquisition of new information to add to the corporate body of knowledge".

Despite this Gambrill (1985) pointed out that the terms “project”, “audit” and “research” are often use synonymously, “thus compounding the existing confusion”. Projects were therefore viewed by many trainers and trainees as a form of research training. Howie (1985) cautioned about the potential negative effects of projects for trainees who embarked on ambitious data collection in attempting to defuse problems incapable of being answered within a reasonable timescale. He hoped, however, that integration of project work into vocational training might lead to general practitioners taking a more positive view of research in the future.

Buckley (1995) described the elements of a project as:

- selecting an area of enquiry which is of relevance and importance to general practice.
- defining a question which is capable of being answered.
- planning an appropriate method of enquiry, negotiating with others on the collection of information.
- analysing data.
- drawing sensible and appropriate inferences from the findings.
- presenting the findings in appropriate ways to appropriate audiences.

Audit became a synonym for a project as the latter was accepted by the JCPTGP as evidence for the former, with neither being defined for training practices.

The audit project was described by Baker (1990a) as one of four types of medical audit in use in general practice along with information systems, peer review and external audit. He described the audit projects as "a method full of potential which is yet to be realised. The new proposals for medical audit must foster this kind of audit or risk losing an important way of improving care".

The results of these two small studies suggests that there is room for optimism for the audit project's role in offering a strong start for possible future audit. Similar positive sentiments are described by Neville and Sowerby (1988) and Spencer *et al* (1989). Although the former were compulsory and the latter optional both reported that medical students and their teachers described the projects as a useful component of the curriculum.

Protected time was still deficient in most examples cited. Kratky (1977) described the importance of his trainer allocating at least three hours per week of protected time for him to devote to his project. He described this technique as "stimulating" enabling him to complete his audit of diabetics in just eight weeks on an introductory attachment to a practice.

Support for deciding criteria, negotiating standards and implementing change as part of an audit project were perceived to be under-resourced. Since this comes mainly from the doctors in the practice it can be assumed that formal teaching of audit method is minimal. Wilson (1993) showed that this parallels hospital training, where few consultants assessed the effectiveness of their teaching and feedback to their juniors in anything more than a rudimentary manner.

Lack of support was mentioned in informal discussion with trainee groups and may reflect trainers' and their partners' lack of confidence in these areas. By comparison, reception staff appeared very helpful, presumably in extracting and filing away case records.

The practice meeting is an obvious place to discuss an audit project with the practice team but the figures suggest that not all trainees have access to such a meeting and that even when they do their audit project is rarely discussed. In an editorial about this research Biggs (1995) described it as presenting "a challenge to medical educators everywhere in ensuring the understanding and support of both trainees and trainers in the educational application of audit". Rademocher (1993) showed that doctors in training grades performed better than their more senior colleagues when taking part in audit and were more responsive to policy change. With so few trainees being encouraged to follow up their predecessor's audit project the opportunity to evaluate change is lost.

SUMMARY

An audit project is having an encouraging effect on the confidence of trainees for effecting future potential change. For most of these trainees general practice training is their first exposure to audit. Trainers and the practice team require to ensure that quality time is devoted to discussing audit and protected time allowed for carrying it out.

It would then be possible for an audit project to provide a format for defining "principles of medical audit" as laid out in part of the JCPTGP criterion.

In addition a system of assessment may offer a means of discriminating between those trainees who are able to demonstrate their competence in understanding those principles from those who do not.

Table 1 – Trainees' ratings of the usefulness of the audit project*(n=103)*

Usefulness of audit project	Rating of project usefulness (expressed as a percentage)				
	1 (poor)	2	3	4	5 (excellent)
In enhancing clinical knowledge of disease process	19	24	27	25	5
In enhancing knowledge of therapeutics	24	26	31	15	4
In enhancing knowledge of practice administration	6	13	38	36	7
As a method for future audits	0	4	11	60	25
As relevant to patient care	2	5	25	44	24

Table 2 - Trainees' estimates of the amount of time spent on the different stages of their audit project

(n=101)

Time spent on	% of respondents reporting length of time taken (in hours)				
	<1	1-3	4-7	8-23	24+
Background reading/literature search	18	40	24	13	6
Discussion of criteria/standard setting	51	40	9	0	0
Preparation for data collection	6	33	61	0	0
Data collection (<i>n = 102</i>)	2	19	25	31	23
Discussion of results/suggests for improvement	39	51	10	0	0

Table 3 - Areas of the audit project where trainees felt more help was required

(n=104)

Protected time	55
Negotiating standards	27
Implementing suggested changes	26
Deciding criteria	25
Choice of project	17
Preparation and planning	16
Use of staff	15
Writing-up project	10
Interpretation of data	9
Evaluating change	8

Table 4 - Reasons given by trainees for not continuing the previous year's audit project

(n=104)

Not seen project	65
Have seen project but never discussed	23
Not interested	5
Too time consuming	4
Not beneficial to practice	3
Poor results	3
Too early to repeat	2
Senior partner interested but trainer was not	1
No answer (blank)	14

CHAPTER 5

AUDIT AND THE ASSESSMENT OF COMPETENCE

Background

The 1990 NHS Act and the preceding White Paper (Department of Health, 1989a) made all health service providers responsible for carrying out clinical audit. The 1987 (Department of Health, 1987) and 1989 Acts together formed part of the Government's legislation defining the NHS reforms started by the 1987 White Paper.

In a paper revisiting the RCGP's Quality Initiative Donald Irvine (1990) described the profession's response to the "1990 Contract" as appearing "to leave a vacuum".

There was conflicting evidence as to whether the vocational training certificate issued by the Joint Committee on Postgraduate Training for General Practice provided sufficient proof of a general practitioner trainee's competence to practise as an unsupervised principal in general practice. Responding to the concern that fewer than 1% of trainees were refused certificates at the end of their training the chairmen of the Joint Committee, the General Medical Services Committee and the Royal College of General Practitioners stated (Irvine *et al*, 1990) that the issue of certification should be determined by "a competent system of assessment" and that a national standard for entry into general practice should be considered. A doctor would therefore have to reach an acceptable standard of competence by the end of their training.

The JCPTGP (1992b) decided therefore that, by the end of their training, a doctor should have:

- adequate knowledge.
- adequate problem solving skills.
- adequate clinical competence.
- adequate consulting skills.
- adequate skills in producing a written report of practical work in general practice.
- adequate performance skills, attitudes and knowledge.

The word “adequate” was not defined. Gray (*personal communication, 1992*) added that “any method devised for assessing trainees summatively should be based on explicit criteria and accepted for publication in a peer reviewed journal.

When the JCPTGP published its policy document on summative assessment (JCPTGP, 1993) there was considerable debate about the need for a broad range of options to represent “practical work”. Examples given were: literature reviews, business plans or a piece of research being carried out during the hospital component of vocational training (*Toby J, personal communication, 1994*).

A written report revealed the ability to communicate an idea or concept which might promote change. Trainees are exposed to many examples of written reports of practical work during the training year. As advocates for their patients many written reports may have crucial implications. Appropriately written referral letters and legal reports are two examples. The urgency with which they are dealt may depend on the manner in which they were written. A badly prepared or poorly written report was therefore deemed to be a demonstration of competence below the standard acceptable of a practising general practitioner. The argument

for including a broad range of material in the report of practical work was therefore persuasive.

In 1991 the Committee on General Practice in the west of Scotland appointed a working party to investigate the possibility of developing a programme of summative assessment for general practice trainees. Its conclusion (Campbell *et al*, 1993) was that the assessment process should take account of the following:

- the trainer's assessment should carry weight.
- there should be an objective external contribution.
- clinical competence must be directly assessed.
- performance throughout the trainee year should count in the assessment.
- a 100% pass rate should be possible.
- the procedure must be feasible.

As a means of achieving the above attributes while taking account of the Joint Committee's criteria it was decided that the assessment process should be divided into four components which would combine to produce a balanced overall assessment:

- A test of factual knowledge and problem-solving skills.
- An audit project.
- Evaluation of consulting skills.
- The trainer's overall assessment.

The inclusion of an audit project within a summative assessment system

An audit project was to be the format of a written report of practical work which would be submitted by a trainee in the west of Scotland. Audit was seen as a

method of identifying learning needs (Coles, 1990) and could be useful in problem-solving (Baker, 1990b). Data collection, awareness of relevant literature, negotiated team work and discussion of change all involved a certain amount of action on the part of the trainee and could therefore be justified as practical work. Committing the audit to a written format helped to focus on the need for change where such change has been clearly identified. The choice of subject for the audit project tested whether the doctor was able to balance the importance of the topic with the feasibility of investigating the quality of care in the time available. In essence, the trainee was demonstrating his or her ability to monitor and, if required, to improve the quality of care being provided, described by the GMC (1995) as "a basic principle of good practice". It was strongly argued that failure to demonstrate an example of this principle was accepted as being important enough to require a period of extra training to ensure that audit method was understood as judged by the successful submission of an audit project.

Assessing an audit project

The definition of audit used in "Working for Patients" (Department of Health, 1989b) was a reminder that measuring the quality of patient care would not be straightforward.

"The systematic critical analysis of the quality of medical care, including the procedures used for the diagnosis and treatment, the use of resources and the resulting outcome and quality of life for the patient."

In its response the Standing Committee on Postgraduate Medical Education (SCPME, 1989) was optimistic in the perceived link between audit development and its "potentially educational outcomes".

The two themes in the many definitions of audit around at this time were the move from medical to clinical audit and the need to “identify opportunities for improvements and provide a mechanism for realising them” (Shaw & Costain, 1989) rather than merely identifying areas for improvement producing what was termed “orphan data” in the USA (Nelson, 1976).

The aim was to ensure that a submitted audit project would allow a trainee to demonstrate the reason for choosing the particular audit and that the method was appropriate.

From August 1992 all trainees in the west of Scotland starting the practice component of their vocational training had been asked to submit an audit project.

The advice given to practices (Irvine & Irvine, 1991) was that the audit project should include:

- subject of audit.
- background.
- reason for the audit.
- methods.
- results.
- changes recommended.
- repeat audit, if possible.

As the implementation of change was deemed to be outwith the trainee’s control, recommendations for change were accepted.

Method – designing an assessment instrument

Between August 1993 and January 1994 a series of six small group discussions was held between the author and the trainers and trainees in the west of Scotland to establish their views on a more structured framework than Irvine's criteria in guiding the completion of an audit project. A number of basic questions were agreed upon. These were:

- is the topic relevant to routine general practice?
- is the present practice situation defined?
- are the criteria chosen relevant to the choice of audit?
- have standards been set?
- have changes been recommended which are realistic?

Between February and July 1994 a further series of workshops was held to explore the issues of teaching and assessing audit in more depth. 18 participants with an interest in and practical experience of audit in the west of Scotland - seven general practitioners who were either trainers or partners of trainers, seven general practitioners who were audit facilitators, three doctors (one non-medical) from the University of Glasgow and one practice nurse who was also an audit facilitator - agreed on the following objectives for the workshops:

- to develop independent criteria based on the elements of a criterion audit which would help to define part of the JCPTGP criterion on "principles of medical audit".
- to produce an instrument which could be used to assess whether the criteria were present in a submitted audit project.

- to develop an appropriate system for dealing with the resubmission of a substandard audit project.
- to provide appropriate instruction and training for assessors of an audit project.
- to ensure that the GMC's principle of good practice in relation to clinical audit was addressed in the successful submission of an audit project.

Crombie and Davies (1992) described the need to answer three questions in any data collection exercise:

- Why was it done?
- How was it done?
- What did it find?

This was used as the starting point in identifying the features which characterised a satisfactory audit.

In order to introduce trainees to considering the management of change a fourth question was added:

- What next?

The four questions were applied to a number of trainee audit projects and a list of elements required to answer the questions was created:

- Why was the audit done?
 - Potential for change.
 - Clear reason for change.
- How was the audit done?
 - Relevance of criteria.
 - Standard setting.
 - Preparation and planning.
 - Negotiation with relevant team members.
 - Timescale.
 - Staff involvement.
 - Money required.
- What was found?
 - Interpretation of presented data.
 - Data presented are relevant to the criteria.
- What next?
 - System for change described.
 - Further change proposed where appropriate.
 - Second collection of data compared.

The list of 14 elements considered to be part of an audit project was sent to all 155 trainers in the west of Scotland. They were asked to assess each element as essential or desirable as part of their trainee's audit project, taking into account their current confidence in and experience of audit method.

The trainers' opinions were rank ordered and applied to a further series of trainee audit projects. It was found that to avoid unnecessary duplication of material in the audit projects 10 of the elements were sufficient. These elements received agreement from 82% or more of the trainers. This confirmed the trainers' and

trainees' feedback in discussions that collecting a second set of data for an audit project was outwith their experience and confidence.

The final assessment instrument was drawn up using the four questions which required to be answered by addressing five dichotomous criteria which were chosen from the 10 remaining elements agreed to be essential or desirable by more than 80% of the responding trainers. It is shown in Figure 1.

An accompanying instruction sheet was prepared for assessors indicating that all five criteria had to be present for the audit project to pass (Figure 2). In addition a structured proforma for the audit project was drawn up with each criterion being assessed placed at the head of each page to minimise any confusion by the trainee, trainer or assessor about what was being assessed.

Method - Designing an assessment system

Due to the inexperience of markers, trainers and trainees with the new assessment instrument, the following objectives in defining an assessment system were agreed upon:

- the purpose was to identify a project which did not describe an audit as defined by the five criteria.
- there would be a referral system for a further assessment for decisions where there was still some doubt.
- there would be feedback given to a trainee where criteria were judged to be inadequately addressed with an ability to resubmit the project.

- if any further teaching was required for a resubmitted project it should not be given by any of the project's assessors.
- further development of the system would be encouraged as the confidence and experience of trainers, trainees and assessors.

Marking experience was gained using a range of previously submitted trainee audit projects. As all previous trainees had submitted projects a wide range of quality was available. Projects were marked in groups of three with maximum opportunity given for discussion on disagreements. The assessors were instructed to read the complete project, to make a judgement on each of the five criteria and to discuss differences with their colleagues. Projects were then passed to different groups and the results compared. A total of 20 assessors took part with the rule that if they were in doubt about any criterion the project would be referred for a further assessment by two more senior colleagues with more experience of audit for a final decision.

22 audit projects were subsequently chosen from the 95 submitted by trainees in 1993. One project where there was unanimous agreement at the workshops (Project Q) was designated a "refer" project as a marker. The other 21 projects were chosen at random and all 22 projects were marked independently over four weeks by the 20 assessors who had completed their previous training, 15 of whom had been involved in the original development of the assessment instrument.

In order to decide the optimum number of assessors from the requirement for a referral system a set of rules was devised which would balance the probability of a poor project being referred and the probability of a good project not being referred. These were calculated from the spread of referrals as a result of the marking exercise (Figure 3).

Analysis

Reliability of agreement between pairs of assessors was calculated using the kappa statistic (Fleiss, 1981). A kappa value signifies a measure of agreement between assessors beyond that which would have been reached by chance alone. If an assessor makes pass or refer decisions in all cases there is no discrimination beyond chance and the kappa value cannot be calculated.

This can have a disproportionate affect on the average of the total kappa values and consequently the proportions of agreements between pairs of assessors were also calculated.

Furthermore, interpreting the kappa statistic is controversial with no absolute definition available. Landis and Koch (1977) suggested the following as a guide:

Value of kappa	Strength of agreement
<0.20	Poor
0.21 – 0.40	Fair
0.41 – 0.60	Moderate
0.61 – 0.80	Good
0.81 – 1.00	Very good

Data were entered into an Excel spreadsheet with 1 for a pass and 0 for a refer judgement.

Average kappa values and proportions of agreements were calculated using SAS version 6.03.

Method - Testing the assessment system

All 102 trainees starting their training in August 1994 and due to complete it in July 1995 were given a pack containing advice on the marking instrument, the proforma and the assessment system of referral and the possibility of a resubmission should that be required. They were advised that this was a pilot exercise and would have no implications for the trainee. They were advised to submit their project no later than three months before the date for completing their training in case of the need for a resubmission.

RESULTS

Designing an instrument and system of assessment

135 trainers (87%) responded to the list of 14 elements which they considered essential or desirable in a trainee's audit project (Table 1).

The five criteria based on 10 of these elements were used by the 20 assessors to mark the 22 audit projects which were returned within four weeks. The results were analysed to decide the optimum number of markers and requirements for referral.

A total of seven or more "refers" for a project from the 20 assessors was taken as the arbitrary cut-off between projects which should have been referred and those which should not. Figure 3 shows the spread of referrals and provides justification for the selection of this cut-off point.

The rules supporting the developing of an assessment strategy were:

- the six worst projects are defined as those with seven or more “refers” from the 20 assessors. These projects should have been referred.
- the sixteen projects with five or fewer “refers” from the 20 assessors should not be referred.

The probability of a poor project being referred - the sensitivity - and the probability of a good project not being referred - the specificity - were calculated for the number of combinations of assessors to achieve a reasonable balance with the emphasis on sensitivity to ensure that a poor project was clearly identified.

System	Sensitivity	Specificity
Use one assessor	45%	92%
Use two assessors, refer if both they refer	20%	99%
Use two assessors, refer if at least one says refer	70%	86%
Use three assessors, refer if all three say refer	10%	100%
Use three assessors, refer if at least two say refer	42%	97%
Use three assessors, refer if at least one says refer	83%	80%

Three assessors were found to give the optimum balance of sensitivity and specificity with a referral for further assessment if one or more of the three felt that a project should be referred.

The results of the marking exercise and summary statistics for average kappa values and average proportions of agreement between assessors are shown in Table 2.

The average kappa values were low, in part due to the disproportionate effect of some assessors such as assessor number 4 who passed all projects.

Some assessors had negative kappa values indicating that their judgements were less reliable than chance. The interquartile range for agreements between pairs, however, was 74% to 80% indicating the importance of combining assessors in judging the projects given their relative inexperience.

Testing an assessment system

All 102 trainees submitted an audit project. Each project was marked by three independent assessors. Projects were sent out in batches of four and assessors were given seven working days to mark them. 77 projects (76%) were passed by all three assessors and the registrars notified of a pass result. 25 projects (24%) were referred by one or more assessors and were further assessed by two more experienced colleagues using the same five criteria. The mean delay in a registrar receiving notification of a pass result or a request to resubmit their project was 5.3 days.

15 projects referred by first level were passed at second level with 10 projects (10%) being sent back to the registrar highlighting the criteria where the assessors felt the registrar had not demonstrated sufficient understanding. An invitation to resubmit the project or, if preferred, to submit a new one was given. All the resubmissions were revisions of the original projects.

The resubmitted projects were then marked by three different assessors independently. A total of 98 projects (96%) achieved a pass with two projects still judged to be below standard and a further two were not resubmitted. There was no disagreement on second level decisions to pass or refer.

All 102 projects were subsequently marked by the author. None which had been passed by the three first level assessors was judged to be a missed referral. The complete results are shown in Figure 4.

A further marking exercise consisting of a 10% sample of the 102 submissions was carried out. One project was chosen which had been referred by all three first level and both second level assessors. This was added to nine randomly chosen

projects with the total representing the 90% which had passed and 10% which had been referred for resubmission.

Only two of the 17 assessors who took part in the marking exercise passed the “poor” project but they also passed all other projects. Even if they had been in the same marking group of three assessors the project would still have been referred to the next level as any third assessor would have referred it. Two other assessors were found to have referred more than 50% of their projects. Feedback to and discussion with the assessors on the leniency or stringency of their assessments compared with their colleagues, were found to be helpful in recognising the reasons for the variations in their judgements.

The results of the marking exercise and the summary statistics of the average kappa values and average proportions of agreements for the assessments are shown in Table 3.

Both the average kappa values and the interquartile range of proportions of agreement (77% to 85%) were higher than in the previous exercise.

Each project took between 10 and 15 minutes to mark with each assessment costed at £12.

DISCUSSION

There were a variety of frameworks available for evaluating audits which were used for a wide range of purposes. The questions raised by Crombie and Davies (1992) provided the basis for the initial development of the assessment instrument for training practices. Their framework, however, focused on general issues in the assessment of audits without necessarily considering more specific issues such as

the setting in which the audit took place. Bhopal's and Thomson's model (1991) was designed to evaluate papers published in audit and covered a level of detail of methodology which would have rendered it unusable in the context of educational assessment.

Walshe and Tomalin (1993) looked specifically at how to evaluate the funding for audit projects by considering five issues related to specific objectives being met and resources being used. The Oxford method for "auditing audits" (Derry *et al*, 1991) was developed during the first year of the existence of medical audit advisory groups in England and Wales and was used to assess the number and completeness of audits taking place during that year. Limitations in the method, however, were quickly identified such as reliance on a form of words such as full, partial and potential audits which caused confusion, did not take account of the range of topics chosen by a practice and gave no indication of team involvement such that a practice could be relatively passive with one enthusiast being responsible for all the audits. It was, however, widely used beyond Oxfordshire and provided feedback to both practices and family health service authorities (Lawrence *et al*, 1994). The Kirklees Method (1993) for classifying audits was a ten point scale which could similarly document audit activity over a number of practices in a region. Problems, however, were recognised (Lister *et al*, 1998) when the scale was used outwith its originating region. This resulted in the need to adapt it for local use thus questioning its generalisability.

The decision therefore to develop and test an assessment instrument, the construction of which considered the users - both trainers and trainees - had important implications for its validity, an issue which will be addressed later in the discussion.

A key issue in the assessment of competence which Neufeld *et al* (1985) described is clarity of purpose. The JCPTGP had identified a necessity for a trainee at the end of his or her training to demonstrate a satisfactory written report of practical work the format of which, in the west of Scotland, was to be an audit project. This concurred with the GMC (1995) standards of competence, one of which was to monitor and improve the quality of health care - to take part in regular and systematic clinical audit. Competence in this area therefore would be assessed summatively but, taking account of the lack of experience with audit found in the training practices, would include a formative element of feedback and an opportunity to resubmit if required. Ultimate failure to demonstrate such competence by submitting a satisfactory audit project was deemed unlikely.

A further important issue in the assessment of competence is blueprinting (Dauphinee, 1994). This process ensures that test content is carefully planned against learning objectives. Conceptual frameworks against which to plan assessments are essential. Blueprinting contributes to the content validity of the assessment process. The five criteria constituting the assessment instrument using consensus development techniques in their construction give some justification to both face and content validity of the instrument. They also provided a definition for the "principles of medical audit" as required of training practices by the JCPTGP.

The importance of setting an appropriate standard in advance of the assessment is another issue in the assessment of competence. In the context of vocational training the certificate of accreditation issued by the JCPTGP is restrictive; by law a doctor cannot practise as a principal in the NHS without it. Thus the setting of a minimum standard defines a level of competence or performance anything less than which is unacceptable (Irvine, 1993). This is in contrast to a good standard at

practice level such as the MRCGP and at the level of excellence such as that accorded by Fellowship by Assessment (RCGP, 1990).

Minimum competency testing for students was at its most popular in the late 1970s, particularly in the USA, in response to worries about falling standards (Wolf, 1995). Berk (1986) described common features of minimum competence testing programmes such as an emphasis on the acquisition of minimum skills or competence, an explicit performance standard for pass/fail decisions so that the competent can be separated from the incompetent with the test results being used to make high - stakes decisions about individual students. Legal challenge in the USA was met by an accumulation of large bodies of evidence to prove the content validity of testing programmes. A benefit has been the setting up of procedures to obtain the consensus of all those involved in the testing such that the definition of the curriculum and requirements for minimum competence have credibility and meaning (Rust & Golomboks, 1989). Norm referencing - comparing one trainee with others - is unacceptable for a situation where there needs to be a clear distinction between those who have demonstrated competence and those who have not. Such standards are set by criterion referencing, a method first introduced by Glaser (1963). The main issue in the development of a criterion-referenced test is "to obtain rigorous and precise domain specifications to maximise the interpretability of an individual's domain score" (Berk, 1980). For minimum competence testing the criteria are dichotomous - they are either present or not. Judging audit projects against explicit criteria should increase the reliability of the assessments. Criterion-referenced measurements allow comparison against well-recognised descriptions representing the attributes or areas of competence being assessed (Mulholland & Tombleson, 1990). Assuming valid criteria against which to judge, the main influence on assessors' decisions is the variation in an individual marker's behaviour, either stringency with a tendency to over-refer or leniency with a tendency to under-refer. Campbell *et al* (1995) showed that

feeding back variations in assessors' judgements allowed "calibration" to ameliorate extremes of assessor behaviour. Preston - Whyte *et al* (1993) showed that further training maintained reliability. Van der Vleuten *et al* (1989) described one solution to the problem of examiner variation as the use of standardised instructions for those assessing and those being assessed and the use of materials which complement the assessment instrument such as the proforma on which audit projects were submitted.

Sensitivity and specificity have been described by the Institute of Medicine (1990) as two of the structural attributes used in the evaluation of quality measurement techniques for detecting deficient or inappropriate care. Although not conventional attributes of an assessment process in this country they were felt to be appropriate given the inexperience of all those with a role in the system - trainees, trainers and assessors. The ability to balance sensitivity and specificity given a variety of combinations of assessors allowed the optimum number to be chosen in developing a referral system should a potentially poor project be identified. The combination of three assessors maximised reliability as judged by the agreement between different pairings of assessors.

Van der Vleuten (1996) described what are now accepted as the desirable components of any process of assessment. The five attributes which require balancing in the design of an assessment system are validity, reliability, acceptability, feasibility and educational impact. The utility of an assessment process is a multiplicative function of the five variables each with its own weight. The implication is that each needs to be present to some extent with the absence of any one implying that the utility of the process is zero. He states that "a reliable, valid and feasible test will have a short life if it's accepted by no-one".

Carmines and Zeller (1979) described measurement as a process involving both theoretical and empirical considerations. The latter focuses on the observable response whilst the former concentrates on the underlying unobservable concept represented by the response. Measurement focuses on the relationship between the two. A strong relationship allows useful inferences to be drawn between the responses observed and the concepts being studied. The challenge for any instrument of measurement is to reduce the amount of error due to chance - random error - to a minimum (reliability) thus maximising the extent to which any series of given indicators represents the concept being measured (validity). A valid instrument therefore measures what it is intended to measure. Invalidity is introduced if other variables or constructs, often unmeasured, interfere with this process - non-random error. According to Cronbach (1971) "one validates not a test but an interpretation of data arising from a specified procedure".

It should be noted, however, that much emphasis has been placed on the quantifiable aspects of testing, ones which originated in the psychometrics movement originating in the earlier years of the last century. More recent literature in educational assessment has questioned their suitability when applied to more complex issues such as competence. Moss (1992) was critical of relying too heavily on quantifiable approaches suggesting that they tended to overshadow the effect the assessment has on learning and, in addition, encouraging a possible political use of test results. The social consequences of assessment use - known increasingly as consequential validity which takes account of concerns about negative consequences and expectations about positive consequences - are now considered central to concepts of validity, more broad ranging than, but including, educational impact.

In considering validity issues in the assessment of competence Miller (1990) has described a conceptual model based on a pyramid with the base representing

factual knowledge or “knows” as tested by an MCQ, with two further layers of clinical context based tests (“knows how”), performance assessment (“shows how”) and finally the apex of the pyramid performance assessment (“does”). This represents the ultimate goal for a valid assessment of clinical competence, i.e. to test what the doctor actually does in his or her work - a doctor’s actual performance. An audit project can be considered as an example of performance assessment in real life.

The five criteria of which the assessment instrument is a construct are widely recognised as integral parts of the criterion audit process. The domain of content is limited by the trainers’ perceived level of confidence in teaching the criteria in the timescale available (realistically nine months). Therefore although implementation of change might be desirable as an end-point in the assessment process the trainers recognised that detailed proposals for change was, on balance, what was achievable given current experience and confidence in training practices. It is precisely because of the limitations imposed by the content domain that content validity is in itself an imprecise standard against which to assess the validity of an instrument. As mentioned previously construct validity measures unobservable qualities (the constructs) by requiring a range of methods to “build a case”. Important in this process is proving that other factors are not confounders. In the context of the audit project the questions posed in the structuring of the assessment instrument help to build the case for a trainee demonstrating his or her competence in critically analysing a piece of work described by the General Medical Council as being able to take part in clinical audit. There are obvious limitations with this approach given the shortcomings of using the assessment on only one submitted audit project. This is recognised by Swanson *et al* (1995). Trainees may not necessarily perform consistently if further audit projects were assessed. Successful demonstration of competence on one occasion however should increase the chances of a similar situation given similar circumstances.

Validity therefore is a trade-off between what is theoretically desirable and practically possible. Ebel (1961) refers to validity research as that which is "universally praised but the good works in its name are remarkably few".

Transparency is the key to acceptability. All involved in the assessment process need to know each others' roles and responsibilities. Both trainers and trainees were clearly aware of the five criteria to be assessed with the trainees having a proforma matched to the criteria which offered guidance with their project. This more structured approach has also been shown to reduce assessor variation (Van der Vleuten *et al*, 1989).

Good assessment is costly. Investing in good assessment, however, can be described as investing in teaching and learning. The feasibility of an assessment system is a measure of what *can* be done in a practical setting and is a major limiting factor when choosing assessment methods and in achieving consistently reliable results. The fairness of the system covers what *should* be done to ensure equal opportunity for all participating in the assessment system. The advantage of a summative assessment system is that all can pass.

With each project being marked three times during the screening process by relatively inexperienced assessors the opportunity for a reliable result was improved. The addition of two further assessments when a decision was in doubt introduced the idea of fairness to a trainee where further work was judged to be needed after five individual assessments. With a project taking between 10 and 15 minutes to mark such a number of assessments was possible, given competing pressures of service work and other training needs.

Educational impact is closely linked to consequential validity. Together with construct validity these two issues are increasingly considered by Gipps (1994) to be at the core of the assessment of clinical competence. As noted in a previous chapter the impact of two years of audit project submissions by trainees was felt by the majority to have increased their confidence in implementing change, despite the audit project being their first ever experience of audit. Such an example illustrates how assessment can be part of the learning process, helping to achieve educational objectives set out early in the training year. The importance of feedback of assessment results for substandard projects provided an example where assessment can drive learning through information imparted as part of the process. The separation of educational and assessment networks limits the risk of the assessment objectives overpowering learning objectives. Where a trainer is also an assessor potential for conflict of interest in certain situations can be avoided by ensuring that the two roles do not overlap for a particular trainee.

Subsequent analysis for construct validity

Between August 1994 and July 1997 333 criterion audit projects were assessed against the five criteria. Explanatory factor analysis was carried out on the results. The purpose of factor analysis in this context is to examine the co-variance relationship among variables. Two constructs were hypothesised: the reason justifying the choice of the audit and how the audit was executed. The analysis was carried out using SAS.

Normally the number of factors retained is determined using Kaiser's criterion (1960) which retains only those factors with an eigenvalue greater than 1. Where the number of variables is fewer than 20, however, this method tends to generate a low number of factors.

Two factors were identified. *Criteria, preparation and planning, interpretation of data and detailed proposals for change* loaded heavily on factor one while *reason for choice* loaded heavily on factor two. Interpreting this suggests that the method of audit (how the audit was done), corresponding to factor one, and the reason for the audit (why the subject for audit was chosen) represented on factor two tended to be assessed differently from each other.

Applying Promax rotation to the original factor pattern can provide more "useful" factor patterns but the factor structure and reference structure, in addition to the factor patterns, require consideration. Varimax rotation was broadly similar. The effect is to remove from each factor the effect of the other factor. The overall impression of the Promax and Varimax rotated factors was similar thus appearing to confirm that reason for choice is being assessed differently from the other four criteria.

The proportion of the total sample variance is obtained with reference to the eigenvalue which indicates its relative importance. The results are shown in Table 4.

The two factors retained in the analysis together explained 69% of the total sample variance.

SUMMARY

The JCPTGP required a trainee to demonstrate his or her ability to prepare a written report of practical work as part of a summative process for the assessment of competence. The choice of such work in the west of Scotland was an audit project. The GMC has described the ability to take part in clinical audit as a basic principle of good practice in its standards of competence. Five objective criteria

were developed to assess this competence and a system developed to allow pass/fail decisions to be made at a standard of minimum competence with the possibility of feedback and further resubmission should it be deemed appropriate. Assessment objectives matched educational objectives. The five accepted attributes of an assessment system were balanced to provide a model of competence assessment which would discriminate between those trainees who were able to demonstrate their understanding of certain principles of clinical audit, as set out in an audit project, from those who could not.

Figure 1**SUMMATIVE ASSESSMENT - AUDIT****MARKING SCHEDULE**

Please tick the box provided if the criterion for answering each question is/are present.

<u>QUESTION</u>	<u>CRITERION</u>	<u>CRITERION PRESENT</u>
<u>Why was the audit done ?</u>	<u>Reason For Choice</u> Should be clearly defined and reflected in the title. Should include potential for change.	<input type="checkbox"/>
<u>How was the audit done ?</u>	<u>Criteria Chosen</u> Should be relevant to the subject of the audit. Should be justified e.g. literature.	<input type="checkbox"/>
	<u>Preparation and Planning</u> Should show appropriate teamwork and methodology in carrying out the audit. If standards are set they should be appropriate and justified.	<input type="checkbox"/>
<u>What was found ?</u>	<u>Interpretation of Data</u> Should use relevant data to allow appropriate conclusions to be drawn.	<input type="checkbox"/>
<u>What next ?</u>	<u>Detailed Proposals for Change</u> Should show explicit details of proposed changes.	<input type="checkbox"/>

A satisfactory trainee audit report should include all 5 criteria to pass.

Please enter your opinion in the box provided.

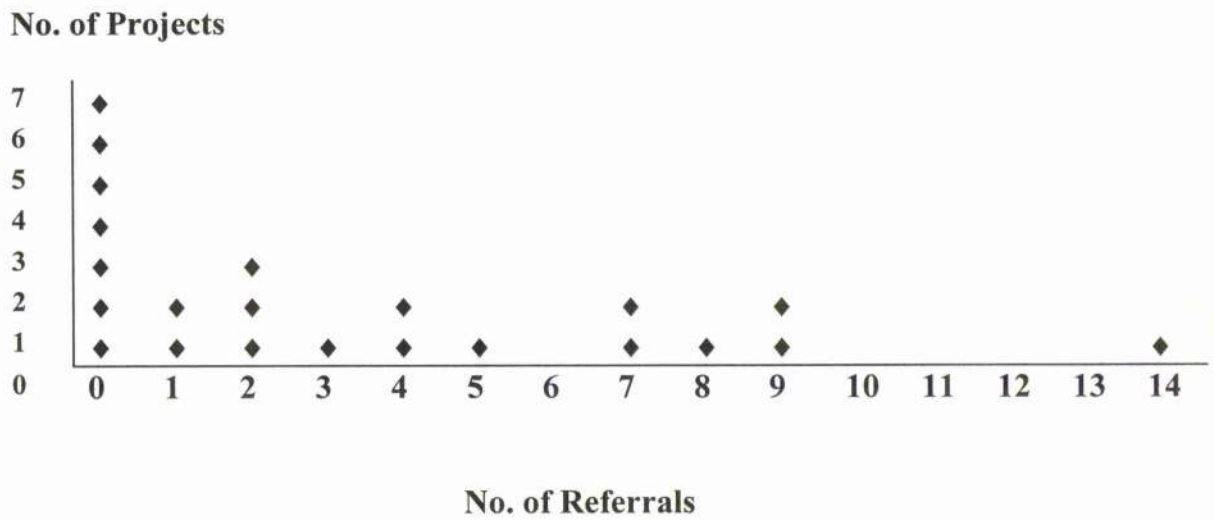
Pass ☐
Refer ☐

If Refer, please comment on your reasons.

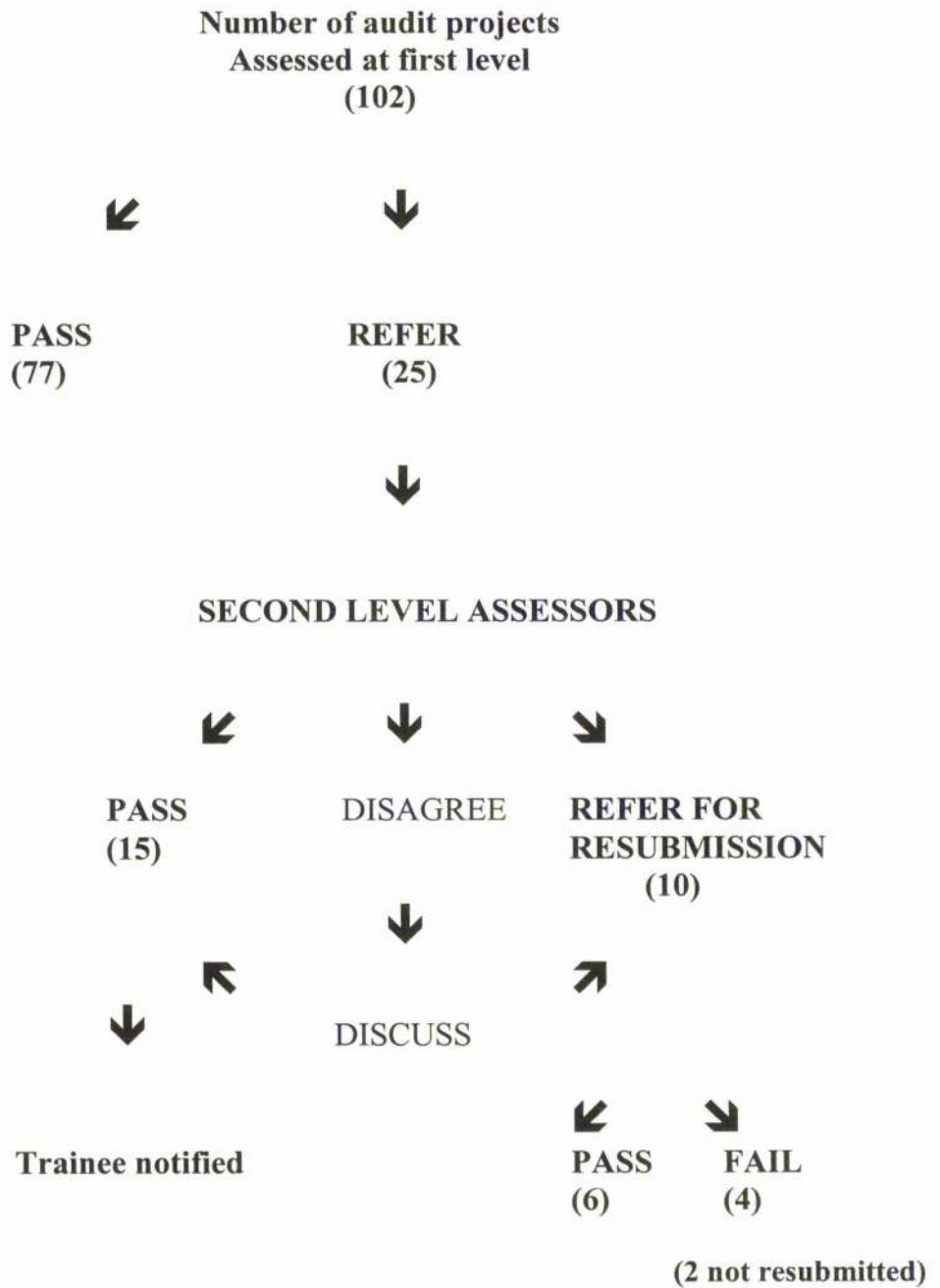
Figure 2***Instructions for audit assessors***

Please use the marking schedule to give your opinion of the general practitioner registrar's audit project. It is crucial that the whole project is read before marking begins. The criteria to be used for marking are in bold print. The statements in less bold print should act as a guide when making your judgement. There are five criteria to be marked for an audit project to pass assessment, all five criteria must be present. Please comment at any stage of the process but specifically if the registrar audit project is being referred.

Figure 3 - Development of strategy for 'referral' of an audit project



**Figure 4 - Assessment of 102 audit projects:
August 1994 – July 1995**



**Table 1 - Trainers' opinions on essential or desirable elements
of a registrar's audit project**

Element of audit project	% of 135 trainers considering element essential/desirable
Relevance of criteria chosen	97.0
Standard setting	96.3
Preparation and planning of project	96.3
Interpretation of presented data	94.1
Potential for change	91.9
System for change described	87.4
Data presented: relevance to criteria	86.7
Negotiation with relevant team members	86.7
Reason for choice of project	85.9
Further change proposed where appropriate	82.2
Second collection of data prepared	74.8
Timescale to complete project	74.1
Staff involvement	68.9
Money required	45.9

Table 2 - Audit assessors' marking exercise (1) on agreement between pairs

Assessors	Projects																			
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	P	P	P	P	P	P	P	P	P	P	R	P	P	P	P	P	P	P	P	P
2	P	P	P	P	P	P	P	P	R	P	R	P	P	R	P	P	R	P	P	R
3	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	R	P	P
4	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P
5	P	P	P	P	P	P	P	P	R	P	P	P	P	P	P	P	R	R	P	P
6	R	R	R	P	P	P	P	P	R	P	P	P	P	R	P	P	R	P	P	P
7	P	P	R	P	P	P	P	P	R	P	P	P	P	P	P	P	R	P	P	P
8	P	P	R	P	P	P	P	P	P	P	P	P	P	R	P	R	R	P	R	P
9	P	P	P	P	P	P	P	P	P	R	P	P	P	P	P	P	R	P	P	P
10	P	P	R	P	P	P	P	P	P	P	P	P	P	P	P	P	R	P	R	P
11	P	P	R	P	P	P	P	P	R	P	R	R	P	R	P	R	R	P	P	P
12	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	R	P	R	P
13	P	P	P	P	P	P	P	P	R	P	R	P	P	R	P	P	R	P	R	P
14	R	P	R	P	P	P	P	R	P	R	R	R	P	R	P	P	R	P	R	P
15	R	P	R	P	P	P	P	R	P	P	R	P	P	P	P	P	R	P	P	P
16	P	P	P	P	R	P	P	P	R	P	P	P	P	P	P	P	P	P	P	P
17	P	P	P	P	P	P	P	P	P	R	P	R	P	P	P	P	P	P	P	P
18	P	P	P	P	P	P	P	P	P	P	P	R	P	R	P	P	R	P	R	P
19	P	P	P	P	P	P	P	R	R	P	R	P	P	P	P	P	R	P	P	P
20	P	P	P	P	P	P	P	R	R	R	R	R	P	R	P	P	P	P	P	P

P = Pass decision

R = Refer decision

	Mean	Median	S.D.	Minimum	Maximum	Q1	Q3
Average kappa values	0.35	0.43	0.19	0	0.54	0.17	0.54
Average proportions of agreements	0.76	0.81	0.12	0.45	0.85	0.77	0.85

Table 3 - Audit assessors' marking exercise (2) on agreement between pairs

Projects										
Assessors	A	B	C	D	E	F	G	H	I	J
1	P	R	P	P	P	P	P	P	R	P
2	P	R	P	P	P	P	P	R	P	P
3	P	R	P	P	P	P	R	P	P	P
4	P	R	R	P	R	R	R	P	R	R
5	P	R	P	P	P	P	P	P	P	P
6	P	R	P	P	P	P	P	P	P	P
7	P	R	P	P	R	P	P	P	R	R
8	P	R	P	R	R	R	R	R	P	P
9	P	R	P	P	P	P	R	P	P	P
10	P	P	P	P	P	P	P	P	P	P
11	P	R	P	P	P	P	P	P	P	P
12	P	R	P	P	P	P	P	P	R	P
13	P	R	P	P	P	P	P	P	P	P
14	P	P	P	P	P	P	P	P	P	P
15	P	R	P	P	P	P	R	P	R	P
16	P	R	P	P	P	P	P	P	P	P
17	P	R	P	P	P	P	P	P	P	P

P = Pass decision

R = Refer decision

	Mean	Median	S.D.	Minimum	Maximum	Q1	Q3
Average kappa values	0.18	0.22	0.12	-0.05	0.35	0.08	0.27
Average proportions of agreements	0.77	0.78	0.04	0.65	0.82	0.74	0.80

**Table 4 – Factor analysis of assessments of 333 audit projects
(1994 – 1997)**

Factor	Eigen value	Proportion of variance
1	2.64	0.53
2	0.79	0.16
3	0.64	0.13
4	0.50	0.10
5	0.44	0.08

CHAPTER 6

TEACHING AND LEARNING: ISSUES IN KNOWLEDGE, METHODS AND CONFIDENCE

The submission of an audit project by all trainees, now referred to as registrars, in the west of Scotland promoted confidence in their ability to effect change. An assessment instrument and system was beginning to identify registrars who had a problem demonstrating an understanding of five criteria which constituted an adequate criterion audit project after two attempts with feedback. The lack of audits presented at training re-accreditation visits had suggested that knowledge of and confidence with audit method might be a problem for some trainers. The role model of the trainer had already been emphasised in prioritising a rigorous approach to the principles of good practice (RCGP, 1985a). Baker and Thompson (1995) identified training practices as being more innovative. This follow up study confirmed that training practices were consistently ahead of their non-training colleagues in implementing audit. It would seem reasonable to infer from this that trainers had a reasonable knowledge of audit method but this had never been formally tested.

Two studies were undertaken:

- the first tested whether the trainers could recognise deficiencies in the criteria of registrar audit projects using the same assessment instrument which they had helped to validate and which was used to assess their registrars' audit projects.

- the second considered the methods used for teaching audit to the registrars and assessed whether the submission of an audit project had increased the trainers' and the registrars' confidence with audit method.

1. Trainers' assessments of audit criteria

Method

In July 1994 158 trainers in the west of Scotland were sent five registrar audit projects. All five projects had been referred by first level and second level for further work, a maximum of five possible referrals. In the first project *reasons for choice* had been judged as substandard by three of the five assessors, *detailed proposals for change* and *preparation and planning* had been judged to be substandard by four of the five assessors and *criteria* and *interpretation of data* were judged to be substandard by all five assessors. The trainers were asked to assess all five projects with no qualifying advice given. They used the same assessment schedule and were given two weeks for the exercise. Two follow up requests were made.

Analysis

Three outcomes were identified:

- trainers' against assessors' judgements on identifying the five substandard criteria.
- for each trainer the number of substandard criteria correctly identified and the total number of criteria judged to be substandard were compared.
- the trainers' overall judgements of pass/refer were compared with the assessors' judgements.

Results

114 trainers (72%) completed the marking exercise.

All five criteria were judged to be satisfactory by the majority of trainers (Table 1).

Three trainers correctly identified all five substandard criteria. In doing so they judged a mean of 8.3 criteria as being substandard. In addition the greater the number of criteria they correctly identified the greater the mean number of other criteria they judged as substandard (Table 2).

All five projects were passed by the majority of trainers (Table 3).

2. *Methods and implications of teaching criterion audit*

A follow up study was carried out in 1996. The study set out to explore the reasons for the difficulty experienced by the registrars who had not submitted a satisfactory audit project. An opportunity was then taken to explore further the methods used by the trainers to teach criterion audit method and the impact of this on their registrars. More specifically, the impact of submitting an audit project as part of a summative assessment process on trainers' and registrars' confidence with their respective teaching and learning was explored.

Method

In 1995 four registrars (4%) and in 1996 five registrars (5%) failed to achieve a satisfactory audit project despite an opportunity to resubmit. One to one interviews with all nine registrars were undertaken by the author to establish the reasons for the "failure". The reasons given (with the appropriate number of registrars) were:

- trainer's lack of understanding of audit (8)
- audit considered unimportant in the practice (5)
- the relatively passive role of partners in the practice (5)
- no discussion of audit at a practice meeting (4)
- registrar's lack of confidence with advice being given (3)
- difficulties with discussing change (2)
- insufficient constructive criticism (1)

The trainer's influence in advising on preparation for and the final submission of a registrar's audit project was felt to be relevant in all cases and, for three of the registrars, had contradicted the registrar's own intuition.

In order to explore more positive influences on the teaching and learning of audit method the author interviewed five trainers picked at random and six registrars who had passed their audit project. The focus of the interviews was the teaching of audit and the role (if any) of the audit project for summative assessment. Both groups were advised that their responses would help to construct a questionnaire for all trainers and registrars to explore more widely the issues raised in the teaching of criterion audit method.

Issues raised by both trainers and registrars were:

- importance of confidence with understanding criterion audit method.
- methods of teaching criterion audit method.
- trainers' needs for teaching.

Issues raised by registrars only were:

- involvement in teaching from the wider primary care team.
- protected time for the audit project.
- the teaching of the use of information technology.

A questionnaire was constructed taking account of the above issues and laid out in such a way as to allow comparisons to be made between trainers and registrars where appropriate. Questions relating to issues of confidence were laid out in a five-point Likert scale from 1 (no confidence) to 5 (very confident).

The questions used are shown above Tables 4 to 6.

The questionnaire was sent to 151 trainers and 116 registrars in the west of Scotland in August 1996.

Analysis

Categorical data on trainers and trainees were compared using Fisher's exact test for any unspecified association in the tables. Where numbers in some of the cells were less than five the adjacent categories, e.g. scores of 1 and 2 on the Likert scale were combined. Analysis was carried out using SAS.

RESULTS

Replics were received from 129 trainers (85%) and 115 registrars (99%).

There was no statistical difference between trainers and trainees in their confidence with criterion audit method. Both groups were reasonably positive in their confidence (Table 4).

Similarly there was no difference in confidence in the teaching of and being taught criterion audit method. Both groups were slightly less confident with this than with their knowledge of audit method. (Table 5)

The submission of an audit project for summative assessment had a significant impact on trainees with 84% being more confident as a result and just over half (54%) of trainers also being more confident. Only a very small percentage of either trainees or trainers (6% for each) were less confident as a result of the submission. Statistically, the difference between trainers and trainees was highly significant at $p < 0.0001$ (Table 6a). Just under half of the trainers (46%) felt more confident with their teaching of criterion audit method following the submission of a summative assessment audit project with an additional 49% feeling that it made no difference (Table 6b).

Table 7 shows the difference in perceived needs between trainers and trainees with the various components of an audit project. There was virtually no difference between the two groups in terms of perceived needs with approximately one third to two thirds of either group feeling little need in any of the areas. The two areas where there appeared to be a statistically significant difference were with literature searches and guidance on interpretation of data where trainees did not appear to need help.

Table 8 shows the range of methods used for teaching criterion audit. Informal discussion such as at a practice meeting was the commonest method for both groups. Seven trainees (6%) claimed to have received no teaching at all.

The trainer was still doing most of the teaching followed by the partners in the practice. The practice nurse and practice manager appeared to play very little role (Table 9).

In addition 66 trainees (59%) felt they had sufficient protected time for their audit project and teaching in information technology was received by 49 trainees (43%).

DISCUSSION

The two studies outlined in this chapter were a consequence of the inability of a small number of registrars to submit a successful audit project after nearly one year in a training practice and feedback on their initial submission. The involvement of the trainers in the construction of the marking instrument should have enhanced their familiarity with its construct and, if anything, help with the marking exercise described.

Walshe (1995) described clinical audit as being founded on "an implicit assumption that health care professionals new what it was and how to do it". The registrars' experience and some evidence from the trainers' marking exercise suggest that this assumption was misplaced. It is likely therefore that many National Health Service clinicians have not had the necessary skills or training in clinical audit. This is in contrast to many other health care systems outwith the United Kingdom where there is evidence of large training programmes on quality improvement initiatives (Joss & Kogan, 1995).

Strict comparisons between different systems however is difficult. Baker and Green (1990) were cautious about drawing conclusions from a questionnaire sent to trainees in general (family) practice in the UK and the USA in 1987. Although a wide of variety of teaching methods was suggested, neither country appeared to promote the teaching of quality assurance as a priority, despite the longer history of quality assurance in training programmes in family practice in the USA. While recognising the difficulties of making firm conclusions from questionnaires the

authors suggested that the teaching of quality assurance to registrar deserved a higher priority in both the USA and the United Kingdom. The position of quality assurance in training curricula should be reviewed with the methods used being appropriate to the teaching of the subject.

The marking exercise carried out by the trainers appeared to confirm that their ability to teach criterion audit could be compromised by a failure to recognise key elements in the construction of an audit project. Some caution is required however in interpreting the results. Although only three trainers recognised the five criteria judged to be inadequate, it should be acknowledged that the "expert" assessors were not in total agreement themselves on three of the five criteria although the level of agreement was strong. The five projects were chosen from a total of 104 submissions, only 10% of which were judged to be inadequate, thus the sampling frame was relatively small.

The strong association between the number of criteria "correctly" identified as being inadequate and the mean number of total criteria identified as being inadequate suggest that stringent trainer marking behaviour rather than their ability to discriminate aspects of criterion audit method resulted in the "correct" criteria being identified. Thus the results of this exercise appeared to confirm some of the registrars' concerns about their trainers' understanding of audit method.

The high response rate from both studies suggests that both trainers and registrars were aware of the importance of the issue. The submission of a criterion audit for summative assessment had had a significant impact on the confidence of both trainers and registrars with criterion audit method. This offered further evidence for the consequential validity of the audit project as seen in the earlier study where registrars were more confident about introducing change as a result of their audit submission. It is possible that part of this confidence may be explained by the fact

that they had just finished their audit projects and the results would have been known to them.

Trainers themselves still carry out the bulk of teaching audit in the practice with non-medical staff contributing less than one quarter of teaching input. A variety of methods of teaching were used although 6% of registrars claimed to have had no teaching at all and protected time was still an issue for just under half of the registrars. Baker and Green showed a very high “exposure to in-practice quality assurance” (99%) compared with 78% in this study with formal tutorials constituting 76% and 65% respectively.

There is still an issue therefore about the failure of audit to achieve its educational potential. Kerrison *et al* (1993) suggested that lack of guidance for those being trained in audit, lack of training of the teachers, the relatively long audit cycle and relatively short attachments of those being trained and finally the lack of mechanism for feedback of audit findings all contributed to a reduced link between education and audit. Vimpany and Nixon (1995) confirmed Kerrison’s points in a qualitative study carried out between 1992 and 1994. They concluded that negative attitudes to audit would ensue in an environment where knowledge and skills were deficient and the educational benefits were blurred.

Experience of audit teaching in the undergraduate curriculum appeared to confirm the reasons for the low exposure to audit and possible increase in the registrars’ confidence after submitting an audit project during their general practice year. Spencer (1992) surveyed all 31 academic departments of general practice in the United Kingdom and Eire. 30 departments responded. Concerns were expressed about the lack of expertise and knowledge of audit among the staff. Also cited were lack of time in the curriculum and difficulties making the teaching relevant and interesting. Only 11 departments provided formal teaching about audit with a

further five intending to introduce it in the near future. "Project work" was the method of teaching in eight departments with seminars and lectures providing the remaining methods. The importance of establishing a culture of self-enquiry, whatever the method used, was acknowledged. Morrison and Sullivan (1993) evaluated 128 medical students (of 153) in terms of their knowledge and attitudes to audit. A case-note review of 10 diabetic patients was carried out by all students during their fourth clinical year attachment in general practice. They found the exercise relevant but a significant proportion found the data collection boring or very boring. Self reported knowledge of audit method showed a significant increase by the end of the evaluation.

A more in-depth survey of teaching audit to undergraduates was carried out by Spencer and Barton (1994) at Newcastle Upon Tyne Medical School. Despite the agreement of 75% of the senior clinical teachers who responded to the questionnaire that audit should be taught to undergraduates, only 17% actually provided such teaching. Of those respondents who were in favour of such teaching most justified it by the need for students to know about audit and quality with some feeling it would be better taught as a postgraduate subject.

Campion *et al* (1992) explored the sharing of an audit project between groups of medical students and the general practitioners in whose practices they worked. The majority of both medical students and general practitioners found the experience valuable with the study concluding that project work was an effective tool for motivating students to learn and led to change in the clinical setting in which it occurred. This appeared to concur with Peterson (1973) who asserted that "timing is an important factor in the acceptance of the review concept" and advocated that exposure to audit principles should be introduced as early as possible to the curriculum. The optimum time according to Barbaccia (1976) is when it has maximum relevance. With an increasing number of medical students

now having clinical experience earlier in their curriculum the relevance of audit should be introduced at an appropriately early time.

Patrick (1992) has described a relationship between knowledge and competence in skills. The skills and attributes required for self-audit and participation in peer review will include critical thinking, gaining access to medical literature, an appreciation of how groups and organisations function, change management and an awareness of personal limitations and an ability to seek help when necessary. In order to teach these skills medical teachers are required to be adequately trained and supported (General Medical Council, 1993). According to Lowry (1993) the training of such teachers requires a balanced programme of theory and practice if the necessary skills are to be acquired, developed and refined.

Calman and Downie (1988) described the importance of both education and training in the preparation of doctors for lifelong learning. They refer to Peters (1967) who greatly influenced the philosophy of education in the 1960s and 1970s and who stated that for any activity to be "educational" it should contain something worthwhile or valuable for its own sake, it should have wide cognitive perspective which can deepen one's understanding and those who are engaged in such activity must care about what they are doing. A registrar's evaluation of an aspect of the quality of care they are providing is, by its very nature, a worthwhile activity, the construct of which through an audit project encourages a deeper understanding of the process and, finally, by having such a project assessed will attach a sense of importance to a successful submission of such a project.

The results of these two studies suggest that registrars are taking full responsibility for achieving a series of goals through a variety of processes involving different methods of learning culminating in the production of a criterion audit project.

Some of their trainers are either passive in this process or possibly giving erroneous advice. Project-based learning for registrars is one strategy for developing self directed learning, the key elements of which according to Brookfield (1986) are that the learner takes the initiative for:

- diagnosing learning needs.
- formulating goals.
- identifying resources.
- implementing appropriate activities.
- evaluating outcomes.

Coles (1998) described self directed learning as an active process is to encourage a deeper approach to an active search for understanding as opposed to reproducing what has been learned. The quality of learning outcomes is largely determined by the approach taken to learning by the learners. Chastonay *et al* (1996) showed that self directed learning is at its most efficacious when the learning process is based on experience and the new knowledge and understanding gained can be integrated into the personal and professional context of the individual. The choice of subject for the audit project is the registrar's responsibility and the process which is then worked through within his or her practice encourages a range of skills such as preparation and planning and negotiating with other team members with whom they may have to work for the rest of their professional lives. The importance of the registrar choosing his/her subject to audit in terms of increasing the likelihood of implementing change has already been shown in the thesis.

The process for developing criteria and standards in order to enable judgements to be made about clinical care involves the use of behaviouristic principles of learning. A clear framework for action is produced based on a need to modify observable behaviour as described by Hilgard (1962). These principles have been

described as the antithesis of self direction and ignore many of the more reflective processes taken up in any other aspect of audit.

The need for critical thinking is exhibited by the knowledge-seeking and understanding-seeking learning styles described by Entwistle (1987). The testing of the reason for a particular choice of subject to audit and the choice of particular criteria and standards reflect the ability to explore deeper concepts and processes involved in learning.

Central to the whole audit process is the concept of reflective learning as described by Schon (1988). He described a process of modifying practice based on interpreting the results of complex problems in an iterative process. There are various levels of reflective learning with different levels of complexity but Kolb (1984) described a learning cycle based on observing current practice, reflection on the reasons for current practice, theorising about principles and standards and finally experimenting in new situations and returning to observing practice. This framework for learning builds on work on the assessment of previous learning and develops further into peer assessment and feedback as described by Falchikov and Boud (1989).

Finally, the importance of personal growth theories of learning is stressed by Roth (1990). Reflection upon experience encourages self confidence and initiative and a deeper understanding of one's self and others. The significant increase in confidence seen in two different studies in this thesis supports this theory and again emphasises the importance of consequential validity as part of the assessment process. The structural framework of an audit project is important in balancing the tension and anxiety which individual choice and responsibility can provoke (Stevens, 1990) and the notion of "the reasonable adventurer" (Heath, 1964) which encourages the more positive aspects of personal growth.

In conclusion therefore a variety of learning opportunities is exhibited in the production of an audit project. Despite the limitations discussed it would appear that trainers will require significant input to encourage an environment where registrars can learn the principles of audit as part of an overall reflective process. The introduction of the audit project has shown that learning can occur resulting in an increase in confidence and personal growth, a process which is increasingly seen as important when considering the validity of an assessment process. The need now is to create an environment where registrars can learn effectively and efficiently. The trainers would also benefit from such an environment in having a better understanding of how their registrars learn as described by Irby (1994).

The next phase of encouraging this change is the development of a programme of audit throughout the training practices using different techniques and covering different areas of care.

Table 1 – Trainers' versus assessors' judgements on identifying substandard criteria

Project No.	Criterion for assessment	Assessors' judgements of Criterion as substandard		Trainers' judgements of criterion as substandard		Left Blank
		(n=5)		(n=114)		
		Yes	No	Yes	No	
A	Reason for choice	2	3	89	21	4
B	Criteria chosen	0	5	87	24	3
C	Preparation and Planning	1	4	81	32	1
D	Interpretation of data	0	5	93	17	4
E	Detailed proposals for change	1	4	64	48	2

Table 2 – Trainers' marking behaviour when assessing criteria as substandard

No. of criteria to be identified correctly	No. of trainers correctly identifying criteria	No. of other criteria identified	Mean no. of times other criteria judged substandard
5	3	25	8.3
4	4	27	6.7
3	7	41	5.9
2	27	102	3.8
1	34	72	2.1
0	39	-	-

Table 3 – Trainers' versus assessors' overall judgements for each project

Project No.	Assessors' overall judgement		Trainers' overall Judgement		Left Blank
	Pass	Refer	Pass	Refer	
A	0	5	61	50	3
B	0	5	66	46	2
C	0	5	66	47	1
D	0	5	92	18	4
E	0	5	63	49	2

Table 4 - How confident are you with criterion audit method?**Actual numbers (%)**

	Less confident	Middle	More confident
Trainee (n=114)	5 (4%)	49 (43%)	60 (53%)
Trainer (n = 129)	9 (7%)	56 (43%)	64 (50%)

Fisher's Exact test, $p = 0.06$ **Table 5 - How confident are you with the teaching of criterion audit method?****Actual numbers (%)**

	Less confident	Middle	More confident
Trainee (n=113)	26 (23%)	48 (43%)	39 (34%)
Trainer (n = 129)	25 (20%)	58 (45%)	46 (35%)

Fisher's Exact test, $p = 0.70$ **Table 6(a) - Has your confidence with criterion audit method changed since the submission of a summative assessment audit project?****Actual numbers (%)**

	Less confident	Middle	More confident
Trainee (n=112)	6 (6%)	11 (10%)	95 (84%)
Trainer (n = 127)	8 (6%)	51 (40%)	68 (54%)

Fisher's Exact test, $p < 0.0001$

Table 6(b) - Has your confidence with teaching criterion audit method changed since the submission of a summative assessment audit project?
(n = 127)

Less confident	No difference	More confident
6 (5%)	62 (49%)	59 (46%)

Table 7 - Is more help needed with the teaching of:

	Trainer	Trainee	p value
Choosing an appropriate audit project?	20 (16%)	27 (24%)	0.12
Choosing criteria?	41 (33%)	35 (32%)	0.61
Advising on a literature search?	51 (41%)	22 (19%)	0.002
Setting appropriate standards?	40 (33%)	37 (32%)	0.03
Appropriate preparation and planning of and audit?	34 (28%)	27 (24%)	0.11
Interpretation of data?	49 (39%)	17 (15%)	0.008
Implementing change?	31 (25%)	23 (20%)	0.31

Table 8 - Did audit teaching involve:

	Trainer	Trainee
Formal tutorial?	84 (65%)	62 (54%)
Informal discussion? e.g. practice meeting	114 (88%)	83 (73%)
Recommending a book to read?	41 (32%)	31 (27%)
Demonstrating a completed audit?	100 (78%)	77 (67%)
No teaching carried out	-	7 (6%)

Table 9 - How much audit teaching input was there from:

	Least	Middle	Most	No Replies
Trainer? (n = 113)	24 (21%)	33 (29%)	56 (49%)	4
Any partner(s)? (n = 105)	46 (44%)	36 (34%)	23 (22%)	2
Practice nurse? (n = 105)	91 (87%)	8 (8%)	6 (5%)	12
Practice manager? (n = 103)	86 (83%)	11 (11%)	6 (6%)	14
Reception staff? (n = 107)	78 (73%)	18 (17%)	11 (10%)	10

CHAPTER 7

A STRATEGIC APPROACH TO THE INTEGRATION OF AUDIT INTO TRAINING PRACTICES

Verification and discussion about audit activity during training practice re-accreditation visits continued to confirm that audit was still dissociated from daily work with a negative message about the purpose of audit being transferred from trainer to registrar. The setting of standards and the use of multi-professional protocols were rare. One of the main recommendations of the Standing Committee on Postgraduate Medical Education for Training - that "medical audit skills should be seen as a high priority" - was not being implemented.

Issues still requiring to be addressed were:

- practice audits were project-based producing data rather than change.
- audits were unfocused with no connection to a wider audit plan for a practice.
- practices were still very sensitive about sharing their data.
- the teaching of audit was still largely based on informal discussions.
- the JCPTGP criterion for audit in training practices was not being implemented.

A wide-ranging debate among trainers and associate advisers took place within the region on how to respond to these issues and agreement was reached on the need to give a sense of direction for the development of audit in training practices.

Coincidentally a further incentive for a defined programme of audit activities was the publication by the Management Executive of the Scottish Office (1995) of "A

Strategic Framework for Clinical Audit in Scotland". Its aim was to "set out the direction of clinical audit over the medium term in all practices in Scotland". Two of its key strategic objectives were that:

- all general medical training practices are to have a rolling programme of audit in place for January 1997.
- two thirds of general medical practices are to have a rolling programme of audit in place by January 1998.

There was no definition of a "rolling programme" and no mechanism described for its verification.

A draft programme of audit activities was constructed by the author and distributed to all associate advisers responsible for the training groups throughout the west of Scotland. They were asked to discuss the programme with their trainers, to comment freely and, if necessary, to suggest alternatives. The programme had two aims:

- to provide training practices with a broad and practical range of audits to which registrars would be exposed.
- to quality assure the JCPTGP's criterion for audit in training practices.

The general principles of the programme were that:

- it should reflect the multi-dimensional nature of "quality", i.e. should contain quantitative and qualitative methods.

- it should address issues of structure, process and outcome of care.
- aspects of the programme should allow measurement of performance against explicit optimal standards for the region agreed by consensus.
- elements of the programme should concentrate on the use of the practice team including medical, nursing and receptionist colleagues with the practice manager being encouraged to have overall operational control of the programme.
- training practices were encouraged to work together in their districts to act as a resource for each other.
- progress with the programme would be assessed twice at three-yearly intervals as part of the rolling programme of re-accreditation visits. Each practice assessment would be carried out in advance of the visit by the associate advisers using a pre-agreed proforma.

The draft programme was discussed for six months by the trainers and associate advisers. No alternatives were forthcoming. A supporting manual with background evidence for the audits being carried out, sample data collection sheets for practices to use and an explanation of the audit methods and key issues around the chosen audits were sent to all the training practices in the region. The programme was implemented throughout the region on 1st January 1995.

The three broad areas covered by the audit programme were:

- organisation of appointment systems.

- chronic disease management.
- significant event analysis.

Organisation of appointment systems

Educational objective:

To demonstrate to a registrar the various components involved in organising an efficient and effective system for patient access to health care and to appreciate the difficulties in achieving this.

Review criteria:

- Non-urgent problems should be seen within three working days (standard, 90%).
- Patients should wait no longer than 20 minutes from their allotted appointment time (standard, 80%).
- Surgeries should start within five minutes of their allotted time and finish within 20 minutes of their allotted time (standards, 90% and 75%).
- Patient satisfaction with the surgery and the consultation should be assessed by an appropriate sample of the practice population (standard, one SSQ per practice and one CSQ per partner and the registrar).

Chronic disease management

Educational objective:

To demonstrate to a registrar the difficulties in managing the complexities of five chronic diseases in primary care - diabetes, asthma, epilepsy, hypertension and

rheumatoid arthritis - with an emphasis on the importance of good organisation and team work.

Review criteria:

For each disease there should be:

- a disease register updated within the previous six months.
- a method for focused data collection (flow charts) of “must do” criteria.
- a written protocol of care covering the members of the practice team.
- one completed audit cycle as chosen by the practice for each disease.

Significant Event Analysis

Educational objective:

To encourage creative and constructive reflection on events which happen in a practice unexpectedly. The registrar should appreciate the importance of trust and confidentiality.

Five significant events should be analysed and discussed with the registrar by the practice. Events should be chosen because:

- they are thought to be important in the life of the practice.
- they might offer some insight into the care provided by the practice.

- discussion should focus on specific reasons for actions and behaviour of an individual or failure of a system.

RESULTS

Data from 114 training practices in 1998 and 113 practices in 2001 were collated using Excel and analysed using SAS. Successful achievement of the various criteria were compared using Fisher's Exact Test. The differences in proportions achieving the criteria and standards between the two timescales with 95% confidence intervals were calculated.

No baseline data collection was possible due to both the work involved in carrying this out and the perceived threat which was transmitted from the trainers by the associate advisers. Previous practice visits had confirmed that few practices had any systems in place to monitor the areas of the programme.

The method inevitably gives an approximation of the overall change throughout the region as a whole. Over the timescale of the programme new practices were accredited and existing practices either stopped training or were de-accredited and this resulted in a difference in overall numbers over each three-year period. The analysis therefore was based on a collation of practices achieving the criteria and standards with differences in proportions compared.

Appointment Systems (Table 1)

There was an immediate increase in completion of the process audits with a smaller percentage achieving the desired standard.

There was a further increase for all process and outcome audits in the subsequent three years with more than 80% achieving the standards for appointment

availability, waiting times, start and finish times. More than 50% achieve the desired standard in each area.

The only area where statistical significance was reached was the increase in the proportion of practices reaching the 90% standard for starting surgeries within five minutes of time ($p=0.02$, 95% C.I. for difference in proportions = 3% to 28%).

Chronic Disease Management (Table 2)

Fewer than 50% of practices had disease registers for the chronic diseases by the end of the first three years. There was an increase in the proportion achieving this for all diseases by the end of the subsequent three years, with all but rheumatoid arthritis achieving statistical significance at the 5% level.

Flow charts for the collection of data were present in the records for diabetes and asthma in the greatest numbers of practices followed by epilepsy, hypertension and rheumatoid arthritis. Again, all proportions increased between data collections for all diseases with a range from 4% for epilepsy to 13% for hypertension with none achieving statistical significance.

The increase in proportions of practices having protocols of care ranged from 5% for rheumatoid arthritis to 9% for diabetes. The small increases may reflect the higher proportions after the first data collection for all diseases except epilepsy. None of the increases reached statistical significance.

By the end of both three-year periods fewer than 50% of practices had demonstrated at least one full cycle of change for any disease. The increases in proportions were modest from 2% for diabetes (which had the highest proportion after the first collection of data) to 10% for hypertension. None achieved statistical significance.

Significant Event Analysis (Table 3)

There was a highly significant increase in the proportion of practices achieving at least five significant event analyses (25%) in the second three-year period ($p < 0.001$, 95% C.I. 13% to 36%).

DISCUSSION

Evidence from previous research showed that both trainers and registrars had little practical experience of audit as a mechanism for improving quality of care. Registrars in particular had found the submission of an audit project for summative assessment a boost to their confidence and trainers had been shown as a consequence of this to be deficient in their knowledge and skills in identifying certain aspects of criterion audit method with inevitable consequences for their ability to teach it. The need to integrate audit into working practice was therefore important and an adequate timescale to allow this to take place had to be allowed.

The time for an appropriate culture change to take place for implementing quality programmes in large organisations has been shown to be at least five years and often longer (Joss *et al*, 1995). Two three-year periods were deemed appropriate for setting up a broad audit programme which would result in the implementation of the JCPTGP criterion for audit in all training practices in the west of Scotland.

The debate on the composition of the programme was rigorous. There was an inevitable tension between the more managerial style and a more developmental approach which might allow more innovative programmes to be implemented at district level. The former was not popular but options on alternative ideas had been sought and none had been submitted. It is likely that this reflected the relative inexperience of the trainers with few ideas to build upon.

The framework for the audit programme tried to reflect the multi-dimensional nature of quality. Three models were considered in the construction of the programme although the final choice of programme content addressed what Maxwell (1992) described as the need to avoid the search for quality becoming "too intellectual, purist and static".

A working group of the World Health Organisation (1983) described quality in a health service as having four dimensions: professional performance (technical quality), resource use (efficiency), risk management (the risk of injury or illness associated with the service provided) and patients' satisfaction with the service provided. Although mainly focused on technical aspects of care delivery it was broad enough to offer a range of options for training practices.

A further classification more familiar in its taxonomy is Donabedian's triad (1980) of structure (for example staff and resources), process (how resources are used) and outcome (the result of using resources). This triad is much more intuitive in that different aspects of quality can be defined in a more structured way. Finally, Maxwell (1984) described six dimensions which expanded the idea of quality but was still essentially a list of roughly equal weighting.

Each of these models had its strengths but none was sufficient on its own. The final choice was the result of a series of discussions with those responsible for delivering training in the region and was implemented against a backdrop of a wider debate taking place (Fulton, 1996) on the possible need to reconsider the goals and methods of audit throughout the United Kingdom. In response to this paper the author argued in a letter to the BMJ (Lough, 1996) that it was far too early to conclude that "audit is not working" and that strong leadership and a sense of direction were required. Baker and Fraser (1996) agreed, stating that

misunderstandings about the true nature of audit and confusion about the debate were giving an overly-negative view of the future potential of audit to deliver better health care.

The two aims and the number of objectives of the programme were relatively modest. The first aim - the integration of an audit programme to provide wider experience for teaching audit method - was built around the idea of Baker's Practice Audit Plan (Baker & Presley, 1990). Rather than each practice designing its own audit plan, which had been requested during the earlier discussion of the regional programme, each training practice would have a common programme. This also addressed the political drive in Scotland to promote audit by the Scottish Executive.

The second aim was a governance issue for the region in meeting its responsibility to the JCPTGP. By the end of the programme all training practices would have in place a defined series of activities which would be assessable by verification, would have addressed key issues for patients, encouraged team work and collaboration in training districts, addressed the measurement of performance against defined standards and finally - crucially - offered registrars the opportunity at least to have demonstrated and, possibly, take part in a range of audits which would build on their confidence which had been gained in submitting their audit project for summative assessment.

The three broad areas of the programme were the result of wide debate which, at times, was heated. Access to practices through their appointment systems has consistently been seen as an obstacle by patients and the difficulties of matching supply of appointments with patient demand against explicit standards expected of training practices in the region was always going to be controversial. Registrars' appreciation of the tensions in delivering an efficient and effective appointment

system was paramount with the straightforward need to start on time to reduce the risk of finishing late being a key message, with a particular emphasis placed on the role of doctor behaviour (Heaney *et al*, 1991).

Baker and Streathfield (1995) showed that training practices were often associated with reduced levels of overall satisfaction and particularly so with availability and continuity of care. In this study training practices were overly-represented. The need for protected time for teaching and the annual changeover of registrars both militate to some extent against continuity although many other features of training practices have been shown to offer advantages such as practice organisation, computers and recall systems. Monitoring patients' satisfaction with aspects of their care was therefore accepted as being a vital part of the programme and the need to use valid questionnaires was emphasised. The consultation satisfaction and surgery satisfaction questionnaires were two such published examples (Baker, 1990c; Baker, 1991; Baker & Whitefield, 1992). Wider discussion with the registrars on the difficulties in executing appropriate patient satisfaction questionnaires was encouraged as was the importance of proper planning and team work to ensure that an appropriate system for issuing the questionnaires was applied.

Managing chronic disease is arguably one of the biggest challenges for general practice. Wagner *et al* (1996) emphasised that one of the components of high quality care for patients with chronic disease is the level of organisation adopted by a practice. Such organisation can be based around a register of patients, regularly updated, with a specific disease matched against known prevalence to ensure that the maximum number of patients with the disease are accounted for. A method of collecting valid and reliable data against evidence-based criteria contained in a protocol which meets the needs of staff and patients alike all

contribute to a series of systems which maximised the chance of providing quality care for these vulnerable groups of patients.

Criteria can be defined as “systematically developed statements that can be used to assess the appropriateness of health care decisions, services, and outcomes” (Institute of Medicine, 1992). Baker and Fraser (1995) describe four key principles for audit criteria:

- they should be prioritised.
- they should be measurable.
- they should be appropriate to the setting.
- they should be based on evidence.

The need to prioritise criteria is paramount given the potential for data overload if too many criteria are being assessed and the procedure for categorising criteria on the strength of the research evidence available into “must do”, “should do” and “could do” allowed practices to make judgements on their performance against the strongest research evidence. The involvement of clinical and non-clinical members of the practice are also vital in providing quality organised care for chronic disease patients. For a registrar, recognition is vital that a range of skills - in particular, appropriate leadership skills - are required to run a multi-disciplinary team. Due to lack of training and support these skills are often deficient or absent (Firth-Cozens, 2001). The increasing involvement of nurses - both practice and community - in managing and auditing chronic disease management is acknowledged. Cheater and Keane (1998) showed that where nurses had a major involvement in making decisions multi-disciplinary audit tended to be more successfully established in the delivery of collaborative audit. This study also acknowledged that hierarchical structures in nursing and medical relationships and other pressures on workload and lack of time still contributed to creating obstacles

to better involvement of nurses in audit. The choice of the five diseases to be audited was somewhat arbitrary although built on data already being collected, such as for diabetes and asthma, as a result of previous political initiatives for health promotion. It was agreed from the outset that one identified change in care should be implemented for each disease over the five year period. The emphasis was on ensuring that structures and processes were in place and as a possible consequence of this four of the five chronic diseases showed a significant improvement in the proportion of practices having updated disease registers in place. Despite the relatively small demand of one change occurring for each disease there was a very small change in the proportion of practices achieving this over the six year period, a reminder of the difficulties in implementing change, particularly when relying on paper-based data collection systems.

The most successful area of the programme was the qualitative component of analysing significant events. The highly significant improvement in the proportions achieving five significant events analysed to a recommended format (Marinker, 1990) was very encouraging given the potentially highly sensitive nature of this qualitative method of audit which involved a degree of trust within the practice team. The importance in teaching the registrar the value of openness and trust in dealing with events which happened unexpectedly was recognised by the trainers and registrars alike as a valuable teaching and learning tool.

The overall programme was designed to be mixed, illustrating the importance of considering the structure, process and outcome of care in association with the technical and inter-personal components involved in their assessment. The importance of appreciating that quality is about making judgements by health care professionals, patients and managers is an important part of this framework (Donabedian, 1980). The need for guidelines using evidence-based criteria (Institute of Medicine, 1992) – the technical aspect of the quality of care - requires

to be balanced by the inter-personal aspects of care where best evidence may not be available. In this programme the criteria for appointment systems represented the patients' journey from their decision to seek a contact with their doctor through to the completion of the consultation. A Which? consumer survey (1995) had shown that patients placed a high priority on the potential for improvement in many aspects of appointment systems.

The criteria to be used by practices were taken from the Eli Lilly National Clinical Audit Centre Patients' Charter Protocols (Set 1) (Baker *et al*, 1994). Criteria for the chronic diseases were to be chosen by the practices themselves, included in the practice protocol and incorporated into a data collection sheet which would be part of the patient record. Examples of appropriate criteria and the evidence base from which they were derived were sent to all practices with the final prioritisation left to them.

The setting of appropriate standards of care has usually been accompanied by controversy as to whether standards should be minimal or optimal. Baker (1988) describes an ideal standard as being appropriate for an educational process. This may help to explain why few practices had assessed their appointment systems and had very few structures and processes in place for chronic disease management. It also allowed for all practices to improve and demonstrate this improvement to a registrar. The emphasis for appointment systems was on process and outcome audits using explicit criteria and standards. Chronic disease management was more concerned with structures and processes with explicit standards of care for outcomes being felt to be beyond the remit of the programme.

Although the original aims for the programme were met there are limitations to its success. Comparisons over the two three-year periods did not match practices in a "before and after" method as new practices were accredited to train and some

practices were de-accredited. Overall proportions achieving the various criteria and standards were therefore calculated thus ignoring the measured impact of true gain in individual practices. An enormous handicap for 85% of the practices was the fact that the GPass database (General Practice Administration System for Scotland) has no clinical component. Some practices had devised their own spreadsheets or databases but most had to rely on manual data entry and collection. This may explain why so few practices achieved an improvement in "change against one criterion" for chronic diseases. Addressing this deficiency in clinical systems is a matter of top priority for the Scottish Executive.

Hearnshaw *et al* (1998) described a quality improvement programme offered to all 147 primary health care teams in Leicestershire in 1994. Five teams ultimately completed their projects and three completed all seven sessions of the programme. The programme was externally facilitated and evaluated by a research team and changes made during the programme were still found to be in place after three years in three of the six teams. External facilitation was felt to be a positive feature for those committed to the programme although the numbers finishing compared with those invited to take part was a reminder of how difficult it can be to constitute a rolling programme of audit without some external influence.

Another example of facilitated audit was carried out by a team in Tayside (Grant *et al*, 1998) in response to MEL (52) from the Scottish Executive in 1995. Fourteen practices were randomly split into two groups - one group acting as a control - with facilitation, education in audit and I.T. and reimbursement for time spent participating in audit being offered to the seven intervention practices. Formal statistical comparison between the two groups was not undertaken although of 20 topics audited in the intervention group, 15 completed one cycle compared with 21 topics audited in the control group with only two completing a

cycle. The project lasted one year. 12 of the 14 practices were training practices and the cost was £30,000.

In conclusion, training practices are required by the JCPTGP to demonstrate that active audit is taking place for the benefit of a registrar. A programme has been defined covering quantitative and qualitative methods with a significant proportion of the programme considering areas of importance for patients. The increases in proportions achieving the various process structure and outcome audits were modest but by the end of the programme the majority of training practices had addressed the audits expected of them.

**Table 1 - Practice appointment systems –
differences in proportions of training practices
achieving programme criteria**

	1998 n = 114 (%)	2001 n = 113 (%)	p value	Difference in proportions (95% C.I.)
Seen by doctor for non-urgent appointment within three days (audit carried out)	89 (78)	100 (88)	0.05	10% (1% to 20%)
90% achieved	55 (48)	69 (61)	0.06	13% (0% to 26%)
Waiting no longer than 15 minutes from appointment time (audit carried out)	91 (80)	100 (88)	0.10	8% (-1% to 18%)
80% achieved	71 (62)	77 (68)	0.40	6% (-6% to 18%)
% surgeries starting within 5 minutes from appointment time (audit carried out)	85 (75)	95 (84)	0.10	9% (-1% to 20%)
90% achieved	48 (42)	65 (58)	0.02*	16% (3% to 28%)
Finishing within 20 minutes of time (audit carried out)	82 (72)	92 (81)	0.12	9% (1% to 20%)
75% achieved	58 (51)	69 (61)	0.14	10% (3% to 23%)
CSQ carried out within practice	79 (69)	85 (75)	0.37	6% (-6% to 18%)
CSQ completed for each partner and registrar	75 (66)	81 (72)	0.39	6% (-6% to 18%)
One SSQ for practice carried out	88 (77)	89 (79)	0.87	2% (-9% to 12%)

* denotes significance at $p < 0.05$ (Fisher's exact test)

**Table 2 - Chronic Disease Management –
differences in proportions of training practices
achieving programme criteria**

Asthma				
	1998 (n = 114) %	2001 (n = 113) %	p value	Difference in proportions (95% C.I.)
DR	54 (47)	72 (63)	0.02*	16% (4% to 29%)
FC	70 (61)	83 (73)	0.07	12% (0% to 24%)
P	81 (71)	87 (77)	0.36	6% (-5% to 17%)
C	41 (36)	45 (40)	0.59	4% (-9% to 16%)
Diabetes				
	1998 (n = 114) %	2001 (n = 113) %	p value	Difference in proportions (95% C.I.)
DR	55 (48)	73 (64)	0.02*	16% (4% to 29%)
FC	76 (67)	84 (75)	0.25	8% (-4% to 20%)
P	77 (68)	86 (77)	0.18	9% (-3% to 20%)
C	52 (46)	54 (48)	0.79	2% (-1% to 15%)
Epilepsy				
	1998 (n = 114) %	2001 (n = 113) %	p value	Difference in proportions (95% C.I.)
DR	51 (45)	71 (63)	0.01*	18% (5% to 31%)
FC	59 (52)	63 (56)	0.60	4% (-9% to 17%)
P	50 (44)	59 (52)	0.23	8% (-5% to 21%)
C	27 (24)	37 (33)	0.14	9% (-3% to 21%)

* denotes significance at $p < 0.05$ (Fisher's exact test)

DR — Disease Register

FC — Flow Chart

P — Protocol

C — Change against one criterion

**Table 2 - Chronic Disease Management –
differences in proportions of training practices
achieving programme criteria**

Hypertension				
	1998 (n = 114) %	2001 (n = 113) %	p value	Difference in proportions (95% C.I.)
DR	52 (46)	70 (62)	0.02*	16% (4% to 29%)
FC	52 (46)	66 (59)	0.06	13% (0% to 26%)
P	76 (67)	83 (74)	0.31	7% (-5% to 19%)
C	33 (29)	44 (39)	0.12	10% (-3% to 21%)

* denotes significance at $p < 0.05$ (Fisher's exact test)

Rheumatoid Arthritis				
	1998 (n = 114) %	2001 (n = 113) %	p value	Difference in proportions (95% C.I.)
DR	50 (44)	62 (55)	0.11	11% (-2% to 24%)
FC	42 (37)	49 (43)	0.35	6% (-6% to 19%)
P	57 (50)	62 (55)	0.51	5% (-8% to 18%)
C	19 (17)	26 (23)	0.25	6% (-4% to 17%)

DR – Disease Register

FC – Flow Chart

P – Protocol

C – Change against one criterion

**Table 3 – Significant Event Analysis –
difference in proportions of training practices achieving criterion**

	1998 n = 114 (%)	2001 n = 113 (%)	p value	Difference in proportions (95% C.I.)
Achieved five in past three years	64 (56)	91 (81)	0.0001**	25% (13% to 36%)

** denotes significance at $p < 0.001$ (Fisher's exact test)

CHAPTER 8

FACILITATING THE COLLECTION OF DATA FOR AUDIT

Lack of time and resources were identified by trainers and registrars as difficulties in implementing audit. The infrastructure for implementing audit at health board level was through area clinical audit committees (ACACs) who were funded each year by the Clinical Resource and Audit Group (CRAG) of the Scottish Office with the responsibility and accountability being taken at a local level with no specific guidance given on how the money should be spent. Some of this money paid for audit support staff, much of whose role was to respond to practice requests for collecting data from case records.

The audit programme implemented in training practices offered an opportunity to assess the costs involved in collecting data for audit purposes in training practices.

METHOD

12 trainers from one Health Board area in the west of Scotland agreed to their practices being monitored as part of the data collection necessary for part of the regional audit programme. Four audit support staff who had already received appropriate training in collecting data were issued with a specially designed chart to record the data they were collecting at 15 minute intervals for a period of approximately 30 weeks. Specially prepared data collection sheets were distributed to the staff and each was given a customised software programme on a Visual Fox-Pro database for the collation and analysis of the data.

a) Appointment system audits

The appointment systems audits involved existing practice staff who were collecting the data as part of their regular work on paper sheets with the external support staff merely transposing these to the computer software.

The data to be collected were:

- appointment availability - date appointment requested and date actually seen.
- waiting times - patient arrival time, patient appointment time, time seen by doctor, time finished with doctor.
- telephone audit - patients were asked if they felt they had waited more than six rings before the phone was answered. The answer “yes” or “no” was recorded.
- doctor recall of patients. Doctors recorded at the end of each consultation whether patients were discharged, asked to return at the patient’s discretion, or asked to return by the doctor within a given time interval.

b) Chronic disease audits

The 12 practices met to agree on specific review criteria covering the five chronic diseases of the audit programme. These were:

- asthma - recording of peak flow, smoking status, use of a spacer device, patients hospitalised due to asthma since their last surgery visit.
- diabetes - smoking status, last recorded blood pressure, most recent HbA1c within the last year, patients registered as blind.

- epilepsy - date of last seizure, number of different drugs being taken for epilepsy, number of drugs for epilepsy being taken on a more than twice daily regime.
- hypertension - the presence in the case records of the mean of three pre-treatment blood pressures, smoking status, most recent blood pressure reading.
- rheumatoid arthritis - diagnostic accuracy based on American Rheumatism Association criteria, a recording of patients hospitalised with upper gastrointestinal haemorrhage.

Again, specially designed data collection sheets were used. Each practice was able to decide which disease(s) they wished to prioritise. They supplied a list of their patients with diabetes, epilepsy and rheumatoid arthritis. For patients with asthma or hypertension an appropriate sample was chosen from the list of patients supplied by the practice each of whom had been given a random number. The sample size was chosen by consulting a sample size table as described by Derry (1993).

The support staff extracted the relevant patients' case records and searched in the records for the pre-agreed criteria and recorded these on the data collection sheets. Data were fed into the computer software at the end of each day.

c) Analysis

The total times taken for the audits were collated and estimations made of comparative costs for using a receptionist or a practice nurse. Salaries for the support staff were calculated at £5.59 per hour. Salaries for the reception staff

(Grade 3, spine point 9) were calculated at £5.19 per hour. Salaries for practice nurses (mid-point scale F of Whitley Scale) were calculated at £9.59 per hour.

RESULTS

The list sizes of the 12 practices ranged from 3380 to 11700 (mean 6943) with a median number of doctors per practice of 4.75 (range 3 to 8).

The appointments system audit was chosen by 10 practices, the waiting times audit by nine practices, the recall audit by eight practices and the telephone answering audit by seven practices. Transfer of data from the data collection sheets took 78 hours for the appointments audit, 125 hours for the waiting times audit, 24 hours for the recall audit and 17 hours for the telephone audit. The cost per 1,000 patients using the support staff and the estimated costs using receptionists or practice nurses are shown in Table 1.

The average time taken to extract the case records of the patients with the relevant chronic diseases or collect the relevant data was 20 minutes for a hypertensive patients, 18 minutes for a rheumatoid arthritis patient, 17 minutes for a diabetic patient, 16 minutes for an epilepsy patient and 14 minutes for a patient with asthma. The total time for collecting the data and the average costs for using support staff with estimations for practice receptionists or practices nurses are shown in Table 2.

DISCUSSION

The implementation of a specific audit programme was the catalyst to explore the time and costs involved in collecting data for routine audit. The waiting time audit - which took most time - involved four different sets of data from the time the

patient arrived at the surgery to the time they left the consultation room. The time and costs of running this audit had to be offset by the benefits of one of the commonest causes of a dysfunctional appointment system - a mismatch between booking interval and consultation length with the doctor persistently running late, a common cause of stress for both doctor and patient.

Finding data on patients with hypertension took most time and therefore had the greatest cost. One of the reasons was recognising doctors' handwriting, feedback of which might encourage doctors to consider a link between careless and potentially unsafe record keeping and the increasing costs which can accrue. Time, and therefore costs, were saved by a variety of components inherent in accredited training practices such as age-sex registers, morbidity registers, summarised records and a higher level of computerisation. Baker and Thompson (1995) implied that such structures could not be assumed in non-training practices to the same degree and therefore an assumption can be made that the time taken and the costs borne by non-training practices will be greater.

Despite this the General Medical Council has stated (1998a; 1998b) that monitoring and improving quality of care is the responsibility of the profession. Without the equivalent overseeing structure of the JCPTGP to quality assure minimum standards of practice as exists in training practices, many non-training practices may not be in a position to deliver on their responsibilities for such monitoring.

The role of computerisation in the routine collection of data is potentially advantageous but is not without its sceptics (Powsner *et al*, 1998). A poorly designed computer record will not replace a well designed paper record. The crucial issue - whether paper or paperless - in the quest to help clinicians retrieve data more quickly, with less effort and therefore less cost is attention to detail in

basic design principles as described by Nygren *et al* (1998). Patient data are easier to find if consideration is given in the design of records to:

- organisation of documents in the record, for example filing documents chronologically. What appears to be important is the consistency in whichever strategy is chosen rather than any one particular strategy over another. Chronological filing of general practice and hospital records is a JCPTGP criterion for training.
- organisation of data on the document page. Errors are reduced when more data needed to support decisions are viewed on one page rather than several. The use of summary sheets is such an example, particularly important with multiple pathology and medication. Summarising of records is another JCPTGP criterion for training.
- highlighting of data with cues, colour coding or pre-designed data collection sheets with the criteria already agreed can be helpful in reducing the search for data as they concentrate the data for a particular disease in one area of the record.

Lack of time to carry out audit was highlighted in an early evaluation of audit in Scotland (National Audit Office, 1994). The report highlighted the tension between service commitments on a daily basis and the benefits of audits which were often perceived as being longer term. Facilitating the audit process using specifically trained individuals external to the practice was raised as a possible advantage thus avoiding doctors and nurses carrying out tasks which could have been done by others. The role of facilitation and its associated costs in the long term are controversial. McCowan *et al* (1997) followed up more than 3,000 children for four years with approximately half being managed by an audit

facilitator and the other half acting as controls. The facilitator did not see patients directly but supported existing practice staff. The study lasted for two years with further assessment over the subsequent two years. There was a reduction in the number of consultations, prescriptions and hospital admissions in the intervention group which did not persist after the facilitator was withdrawn. By the end of the study process and outcomes measures were similar in both groups. The conclusion of the study was that intervention by a facilitator improved process and outcome of care. Costs saved during this period covered the cost of employing the facilitator. Sustained improvement without the facilitator, however, was not possible.

This study in training practices as part of the regional audit programme confirms that external support for collecting and collating data for audit purposes is likely to be too costly to sustain in the longer term. Using this study's figures for calculating the time taken for data collection Shepherd (2000) estimated that it would take a full-time worker over 19 weeks to extract data (without subsequent analysis) for seven chronic disease audits carried out at their practice annually. He advocated appropriately customised computer software to reduce the time and therefore the cost for extracting and analysing such data. Data were available almost immediately and data input costs had been all but eliminated due to the integration of data collection into routine patient care. This is not possible with current computing systems for the majority (around 85%) of practices in Scotland who use GPass.

In conclusion, the integration of audit into daily practice carries a cost in both time and money. External support is probably unsustainable. Making most efficient use of available practice staff while trying to improve computing software is the current way forward.

Table 1 - Average cost (in pounds sterling) of transferring data for workload audit per 1000 patients

	Appointments (n=61000)	Waiting Times (n=64300)	Re-call (n=48900)	Telephone (n=52600)	Total
Total time for data collection (hours)	78	125	24	17	244
Average cost per thousand patients – receptionist: estimated	6.64	10.09	2.55	1.68	20.96
Average cost per thousand patients – support staff	7.15	10.87	2.74	1.81	22.57
Average cost per thousand patients – practice nurse: estimated	12.26	18.64	4.71	3.10	38.71

Table 2 - Average cost (in pounds sterling) for retrieving data for each chronic disease

	Diabetes (n=205)	Asthma (n=711)	Epilepsy (n=349)	Hypertension (n=1538)	Rheumatoid Arthritis (n=338)	Total
Total time for data collection (hours)	57	164	95	527	103	946
Average cost per patient - reception staff	1.44	1.20	1.41	1.78	1.58	1.48
Average cost per patient – support staff	1.55	1.29	1.52	1.91	1.71	1.60
Average cost per patient – practice nurse	2.67	2.21	2.61	3.29	2.92	2.74

CHAPTER 9

COMPLETION OF THE AUDIT CYCLE - EVALUATING VERSUS PROPOSING CHANGE

Between 1996 and 1997 an increasing number of registrars was evaluating the change they had proposed in their audit project for summative assessment. By completing a cycle of audit they were going beyond what was expected from the five criteria against which their project was being assessed. Thus within four years of its implementation in the west of Scotland the confidence which previous registrars had expressed from submitting an audit project was being translated - at least by a sizeable minority - into a peer motivated rising of standards in the completion of their project.

Summative assessment was implemented as a professionally-led system throughout the United Kingdom in September 1996 (UKCRA, 1995). In response to this the regulations for vocational training for the United Kingdom required to be changed. Officers of the JCPTGP and officials from the Department of Health drew up a revised list of competencies which required to be addressed by the summative assessment process with the latter acquiring legal status in January 1998 (National Health Service Regulations, 1997). A total of seven competencies was included in the legislation. The submission of a criterion audit project now required that a registrar to demonstrate that he/she had acquired "the ability to review and critically analyse the practitioner's own working practice and manage any necessary changes appropriately".

The marking schedule in use satisfied the first part of the competency-definition but fell short of managing the change process beyond suggesting proposals. The implication in the definition was that a registrar should be able to complete an

audit cycle and, as a proportion of registrars in the west of Scotland were attempting to demonstrate, the assessment process would need to be modified to assess more closely the competence defined in law.

METHODS

System Development

The original list of 14 elements considered essential or desirable as part of a registrar's audit project was sent to the 144 trainers in the west of Scotland in 1997. They were asked to reconsider whether each element should now be considered an essential or desirable part of the constitution of the audit project by a registrar. In making their decision they should take into account their current confidence in teaching audit method and the fact that an increasing number of registrars were submitting completed audit cycles. The group of assessors considered the trainers' opinions and crafted a new marking schedule from the results.

In order to test whether the assessors could use the new marking schedule to identify a completed audit cycle, 20 projects were chosen from the 98 audit projects submitted for summative assessment in the previous year. These projects were divided into complete and incomplete audit cycles and 10 from each group were chosen at random and sent to 26 assessors who marked them individually.

Following the completion of this exercise levels of agreement on individual projects were assessed and those with more disagreement were considered in detail with particular attention being paid to the fact that the previous five criteria marking schedule and its associated pro-forma had been used in the original assessment.

System Implementation

All 57 registrars beginning their training in August 1997 were given instruction on the new assessment system of eight criteria for their audit project. They were also advised of the five criteria system by which they would ultimately be assessed should they not pass under the new system. This was to ensure that the registrars undergoing the pilot of a new system were not disadvantaged compared with their peers throughout the United Kingdom.

System Testing

Eleven projects were chosen at random - nine (from 42) which had passed at first level by two independent assessors and two (from seven) referred by the second level assessment for resubmission - to test the sensitivity and specificity of the new marking schedule.

System Acceptability and Feasibility

The registrars were asked whether they found completing the audit cycle easier than expected, as expected or more difficult than expected. They were also asked if they had ever completed an audit cycle prior to their summative assessment audit project.

All the assessors gave their opinions on whether marking with the new instrument had been easier, as expected, or more difficult than expected compared with experience of the previous instrument.

Construct Validity

Between 1996 and 1998 factor analysis was applied to the judgements of 210 audit projects which had been assessed using the eight criteria method. Analysis was carried out using SAS.

Reliability

The 210 audit projects submitted between 1996 and 1998 were each given a random number and 100 were chosen for an exercise to test the inter-rater reliability, as defined by the kappa statistic, and to assess the proportions of agreements of judgements between pairings of assessors.

The 100 projects were divided into five equal batches, with each of 20 assessors marking a batch of 20 projects. There were thus four assessors marking each batch of 20 projects, resulting in a possible six potential pairings of assessments for each project.

RESULTS

System Development

There was a response from 129 (89%) of the 144 trainers to the list of elements for a completed audit cycle. A comparison of the opinions from 1995 is shown in Table 1 with differences calculated in the proportions (with 95% confidence intervals) of each cohort of trainers for each element. Overall there was a much tighter spread of opinion on whether the elements should be an essential or desirable part of an audit project.

There was an increase in the proportion of trainers at the 1% level of significance who felt that the following elements were an essential or desirable part of an audit project from a registrar in general practice:

- *reason for choice of project given.*
- *relevance of data to criteria.*
- *system for change proposed.*
- *staff involvement.*

- *negotiation with relevant team members.*
- *protected time required.*
- *second collection of data required.*
- *money/funding required.*

In addition, there was an increase at the 5% level of significance in the proportion of trainers who felt that *further change proposed where appropriate* should be included.

The identification of a completed audit cycle using the eight criteria marking instrument by the 26 assessors is shown in Table 2(a). A “pass” project represents identification of a completed audit cycle with “refer” representing an incomplete audit cycle. Projects B, C, E, H, I, K, M, P, Q and S were the incomplete audit cycles. There was 95% agreement with the previous assessments on decisions to refer with 66% agreement on decisions to pass. This resulted in a small number of incomplete cycles being missed but highlighted some completed cycles which should have passed. In order to maximise the probability of identifying a completed audit cycle and to minimise the possibility of one being missed, two assessors were used to mark the projects independently.

Summary statistics on assessor agreement are shown in Table 2(b). On average 81% of the assessments were in agreement with 95% of the agreements being between 67% and 95% of the total assessments made.

The new marking instrument based on eight criteria with advice to the assessors is shown in Figure 1.

The new system was based on a similar method of screening a project by two independent assessments. If the two assessments agreed that a completed audit

cycle was present by identifying all eight criteria it passed. If one or both assessors had any doubts about any of the criteria being present the project was referred to two more experienced assessors at a second level who again assessed the project independently against the same criteria. They ultimately had to agree on whether the project should pass or be referred with feedback to the registrar for an opportunity to resubmit the project.

Of the 57 audit projects submitted, 42 passed and 15 were referred for further assessment. Of these 15, eight were judged to be a pass and seven were referred with advice for resubmission. Four of these projects required further work on criteria and standards and the remaining three had problems comparing data collections one and two. All seven projects subsequently passed. In order to ensure that none of the 42 projects which had passed at first level had been missed as an incomplete audit cycle all were assessed independently by the author who agreed with all the assessments.

The results of the marking exercise to calculate the sensitivity and specificity of the instrument in identifying a completed audit cycle are shown in Table 3(a). The projects used in the exercise were a sample of nine passes (of 42) and two (of seven) refers or resubmission (projects E and J) drawn from the first cohort of registrars to complete the new system. All 11 projects were marked by 24 of the original 26 assessors, two being unavailable for marking.

The summary statistics on assessor agreement are shown in Table 3(b). Again, agreement between pairs of assessors was very reasonable with 95% of the agreements being between 62% and 94% of all assessments.

The sensitivities and specificities were calculated by multiplying the number of possible assessments able to identify the two refer projects using one or two

assessors and applying a set of “pass” and “fail” rules. The results are shown in Table 4. The best balance of sensitivity and specificity was achieved with two assessors where a ‘fail’ was identified by one assessor and a ‘pass’ by both assessors.

Acceptability and Feasibility of the System

20 (83%) of the 24 assessors gave their opinions on the ease of marking. All felt that the new marking schedule had been easier to use in making assessments than the previous one.

54 registrars (95%) responded with their comments on the ease of completing an audit cycle. 47 (87%) claimed to find that completing an audit cycle was as expected or easier than expected. Only 11 (20%) had ever completed a cycle of audit prior to the submission of their audit project.

These projects took about 10 minutes to mark with the cost per registrar dropping from £42.60 to £36.75 (based on a payment to assessors of £12 per project). A move from the five criteria to the eight criteria system would therefore result in a saving of £585 per 100 registrars.

Construct Validity

A two-factor model Promax rotated factor pattern suggested that the majority of items loaded heavily on factor 1 with *reason for choice* loading on factor 2. Also represented on factor 2 was *criterion/criteria chosen* and, to a lesser extent, *standards set*. *Criterion/criteria chosen* and *standards set* were also represented on factor 1. One interpretation of this pattern is that *reason for choice* tended to be marked differently from the other criteria.

In a three-factor model the Promax rotated pattern suggested a different grouping of variables. One interpretation of the pattern is that as factor 1 shows fairly heavy loadings for criteria 4 to 8, this might represent "the execution" of the audit project. Criteria 2 and 3 showed heavy loadings on factor 2 which could correspond to "definition and scope" of the project. Finally, factor 3 exhibited a heaving loading for *reason for choice* which might represent the "choice of subject" for the project.

In considering the eigen values of the correlation matrix, factor 1 explained 48% of the total sample variance with factor 2 explaining a further 12% and factor 3 a further 11%. The three factors retained therefore accounted for 71% of the total variance (Table 5).

Reliability

The summary statistics for the average kappa values and average proportions of agreements between assessor-pairings are shown in Table 6.

The results of both are very similar to the reliability values calculated for the five-criteria system.

DISCUSSION

The implementation of an audit programme in training practices provided a framework within which registrars could see and participate in a variety of relevant audits. The interest in evaluating change in their audit project for summative assessment went beyond the criteria expected of them. Rogers (1983) described the characteristics and values of five categories of "adopters" in the uptake of new ideas. Although originally used in the uptake of new farming practices among seed farmers in Iowa it has been applied and confirmed by other

studies in many countries around the world. In a study alluded to earlier in the thesis Baker and Thompson (1995) showed that although patterns of diffusion of innovation can be similar over a timescale - in their case between 1982 and 1990 - the gap between the innovators and laggards became wider. There was an increase in scores associated with practice developments between 1982 and 1990 of 26 among the top ten innovators compared with 15 between the lowest ten. Many of the developments were related to practice structure but were felt to be crucial in the development and implementation of systems towards improving the quality of care.

The registrars who were spontaneously evaluating change in their audit project could be seen as innovators. With their personal characteristics and values described as "venturesome" which has an influence on their leadership-style they can be considered important in leading change among their peers.

The impact of an assessment process on the learner is well documented such as in medical students (Newble & Jaeger, 1983) with particular reference to their styles of learning and the way they are taught as described by Entwistle (1987). Popham *et al* (1985) described the challenge for those developing testing systems to take a strategic view of "measurement-driven instruction" and to promote desirable learning behaviour.

Further, Moss (1992) has described an emerging consensus among measurement researchers about the importance of expanding the concept of validity to include the explicit consideration of the consequences of the use of assessment.

Thus the significant proportion of registrars willing to evaluate change, possibly as a consequence of improving structures and processes in their training practices and the "direction of travel" of the assessment process for audit, dictated a need to

modify the existing testing procedure in the light of their experience as suggested by Bennett and Hayden (1995).

The five attributes of an assessment process described by Van der Vleuten (1996) are similar to those outlined for the five criteria assessment instrument. The high return from the trainers and their more positive response to the elements felt to be essential or desirable as part of an audit project, possibly through an increase in their confidence in teaching the subject, suggest a higher acceptability than previously. Registrars' acceptability with the change was also positive. A reduction from three assessors to two at the first level with no loss of sensitivity in the process reduced costs and simplified marking thus increasing the feasibility.

Inter-rater reliability as defined by the kappa statistic was modest which would be considered unacceptable in a high-stakes test such as losing one's licence to practise. With feedback, however, and a chance to resubmit the project with appropriate teaching reliability as defined by the kappa statistic has been "traded-off" with other attributes of the system. Overall, the referral process at three levels (two within the region, one outwith the region for any potential fail) should increase the reliability by increasing the proportion of agreements on decisions giving a fairer system for quality assurance across the United Kingdom. A project could ultimately therefore only fail after a total of eight independent judgements (six within the deanery - two at first level, four at second level - with a further two in another deanery in the United Kingdom).

In contrast to the kappa values, the proportions of agreements between pairings of assessors were more reassuring.

The scientific rigour with which the eight-criteria audit project and its assessment should be carried out is contentious. Russell and Wilson (1992) gave a robust

challenge for audit to be described as “the third clinical science”. They defined science as the “pursuit of knowledge” and scientific method as having principles laid down for performing observations and testing the soundness of conclusions. They described a taxonomy for clinical research incorporating an explanatory design for generating knowledge which explains bio-medical phenomena and is analysed by hypothesis testing. They also described a pragmatic design which generated knowledge which improved decisions for allocating resources and was analysed by statistical estimation of confidence intervals. The former can be described as “bio-medical science” i.e. pure clinical science and the latter as “health care science” i.e. applied clinical science.

Audit design may be described as generating knowledge which enhanced quality of health care. They described nine distinct steps for “scientific audit” synthesising the practical stages of clinical science and the essential steps of an audit cycle. The importance of basic statistical concepts such as an appropriately calculated sample size and the formulation of a specific hypothesis which would be tested by a proposed audit were central to their argument. Without this level of rigour Smith (1992b) had previously expressed the view that audit was bad research. As Barton and Thomson (1993) pointed out, however, “the goal of audit is not to become research but to become good audit defined by its capacity to improve the quality of patient care”. They expressed concern that adopting Russell’s and Wilson’s approach to audit might limit the potential to create change and, at times, may be totally inappropriate. A range of methods and project designs was more likely to be effective with the challenge being to choose and apply the most appropriate method for the task. Moreover a method which can make statistical sense may be clinically counter-intuitive, such as Russell’s and Wilson’s example involving treatment after myocardial infarction with aspirin and a beta blocker with a standard of 90% appropriately treated with aspirin and beta blocker but a level of less than 80% at which clinical concern would be merited

which would determine the number of cases required to reliably detect a clinically-defined difference from the chosen standard. Setting a standard of 90% but being willing to accept more than 80% may be statistically helpful but does not make much sense clinically.

The design of the audit project to be undertaken by a registrar to demonstrate competence in the ability to critically analyse an aspect of care and implement any necessary changes was, in reality, an uncontrolled before and after the study. The aim was to challenge the registrar to clearly identify a cause for a problem and address it. Attention to the principles on implementing change would not be sufficient. The construct of the audit project therefore addressed the three stages of effecting change as described by Crombie and Davies (1993). The three stages which they outlined were:

- confirm the problem.
- identify reasons for the problem.
- devise a strategy for change.

The project was therefore closer to the concept of continuous quality improvement which consists of multiple small cumulative changes as described by Bucknall *et al* (1992) in the context of asthma management and audit as the “audit spiral”. It was accepted therefore that it might be difficult to show conclusively that the intervention implemented under the direction of the registrar had been truly responsible for any demonstrated change and took account of the fact that more than one intervention may be used concurrently. It was also possible that the intervention might lead to a reduction in the element of care being measured and this would require reflection in the conclusion of the project. It was the responsibility of practices rather than the individual registrar to ensure a sustained change. This allowed the registrar to be involved in the setting of standards with

the practice but with the time constraint of the training period, possibly as little as six months, movement from one target to another towards the standard was more realistic. Statistical difference between measurements became less relevant in this context.

Nearly 25 years ago Nelson (1976) described the experience of the Utah professional review organisation in conducting medical audit using diagnosis-orientated process audits as resulting in an accumulation of vast amounts of unusable data which he termed "orphan data". Their experience suggested that audit could be more effective by focusing on a particular element of care, prioritised for its importance and potential for improvement. The importance of strongly validated criteria in this process was emphasised. Fowkes (1982) was one of the first to assign the concept of a cycle to this process by describing five activities which would constitute a cycle of audit. The twin concepts of orphan data and the audit cycle and their inherent problems are well known and described. A clear strategy for encouraging movement from the former to the latter is less well described. Placing the registrar at the centre of training practice audit, in particular encouraging discussion on the setting of explicit standards, against a wider framework of audit activity, offers both teacher and learner the opportunity to encourage a broader cultural change within the practice as a whole.

The impact of moving to eight from five criteria was investigated by McKay *et al* (2002). 261 audit projects assessed by the five criteria in the three years prior to the change were compared with 210 projects in the three years after the change.

A null hypothesis was tested of no difference in the number of criteria chosen by the registrar for a project. This was taken as a useful measure of the ultimate size of a project – the more criteria, the more data to be collected, analysed and

compared. Large data collections had already been identified as one reason for "orphan data".

97 projects (37%) contained four or more criteria prior to the introduction of eight criteria for a completed audit cycle compared with 29 projects (14%) after its introduction (χ^2 16.23, $p < 0.001$).

In addition 84 projects (32%) compared with 107 projects (51%) contained only one criterion to be measured before and after the change (χ^2 31.23, $p < 0.001$).

As a result of the change to a completed cycle with eight criteria therefore registrars were producing smaller, more focused audits, 90% of which were completed with six months.

In conclusion the move to an eight criteria based assessment of an audit project has addressed the legal definition of the competence to measure an aspect of the quality of care and the professional attributes as expected by the General Medical Council. In combination with the broader audit programme it offers a challenge to the JCPTGP to make more explicit the standards of teaching of audit method undertaken by training practices, already recognised as the innovators of general practice. The need for pragmatism in the light of trainers' varying abilities to teach has allowed for a move to incremental change but still lacks the rigour of evaluating the change in the context of the wider practice population such as by apportioning confidence intervals. It is possible that registrars again will dictate the need for this advance.

Table 1 – Change in trainers’ opinions on essential or desirable elements of a registrar’s audit project

Element of Audit Project	Current system n = 135(%)	Proposed system n = 129 (%)	Difference in proportions (95% CI)	p value	
Reason for choice of project given	116 (86%)	126 (98%)	0.12 (0.05 to 0.18)	<0.0001	**
Criteria applied	125 (97%)	126 (98%)	0.07 (0 to 0.10)	0.05	
Potential for change	124 (92%)	125 (97%)	0.05 (0 to 0.11)	0.07	
Standards set	130 (96%)	125 (97%)	0.01 (-0.04 to 0.05)	0.79	
Preparation and planning	130 (96%)	125 (97%)	0.01 (-0.04 to 0.05)	0.79	
Relevance of data to criteria	117 (87%)	125 (97%)	0.01 (0.04 to 0.17)	0.002	**
Interpretation of data	127 (94%)	125 (97%)	0.03 (-0.02 to 0.08)	0.23	
System for change proposed	118 (87%)	125 (97%)	0.05 (0.03 to 0.16)	0.003	**
Staff involvement	93 (69%)	124 (96%)	0.27 (0.19 to 0.36)	<0.0001	**
Negotiation with relevant team members	117 (87%)	124 (96%)	0.09 (0.03 to 0.16)	0.005	**
Proposed time required	100 (74%)	123 (95%)	0.21 (0.13 to 0.29)	<0.0001	**
Further change proposed where appropriate	111 (82%)	118 (91%)	0.09 (0.12 to 0.17)	0.02	*
Second collection of data compared	101 (75%)	116 (90%)	0.15 (0.06 to 0.24)	0.001	**
Money/funding required	62 (46%)	109 (84%)	0.38 (0.28 to 0.49)	<0.0001	**

* denotes $p < 0.05$

** denotes $p < 0.01$

Table 2(a) - Audit assessors' initial marking exercise to identify an incomplete audit cycle (columns in italics)

Assessor	Projects																			
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	R	R	R	P	R	P	P	R	R	R	R	P	R	P	P	R	R	P	R	P
2	R	R	R	R	R	P	P	R	R	R	R	P	R	P	P	R	<i>P</i>	P	R	P
3	P	R	R	R	R	P	P	R	R	R	R	R	R	P	P	R	R	P	R	P
4	P	R	R	R	R	P	R	R	R	R	R	R	R	P	P	R	R	P	R	R
5	P	R	R	R	R	R	P	R	R	R	R	R	R	P	R	R	R	P	R	P
6	P	R	R	R	R	P	P	R	R	P	R	P	R	P	P	R	R	P	R	P
7	P	R	R	R	R	P	P	R	R	R	R	R	R	P	P	R	R	P	R	P
8	R	R	R	R	R	R	R	R	R	R	R	R	R	P	P	R	R	P	R	P
9	R	R	R	R	R	P	R	R	R	R	R	R	R	P	P	R	R	P	R	P
10	P	R	R	P	R	R	P	R	R	P	R	R	R	P	P	R	R	P	R	P
11	P	R	R	R	R	P	P	R	R	P	R	P	<i>P</i>	P	P	R	R	P	R	P
12	P	R	R	R	R	R	P	<i>P</i>	R	R	R	R	R	P	P	R	<i>P</i>	P	R	P
13	P	R	R	R	R	R	P	R	R	R	R	R	R	R	P	R	R	P	R	P
14	R	R	R	R	R	P	P	R	R	R	R	R	R	P	P	R	R	P	R	P
15	P	R	R	R	R	P	P	R	R	R	R	R	R	P	P	R	R	P	R	P
16	P	R	R	R	R	R	P	R	R	R	R	R	R	P	P	R	R	P	R	P
17	R	R	R	P	<i>P</i>	P	P	R	R	R	R	R	<i>P</i>	R	P	R	R	P	R	P
18	P	<i>P</i>	<i>P</i>	P	R	R	P	R	<i>P</i>	P	<i>P</i>	R	R	P	P	R	R	P	R	P
19	R	R	R	P	R	P	R	R	R	R	R	R	R	R	R	R	R	P	R	P
20	P	R	R	R	R	P	P	R	R	P	R	R	R	P	P	R	R	P	R	P
21	R	R	R	R	R	R	R	R	R	R	R	R	R	P	P	R	R	P	R	P
22	P	R	R	R	R	P	P	R	R	R	R	R	R	P	P	R	R	P	R	P
23	P	R	R	P	<i>P</i>	P	P	R	R	R	R	P	R	P	P	R	R	P	R	P
24	P	R	R	R	R	P	P	R	R	R	R	P	<i>P</i>	P	P	R	R	P	R	P
25	P	R	R	R	R	P	P	R	R	R	R	R	R	R	P	R	R	P	R	P
26	P	R	R	P	R	P	P	R	R	P	R	P	R	P	P	R	R	P	R	P

Table 2(b) – Summary statistics of 26 assessors' judgements of 20 audit projects to identify an incomplete audit cycle

	mean	median	S.D.	minimum	maximum	Q1	Q3
Average kappa values	0.58	0.59	0.13	0.17	0.73	0.55	0.68
Average proportions of agreement	0.81	0.81	0.07	0.57	0.88	0.79	0.85

Table 3(a) - Audit assessors' final marking exercise to identify a 'poor' project (columns in italics)

Assessor	Projects										
	A	B	C	D	<i>E</i>	F	G	H	I	<i>J</i>	K
1	P	P	P	P	<i>P</i>	P	P	P	P	<i>P</i>	P
2	P	P	P	R	R	P	P	R	P	R	P
3	P	P	P	P	R	P	P	P	P	R	P
4	P	P	P	P	<i>P</i>	P	P	P	P	R	P
5	P	P	P	P	<i>P</i>	P	P	P	P	R	P
6	P	R	P	P	R	R	R	R	P	R	P
7	P	P	P	P	R	P	P	P	P	R	P
8	P	P	P	R	R	P	P	R	R	R	P
9	R	R	P	R	R	P	R	R	P	R	P
10	<i>P</i>	P	P	P	R	P	P	P	P	R	P
11	P	P	P	R	R	P	R	P	P	R	P
12	P	P	P	P	R	P	R	P	P	R	P
13	P	P	P	P	<i>P</i>	P	P	P	P	R	P
14	P	P	P	P	R	P	R	P	P	R	P
15	P	P	P	P	<i>P</i>	P	P	P	P	R	P
16	P	P	P	P	<i>P</i>	R	P	P	P	R	P
17	P	P	P	P	R	P	P	P	P	R	P
18	P	P	P	P	<i>P</i>	P	P	P	P	R	P
19	P	P	P	P	<i>P</i>	P	P	R	R	R	P
20	P	P	P	P	R	P	P	P	P	R	P
21	P	P	P	P	R	P	P	R	P	R	P
22	P	P	P	P	R	P	P	P	P	R	P
23	P	R	P	P	R	R	P	R	P	R	P
24	P	R	P	R	R	P	P	P	P	R	P

**Table 3(b) – Summary statistics of 24 assessors’ judgements
of 11 audit projects**

	mean	median	S.D.	minimum	maximum	Q1	Q3
Average kappa value	0.45	0.50	0.15	0	0.61	0.33	0.59
Average proportions of agreement	0.78	0.81	0.08	0.56	0.86	0.74	0.85

**Table 4 – Sensitivity and specificity of the instrument
using one or two assessors**

Fail rule	Sensitivity	Specificity	Pass rule
Fail if a single assessor fails it	83%	88%	Pass if a single assessor passes it
Fail if either one of a pair of assessors fails it	95%	77%	Pass if both of a pair of assessors passes it
Fail if both of a pair of assessors fails it	71%	98%	Pass if either one of a pair of assessors passes it

**Table 5 – Factor analysis of judgements of 210 eight criteria projects
(1996 – 1998)**

Factors	Eigen value	Proportion of Variance
1	3.83	0.48
2	0.99	0.12
3	0.85	0.11
4	0.62	0.08
5	0.50	0.06
6	0.47	0.06
7	0.41	0.05
8	0.34	0.04

**Table 6 – Reliability testing: summary statistics of 20 assessors' judgements
on assessments of 100 audit projects**

(1996 – 1998)

	mean	median	S.D.	minimum	maximum	Q1	Q3
Average kappa values	0.36	0.32	0.22	0.04	0.74	0.19	0.49
Average proportions of agreement	0.74	0.74	0.15	0.32	0.92	0.65	0.87

Figure 1 - Instructions for audit assessors and audit marking schedule

Please use the marking schedule to give your judgement on the registrar's audit project. It is crucial that the complete project is read before marking begins.

The criteria to be used for marking are in bold print. The statements in less bold print should act as a guide when making your judgement.

This will result in eight criteria to be marked. A "Pass" audit project will require all eight criteria to be present.

Please comment at any stage of the process but specifically if the registrar audit project is being referred.

REGISTRAR I.D. NUMBER:

PROJECT TITLE:

ASSESSOR:

Marking Schedule		
CRITERION		CRITERION PRESENT
Reason for choice of audit	Potential for change Relevant to the practice	<input type="checkbox"/>
Criterion/criteria chosen	Relevant to audit subject and justifiable (e.g. current literature)	<input type="checkbox"/>
Standards set	Targets towards a standard with a suitable timescale	<input type="checkbox"/>
Preparation and planning	Evidence of teamwork and adequate discussion where appropriate	<input type="checkbox"/>
Data collection (1)	Results compared against standard	<input type="checkbox"/>
Change(s) to be evaluated	Example supplied	<input type="checkbox"/>
Data collection (2)	Comparison with data collection (1) and standard	<input type="checkbox"/>
Conclusion	Summary of main issues (e.g. bullet points)	<input type="checkbox"/>

A satisfactory registrar audit project report should include all 8 criteria to pass

Pass ☐

Refer ☐

If refer, please comment on your reasons overleaf.

CHAPTER 10

QUALITATIVE AUDIT: ASSESSING AN ANALYSIS OF A SIGNIFICANT EVENT

The introduction of significant event auditing into the regional audit programme in 1995 highlighted the same educational challenge as that experienced in the introduction of criterion based audit in 1992, i.e. the need for an adequate definition. The introduction of a qualitative method of audit based on reflection of an expected event and a discussion surrounding its cause based on trust were potentially threatening but intuitively attractive in offering an opportunity for doctors in training to demonstrate an understanding of why certain events happen and to learn from them by disseminating good practice or reducing the risk of a similar event recurring should that be deemed likely.

Significant event analysis is a form of case-based audit (Pringle *et al*, 1995) which tends to rely retrospectively on individual events which often have strong emotional impact, either because they confirm an example of good practice or, more likely, identify problems with people or systems.

There are, however, drawbacks in analysing a significant event when opinions or perceptions about patient care may result in a lack of a deeper understanding as to the cause of such an event. In order to ensure that insight into the cause is demonstrated and subsequent lessons learned, defining a format for the analysis of a significant event is desirable, particularly in an educational context.

METHOD

Defining the analysis of a significant event

Marinker (1990) described six essential steps in formulating an enquiry into a significant event covering the reason for the enquiry, presentation of the evidence, a resulting policy, a sense of ownership, devising a system and finally an enquiry at some agreed time to ensure compliance with the policy (REPOSE).

20 general practitioners already experienced in the assessment of criterion based audits provided a focus group to consider how Markiner's model could be used to demonstrate a satisfactory analysis of a significant event. A selection of events considered significant by members of the group was discussed informed by Marinker's six elements.

Four questions were agreed upon which provided a frame work for analysing a significant event, submission of which would be on a proforma containing all four questions to be addressed.

Each of the above questions was matched with a corresponding question which together formed the framework assessing whether the analysis of the significant event had been adequate.

Thirty five significant events produced by the group to the agreed framework were circulated to all 20 assessors. Each was asked to judge them using the assessment questions but at the same time to record the reasons justifying their assessments. From the list produced, the most cited reasons - after discussion to minimise duplication - provided support for the judgements being made in assessing the analyses of the significant events. A combination of the assessment questions and their supporting statements were used to construct the final instrument. The staged process is shown in Figure 1.

Testing the assessment instrument

A pack containing background material on significant events, the proforma containing the four questions and a copy of the marking instrument were given to senior house officers (SHOs) in vocational training schemes from February 1998 in the west of Scotland who were asked to submit a significant event analysis in at least one of their four six month training posts. In addition the pack was offered to all principals in general practice in the west of Scotland who might wish to submit a significant event analysis for which they would be awarded one session of postgraduate education allowance (PGEA). The complete package was sent to all training practices as part of the regional audit programme but submission for assessment was on a voluntary basis.

All doctors were advised that events submitted should be anonymised by removing identifying features such as the place of work or named individuals at work as these could possibly compromise confidentiality. They were also aware that their work would be seen by other doctors in an educational setting as part of an external assessment.

The system used for assessing the analyses of significant events was based on two assessors marking independently with the responses to each of the four questions to be addressed being rated 0 (absent) or 1 (present). All four questions had to be adequately addressed for a satisfactory analysis of a significant event.

100 significant event analyses were chosen at random (48 from SHOs and 52 from GP principals), each having been given a random number from a computerised random number generator programme. The sampling frame consisted of 52 significant events from SHOs and 97 from principals in general practice. The 100 projects were then divided into five batches of 20 projects each and the 20

assessors were divided into "cells" of four assessors. Each cell marked a different batch containing 20 projects resulting in each project being marked by six potential pairs. Each assessor worked in ignorance of the marks of any other assessor.

Data on the overall dichotomous judgement of each event were entered into an Excel spreadsheet and the kappa values and proportions of agreements between pairings were calculated using SAS.

RESULTS

The four questions which were adapted from REPOSE differentiated between a description and an analysis of a significant event. A total of 84 statements were recorded in justifying the assessments of the questions being addressed in the analysis. These statements took into account the widely varying contexts of the significant events being experienced. In order to minimise duplication and overlap the ten most cited statements were used as part of the assessment process.

The final assessment instrument is shown in Figure 2.

Of the 71 projects where there was agreement between assessor-pairs, 42 were judged to be a satisfactory analysis of a significant event and 29 were judged to be an unsatisfactory analysis.

Summary statistics of average kappa values and proportions of agreements between assessors are shown in Table 1.

The mean kappa value was "fair" and 95% of the agreements occurred between 51% and 95% of all assessments made.

DISCUSSION

The word "audit" in the JCPTGP criterion is not defined. Criterion based audit was considered to be the method most appropriate to teach the "principles of audit" and to fulfil the GMC's interpretation of the professional responsibilities relating to quality of care assessment. Thus a quantitative method was initially chosen which would be of most benefit to young doctors in practice.

More recently, however, the problems inherent in criterion audit, for example, lack of time and resources, particularly with the lack of supportive software systems, have resulted in an increasing interest in more qualitative audit methods, such as case-based audit or significant event analysis. Many of the disadvantages for criterion audit are not such a problem with single events. As Buckley (1990) stated, large and important areas of clinical practice will be excluded if audit is restricted to the more quantifiable and measurable aspects of care.

Many of the principles which underpin audit in general are particularly appropriate to significant event analysis. McIntyre and Popper (1983) described the importance of tolerance in the search for mistakes, the goal of which should be educational and practical. They also looked forward to "a robust independent profession" which would openly welcome self critical analysis thus freeing the profession from threats of litigation and increasing government imposition. Berwick (1989) stated that the continuous search for improvement in health care could benefit from similar theories of quality improvement in industry. In most cases "for the average doctor, quality fails when systems fail".

The development of discussing clinical cases in small groups had a long tradition in general practice (Balint, 1957) and has been quoted frequently in the context of

vocational training. Although primarily concerned with the relationship between a doctor and the patient the critical analysis applied to their interaction was increasingly seen as appropriate in the context of a single event taking place in daily practice. The trust engendered by such discussion was considered crucial for honesty to prevail, lessons to be learned and significant improvements made. Such an environment was considered vital to protect a doctor in training lest psychological harm resulted from inadequate management of a sensitive issue.

The origins of analysing significant events are now generally considered to lie in a system which was developed in 1941 when the United States Army Air Force urgently needed to select, classify and train air crew in the shortest possible time. The often subjective reasons given for failing certain pilots in their training programmes persuaded a psychologist, John C Flanagan, to construct and distribute to instructors a short questionnaire in order to establish some harder evidence for effective and ineffective training of pilots. He then interviewed experienced pilots asking them why missions had failed or succeeded, what had led to critical situations, what the pilots had done and why their actions had been effective or ineffective. By doing so he reduced opinion, generalisation and personal judgement to a minimum and prepared the basis for a training programme designed to ensure that pilots had the competencies necessary for their flying activities. This factual approach to gathering information became known as the critical incident technique - a collection of information based on first-hand observation (Flanagan, 1954).

An incident was described as "any observable activity sufficiently complete in itself to permit inferences and predictions to be made about the person performing the act".

For an incident to be critical it should occur "in a situation where the purpose of the act seems fairly clear to the observer and where its consequences are sufficiently definite to leave little doubt concerning its effects".

Each incident gathered had to include details of the setting in which the event took place, exactly what occurred, the outcome and why it was considered to lead to effective or ineffective practice. It was crucial that in all cases the focus was on the incident not the individual.

Case-review has a long history in helping to increase understanding in medicine. A structured review of particular cases as a means of learning for informing future cases was the basis of clinical pathological conferences and formed the framework for post-mortems. In all cases the format was similar:

- performance is reviewed systematically.
- features are discovered which contribute to an outcome.
- procedures or practices are modified to avoid or encourage future recurrence.

Such case-based audit, however, was actively discouraged by the Standing Medical Advisory Committee's statement (1990) that this type of review "does not meet the requirements of medical audit". Despite this, however, random case analysis and problem case analysis were being introduced in a formal way to vocational training for general practice. Occurrence screening was a similar method as described by Bennett and Walshc (1990) which used retrospective review of individual cases to highlight errors with a view to limiting exposure to medico-legal risk.

Bradley (1992) was concerned at relying on single cases or events because of their possible subjectivity. He applied the principles of critical incident technique to

attempt to overcome this subjectivity and the term "significant event analysis" was coined. Because of its potentially high emotional tariff it incorporated the feedback approach of Pendleton *et al* (1984) where both positive and negative aspects of behaviour are considered with the former being rewarded as a pre-requisite to constructive advice about perceived deficiencies.

The study by Pringle *et al* (1995) of case-based auditing set out primarily, among other things, to explore the feasibility and acceptability of significant event auditing in a variety of general practice settings. He recruited 20 multiple-partner practices in Lincolnshire and Manchester representing both rural and urban practices respectively. The practices were randomly assigned either to carry out audit over a one year period using the conventional quantitative approach or to using significant event analysis. Each practice was asked to hold a minimum of six audit meetings over one year, regardless of the methods being used. All were required to audit the care of diabetes and doctor availability but in addition could make their own choices from any other clinical and administrative aspects of care. There was no statistical difference in the processes and outcomes of diabetic care between the two groups. The doctor availability audit had to be abandoned due to difficulties with comparisons. The crucial difference was that the conventional arm participated in fewer but more in-depth audits than the significant event arm which covered a much wider range and larger number of topics with less investment of their time. Their conclusions of the study were that this method of single case analysis was both feasible and acceptable in practices of various sizes in both urban and rural settings.

As a result of this study Pringle recommended significant event auditing as complimentary to and not a substitute for more quantitative audit methods. He described its inclusion in an audit programme as balancing "the intellectual and emotional content of performance review" given the high levels of mutual trust

and communication required to be in place for significant event analysis. He recommended that more conventional quantitative methods of audit should be undertaken before introducing the concept of significant event analysis. This is therefore of particular importance in the training of a young doctor.

The role of external facilitation in conducting significant event analysis is controversial. In the above study Pringle recommended an enhanced role for the practice manager who, in most cases, had the trust of the individuals in the practice and, at the same time, could remain dispassionate in conducting such an analysis unless he or she was part of the event being discussed. For many events, however, such facilitation will be unnecessary if the appropriate level of trust among those taking part in the analysis is implicit.

The educational setting is therefore an ideal opportunity to teach and test for an understanding of the principles of significant event analysis. Many of the strengths identified by Pringle as important in significant event auditing are present in a training environment. The audits are often outcomes-focused and deal with practical and relevant issues in day to day practice and cover a wide range of issues often perceived as more difficult to measure than those amenable to quantitative audit methods. The importance of appropriate feedback involves learning to work in teams, not all of which will be fully functional. A registrar in practice is under the protection of his/her trainer and the SHO can seek the advice of an educational supervisor.

What drove Flanagan to suggest and ultimately implement his critical incident technique was the need for objectivity over opinion. He sought facts from real

events and concentrated on these rather than the individuals involved in the events as a source of learning. By considering processes which resulted in a satisfactory outcome as well as less satisfactory results he was able to produce a framework within which he could work to develop fair and constructive decisions about advising on better systems.

The description of a significant event can be likened to the subjective opinion with the analysis demonstrating the benefit of reflection and understanding into the cause of the event. The competence being demonstrated therefore is the ability to learn from and suggest or implement change as a result of an event which happened unexpectedly. The assessment of this competence would therefore offer an opportunity for those in training and in service practice to ensure that they both understood the process of analysing a significant event and if necessary receive feedback on areas of deficiency in such an analysis. Miller's framework of assessment (1990) identifies the components of a doctor's work which take place in real clinical practice as the most difficult to measure accurately and reliably. Assessing what a practitioner "does" (action) as opposed to "knows" (knowledge) will be more clinically authentic, the apex of a pyramid of competence. Significant events are experienced in real time and therefore reflect real clinical practice and can be presented in a written format as a project.

The utility of the assessment system under consideration has to consider the various trade-offs from a number of attributes with their relative weight within the system depending on the purpose intended (Van der Vleuten, 1996). A formative system, such as assessing the analysis of a significant event in a training environment, would benefit from a high validity at the expense of reliability. Validity in this context is difficult to measure as the aim would be to predict whether a satisfactory analysis would increase the chance for a satisfactory event

or reduce the chance for an unsatisfactory event recurring given similar circumstances. The modest reliability found in this system would be considered adequate for providing feedback, with more experience in assessment eventually resulting in a more robust process with higher reliability although the level of agreement between pairings was reasonable.

Acceptability and feasibility both consider the resources available. In order to be acceptable to those involved in implementing the system (assessors and those being assessed) the competence being assessed should be recognised in both a hospital and community setting in the context of training. Although currently assessing only doctors the lessons learned through the educational process should be equally acceptable to teams, both in hospital wards and primary care. One indicator of acceptability is its uptake by clinicians in practice and a steady increase in events being submitted year-on-year has occurred to the current total (March, 2002) of around 400.

Significant events occur frequently in daily practice, are easy to prepare as a report and are quick to assess using the system described. Time and money are important issues and each analysed event took on average 10 minutes to assess. The exercise to check assessor – agreement therefore took between three and four hours of an assessor's time. As all the assessors are working general practitioners time away from practices is an important consideration in assessment.

Feedback is a crucial element in the formative assessment process. If risk reduction is to become part of the educational process it is intuitive that any assessment process linked to this should have similar objectives. Strategic use of an assessment process enhances learning through its *content* (assessing a real event in real practice), its *format* (the criteria against which it will be assessed should be easily understood), through the *information given* (constructive

feedback either endorses a successful analysis or, if necessary, suggests strategies for improvement - itself a form of risk reduction) and through its *programming* (it fits easily into a tight curriculum - for example, a registrar year in general practice or a six-month senior house officer post).

The challenge, particularly in the short to medium term, may be to link significant event auditing with conventional criterion based audit with the former providing an audit needs assessment for a practice with areas of good practice being disseminated widely and lessons learned from deficient practice resulting in opportunities for continually improving quality of care. Pringle (1998) cites one example in his practice which started with a significant event analysis about a sudden death, resulting in a formal audit of the care being provided for ischaemic heart disease followed by a discussion on a health needs assessment for providing adequate care for the risk factors of ischaemic heart disease using such evidence base as was available and ultimately to the cost of providing such care, in this case through commissioning. Although admitting that this case study cannot be generalised and that not all practices will be as committed to or skilled in taking part in such an in-depth review of care the suggested model of using significant events to give context and "emotional relevance" can help to motivate change. The conventional criterion audit method allows for continual reinforcement and checking on progress with the two types of audit method - qualitative and quantitative - integrating in such a way as to maximise the opportunity for best quality care.

CONCLUSION

Training practices have been analysing and discussing significant events since 1995 as part of the programme of audit activities. Registrars have been expected to submit a significant event analysis for formative assessment since August 2001. Results from the last year are currently being evaluated.

By defining audit in qualitative and quantitative terms and by introducing both to a mixed programme within a training environment both in hospital and general practice a more integrated approach to improving quality of care has been offered to training practices in the west of Scotland. The ultimate purpose behind both systems, however, is to ensure that doctors in training for general practice have demonstrated their competence in being able to critically analyse part of their work, as chosen by them, and implement any changes deemed appropriate.

Figure 1 – The development of the process for assessing a significant event analysis

REPOSE	ANALYSIS	ASSESSMENT	SUPPORTIVE STATEMENTS FOR THE ASSESSMENTS
Reasons for the enquiry	What happened?	What makes an event significant?	<ul style="list-style-type: none"> • Has personal impact • Important to an individual or organisation • Causes reflection
A source of ownership			
Presentation of the evidence	Why did it happen?	Is evidence clearly presented?	<ul style="list-style-type: none"> • A clear reason sought
A resulting policy	What have you learned?	Was insight demonstrated?	<ul style="list-style-type: none"> • Awareness of sub-optimal care • Decision-making process altered • Assessment of risk demonstrated • Level of personal responsibility linked to circumstances
Devising a system	As a result, what have you changed?	Was change implemented	<ul style="list-style-type: none"> • Yes – describes implementation of relevant change • No – risk of similar event unlikely • No – unable to influence change but suggestions for change given
An enquiry to ensure compliance with the policy			

Figure 2 – Significant Event Analysis – Assessment Schedule

Name		Project No.	
1. <u>What happened?</u>	<ul style="list-style-type: none"> • Has personal impact • Important to individual or organisation • Causes reflection 	<input type="checkbox"/>	
2. <u>Why did it happen?</u>	<ul style="list-style-type: none"> • Clear reason sought 	<input type="checkbox"/>	
3. <u>Was insight demonstrated?</u>	<ul style="list-style-type: none"> • Aware of previous suboptimal care • Decision-making process altered • Assessment of “risk” demonstrated • Level of personal responsibility linked to circumstances 	<input type="checkbox"/>	
4. <u>Was change implemented?</u>	<ul style="list-style-type: none"> • Yes – Describes implementation of relevant change • No – risk of similar significant event unlikely 	<input type="checkbox"/>	
<u>Satisfactory analysis of significant event</u>		Yes <input type="checkbox"/>	No <input type="checkbox"/>
<u>Comments for feedback (continue overleaf if necessary)</u>			
<p>.....</p> <p>Assessor signature</p> <p>.....</p> <p>Capitals</p>			
		Date	

Table 1 – Summary statistics of marking exercise for proportion of agreements on 100 SEAs by 20 assessors

	mean	median	S.D.	minimum	maximum	Q1	Q3
Average kappa values	0.34	0.34	0.15	0	0.60	0.28	0.48
Average proportions of agreement	0.73	0.77	0.11	0.4	0.83	0.68	0.8

CHAPTER 11

INTRODUCING AUDIT INTO SHO VOCATIONAL TRAINING SCHEMES FOR GENERAL PRACTICE

The General Medical Council (1998c) has recommended that senior house officer (SHOs) vocationally training for general practice should have “an understanding of the principles of clinical audit and self-appraisal”. The Postgraduate Dean has the responsibility for ensuring that this is taking place. Hospital posts in specialties accredited for vocational training for general practice are visited up to five-yearly to ensure that the necessary criteria for training are being implemented. These include an induction course, a named educational supervisor, protected teaching time of a minimum of four hours per week and an understanding of clinical audit (JCPTGP, 1998).

Using a structured questionnaire, confidential feedback from the SHOs at the end of their six month posts consistently suggested that “opportunities for audit” was a problem area to be addressed. 67 hospital posts in ten different specialties in the west of Scotland were visited over a period up to April 1997. Eight units had no evidence of audit whatsoever, 32 units were “doing audit” but with no SHO involvement, 17 units involved SHOs in the collection of data and in ten units the SHOs were “doing an audit project”. (*Kelly M H, personal communication, 1998*).

In a previously unpublished study the author found evidence that consultants’ knowledge of audit method showed wide variation among the specialties accredited for teaching SHOs in the deanery. There was therefore a need for direction in the teaching of audit with guidance given towards an appropriate size of project in the relatively tight timescale which occurs in six-month posts.

An opportunity therefore arose to use the methodologies developed for assessing quantitative and qualitative audit methods and to introduce these into the vocational training schemes. Support from the Postgraduate Dean was given and the output of externally assessed submissions of completed audit cycles and significant event analyses by cohorts of SHOs vocationally training for general practice was monitored for each post in each specialty.

METHOD

Background

This study restricted itself to SHOs on vocational training schemes because they are easily identifiable and therefore more easily followed through their various posts. Approximately two-thirds of all SHOs are "self-constructs" who prepare their own training plan and are consequently even more difficult to identify and follow up.

Between February 1998 and July 2000 all SHOs in hospital posts in vocational training schemes in the west of Scotland were asked to submit an audit project in each of their four 6 month posts. There is a maximum of 58 SHOs per year in the thirteen vocational training schemes in the west of Scotland. Each scheme consists of four six-month posts with two exceptions which are of 18 months duration, giving a potential for 232 SHOs if all posts were filled. During the study SHOs were entering and completing their posts at different stages.

Supporting Material

A teaching pack containing background information on criterion-based audit and significant event analysis was provided. Those SHOs with one post to complete when the study started could choose either a criterion-based "audit cycle", i.e. two

data collections separated by a change in practice or a significant event to analyse. Those with two or more posts had to complete a minimum of one audit cycle and one significant event with their own choice of options for the remaining posts. There was no maximum number of submissions.

Supervision and Instructions

At the beginning of the study period a half-day induction programme was carried out for the SHOs, their educational supervisors and the audit support staff member in the relevant hospital. Local course organisers responsible for supporting vocationally training SHOs were also informed of the programme and advised that thereafter they should provide a similar programme at the beginning of every six-month post. They were asked to record any problems they experienced over the course of the study. A reminder was sent to each SHO two months before the end of each post.

External Assessment

Audit cycles and SEAs were independently assessed by two trained assessors and, if requested in advance by the SHO, the former project could be reviewed as their submission for summative assessment.

For each type of audit, feedback from the assessors was sent to the SHOs to inform them of possible improvements to their project where appropriate. At the end of each post the educational supervisor was provided with an update on the submissions received. This information would then form part of the discussion at the next re-accreditation visit.

The percentage of submissions of completed cycles and significant events from each hospital speciality along with the output from each unit within each specialty were collated.

RESULTS

There was an expected maximum of 341 projects from the 65 units in the six specialties over the study period of 30 months. During this time 138 SHOs were in vocational training posts.

A total of 137 audit projects were received from 92 SHOs, three of whom submitted more than the one project required in a single post. 47 SHOs submitted neither an audit cycle nor a significant event analysis throughout their SHO training.

84 projects (60%) were criterion audit cycles, 41 of which were assessed as being of a standard to pass summative assessment. 53 projects (40%) were analyses of significant events.

Table 1 shows the specialties with their respective number of units. For each unit within each specialty the actual number with the potential number of audit project submissions is shown. The actual and expected totals with their percentages are also shown.

Geriatrics had the best overall percentage of submissions (45% of the possible maximum) and accident & emergency the least (28% of the possible maximum). All specialties provided fewer than half of the expected number of submissions.

Each specialty had at least one unit which did not submit a single project. Only four of the 65 units (two geriatric, one medicine and one obstetrics & gynaecology) submitted the total number of projects expected. Geriatrics therefore supplied the highest percentage of expected submissions. Within the

nine geriatric units, however, one SHO submitted more than what was expected while two other units submitted no projects at all from an expected seven. This variation was seen throughout the units across the specialties.

Submissions for each subsequent cohort (Table 2)

With one exception there was a reduction in the number of SHOs in each six-month cohort submitting a project over the 30 months of the observation period.

Problems Identified

Problems recorded by the course organisers at the end of the study related to three areas:

- educational supervisors frequently changed and were often unknown to the SHO.
- supervision - SHOs were often given conflicting advice on audit method and encouraged to take part in large departmental audits instead which were described by the SHOs as "bean counting". Projects were therefore often started but left unfinished.
- "tracking" SHOs through their posts was a problem thus reducing the opportunity to provide regular feedback to all responsible for educational support.

DISCUSSION

Under Title IV of Council Directive 93/16/EEC the JCPTGP as the competent authority has a duty to ensure that the training it approves satisfies the requirements of all articles of Title IV of the Directive.

The JCPTGP is also responsible for the approval of all training posts for general practice both in hospital and general practice. A list of 13 'Quality Standards' have been approved for the selection and reselection of hospital posts (JCPTGP, 1998). One criterion and standard states that "clinical audit should be in place in all units selected for general practice training". The monitoring of the standards is carried out by the Royal College of General Practitioners on behalf of the Joint Committee through the visiting of SHO posts along with other medical Royal Colleges. This joint hospital visiting model was established in the 1980s to approve SHO posts for general professional training. Recommendations are made to the Education Committee within the deanery and it is up to this Committee to decide whether or not to approve posts for general practice training. They then inform the Joint Committee of their decision. It, in turn, carries out a sample of visits covering all deaneries in the United Kingdom on roughly a three yearly basis where the decisions of the Education Committees are verified.

Such reviews of individual posts may well be superseded by giving more responsibility to Postgraduate Deans to review deaneries to consistent standards overseen by the newly emerging Medical Education Standards Board (MESB) (Department of Health, 2000). Joint working therefore between the various Royal Colleges and the deaneries may well become the norm (Hayden & McKinlay, 2001).

The advantages of project work in an educational context have been described earlier in the thesis (Henry, 1994). The introduction of one quantitative and one qualitative method of audit into SHO posts accredited for vocational training for general practice offered an opportunity to monitor in more detail whether the relevant criterion on clinical audit as advocated by the JCPTGP was being implemented.

There was considerable variation in the submissions of audit projects from the different hospital specialties. Furthermore, individual units within a specialty also showed wide variation. One reason for this variation is the tension between service commitment and educational priorities for SHOs in vocational training posts for general practice. This appears to be confirmed by the low percentage of audits seen from Accident & Emergency where service commitment is considered to be high. In a previous study by Kelly and Murray (1997) Accident & Emergency scored highly for dissatisfaction with the quality of teaching. Lack of protected time for teaching was cited as the main reason.

In a larger study carried out by postal questionnaire in 1997 on a sample of 545 doctors who graduated from medical schools in 1995 a wide variation in the quality of training received by SHOs throughout the United Kingdom was found by Cooke and Hurlock (1999). The response rate was 95% and covered eight specialties accredited for the training of SHOs, including general practice. The minimum protected teaching time received was 2.6 hours per week (Accident & Emergency, Surgery and General Medicine) with a maximum of 5.4 hours in Psychiatry. General practice SHOs received a mean of 3.7 hours per week.

Despite the support for the Postgraduate Dean and all those responsible for ensuring an understanding of audit method in posts accredited for teaching, the number of submissions dropped with each successive six month post - with one exception - once the induction training was undertaken locally. As prior consent had been established from those responsible for providing support for audit it is important to discover why this criterion for training was not being applied in many of the units and how this situation could be remedied.

It is possible that the educational supervisors themselves were unsure of audit method and how it should be taught, a situation similar to that found among the G.P. trainers where prior experience with audit and its teaching had been assumed.

The unpublished survey on "opportunities for audit" in 67 hospital posts in 10 different specialties all accredited for training SHOs in the west of Scotland in April 1997 was followed up by a questionnaire to 127 consultants in these "List A" specialties on their personal experience of clinical audit, knowledge of basic audit method and attitudes towards audit. Replies were received from 72 consultants (response rate of 57%). 42 consultants (58%) had claimed to have completed one cycle of audit. An example of structure of care was correctly identified by 51 consultants (71%), process of care by 47 consultants (65%) and outcome of care 64 consultants (89%). Attitudes were generally positive towards the potential benefits of audit but they were less likely to feel that time spent in clinical audit was time well spent and felt it was less important to involve all relevant team members in the audit. Comments were made that many audits were often left unfinished with more need to focus on smaller, more relevant projects such as those now being received from registrars in practice. The value of verification through an external assessment by assessors experienced in this type of work was therefore important in providing a consistency in the quality of feedback to SHOs and the units.

The small number of significant event analyses being submitted is surprising given the popularity of this format in the hospital environment. Although this might have been due to the potentially sensitive nature of the content this was not borne out by many of the projects submitted, some of which were of an extremely personal nature. It was obvious that many of the SHOs had not had adequate protection, and in some cases counselling, for what were obviously very unpleasant experiences. The assessments of the SEAs allowed for the fact that

SHOs might not be in a position to ensure that change in practice could be effected but an adequate explanation of this had to be given to ensure that insight into the cause of the significant event had been demonstrated.

The regulations for summative assessment allow for the submission of an audit project at any point during the three years of vocational training. With just under half of audit cycles submitted passing the summative assessment process it is likely that there is more support for audit teaching in some units than others, a situation which is unfair to many SHOs. One third of SHOs submitted no projects, the reasons for which are not known. A decision to try to contact them was not taken as the effort required was judged to be too great in view of some of the problems recorded by the course organisers. Identification and follow-up of SHOs are issues which should act as a catalyst for tighter database management between posts to ensure consistent advice and appropriate feedback are given to those SHOs experiencing problems. This will ensure that submission problems can be linked to a particular SHO or a particular unit.

Many of these SHOs are therefore entering the general practice component of their training with no personal experience of carrying out audit at all. This is similar to the figures shown earlier in this thesis highlighting the small number of registrars with prior experience of audit method.

CONCLUSION

Hospital specialties accredited for the training of SHOs for general practice show considerable variation in their output of criterion audit cycles and significant event analyses. There is still therefore a significant number of doctors whose first experience of criterion audit or significant event analysis is the ninth year of their training, i.e. in the registrar year in general practice. Targeted feedback to units of

numbers and the quality of submissions should identify those SHOs and posts where there are problems which can then be addressed by the Postgraduate Dean. Without verification and assessment, assuring the quality of audits being produced may remain speculative.

Table 1 - Actual (with expected) audit project submissions per unit per specialty

A & E (12 units)	Geriatrics (9 units)	Medicine (10 units)	Obs & Gyn (15 units)	Paediatrics (6 units)	Psychiatry (13 units)
0 (2)	0 (1)	0 (2)	0 (4)	0 (1)	0 (2)
0 (2)	0 (6)	1 (4)	1 (4)	1 (4)	0 (1)
0 (5)	1 (2)	1 (4)	1 (6)	2 (11)	1 (3)
1 (3)	1 (2)	1 (2)	1 (3)	3 (5)	2 (4)
1 (4)	2 (1) *	2 (3)	1 (5)	3 (5)	2 (4)
1 (2)	3 (3)	2 (2)	1 (6)	4 (7) *	2 (4)
1 (7)	3 (8)	3 (15)	2 (2)		2 (3)
2 (5)	4 (11)	3 (4)	2 (3)		2 (9)
2 (5)	5 (8)	4 (5) *	2 (5)		3 (8)
2 (3)		4 (7)	3 (5)		3 (5)
2 (6)			3 (8)		4 (12)
3 (10)			4 (13)		4 (8)
			4 (10)		6 (10)
			6 (8)		
			7 (9)		
TOTAL	19 (42)	21 (48)	38 (91)	13 (33)	31 (73)
28%	45%	44%	42%	40%	42%
					137 (341)

* denotes more than one submission by an individual SHO

**Table 2 - Number of audit projects submitted by SHOs over
each six-month post**

Month	1	2	3	4	5
	02/98 – 07/98	08/98 - 01/99	02/99 - 07/99	08/99 – 01/00	02/00 – 07/00
Actual Submissions	34	32	24	35	12
Expected Submissions	42	63	66	87	83
%	76	52	36	51	14

CHAPTER 12

INTEGRATING QUANTITATIVE AND QUALITATIVE AUDIT INTO VOCATIONAL TRAINING FOR GENERAL PRACTICE – FUTURE IMPLICATIONS

This thesis began as a response to the establishment of the JCPTGP criterion and standard for audit which stated that “all training practices must provide opportunities for registrars to become familiar with the principles of medical audit and to participate in medical audit; and they must be able to demonstrate that registrars have actually done so”. Verification in training practices showed that little was happening to meet this criterion. The resulting system of assessment and audit programme were designed and implemented to address this.

If audit was not being demonstrated and taught as part of the training of future general practitioners it was inevitable that competence in this area could not be assumed. Competence was therefore defined using a quantitative and qualitative model based on a written project and assessed externally to the practice using a trained team of assessors. The combination of a quantitative and qualitative approach to understanding audit method and its application to both registrars in a practice environment and SHOs in a hospital environment have encouraged an approach where audit experience is integrated into the training programme for general practice.

National Implications

By the end of their training for general practice a registrar should have demonstrated his/her competence in an “ability to review and critically analyse their working practice and manage any necessary change appropriately”. Between September 1st 1996 and September 1st 2001, 7340 registrars in general practice

throughout the United Kingdom submitted an audit project based on the five criteria described in this thesis. 913 registrars (12.4%) required two attempts to demonstrate this competence. 74 of these (1% of the total) failed after resubmission and had to undergo extra training the length of which was at the discretion of the director of the deanery. A project which failed had had seven independent assessments within the deanery and a further two outwith the deanery before further training was recommended. The implications suggest that experience in the west of Scotland was not an isolated phenomenon and that the teaching of audit method based on actual experience could not be assumed well beyond this deanery. (*M Attwood, personal communication, 2001.*)

The validity of the assessment process developed for audit was endorsed in a systematic review of the published evidence on medical postgraduate certification processes by Hutchinson *et al* (2002). Of the 55 papers identified from 1985 to 2000 in the systematic review only two tested consequential validity, exploring the effect of the assessment process on candidate behaviour. One of these papers is described in detail in chapter four of this thesis. The authors of the review described consequential and construct validity as the two forms of validity considered central in recent general education research in contrast to the emphasis on more quantifiable aspects of reliability.

Predictive validity was addressed in a study of GP non-principals in the west of Scotland by Bowie *et al* (2002). A survey of 200 non-principals achieved a response rate of 79% of whom 67 (42%) had undergone summative assessment in the west of Scotland from 1992 onwards with 91 (58%) trained either in the west of Scotland or elsewhere before summative assessment was introduced UK wide in 1996. Respondents rated their perceived level of knowledge of the different stages of criterion audit method. Significantly higher mean scores were recorded for the post summative assessment group in every area of audit method. 44

respondents (28%) were able to recognise all of the accepted audit criteria statements listed. Significantly more respondents from the post summative assessment group were able to recognise the accepted statements. Minimal statistical difference was noted between groups in their responses to most of the attitude statements listed. Significantly more respondents from the pre summative assessment group required training in audit method ($p < 0.001$). The conclusion was that the criterion audit project embedded in summative assessment had a positive impact on GP non-principals' perceived and actual knowledge of audit method which was sustained long after their training was completed.

In recognition of the problems highlighted by the wording of its 1991 criterion and the increasing evidence from summative assessment the officers of the JCPTGP agreed that two new criteria for audit should be implemented from September 2000. More practical audit required to be in place in training practices if the GMC competence covering clinical audit was to be addressed as would be required for revalidation and clinical governance. Agreement on the wording was reached after consultation with the Conference of General Practice Education Directors (COGPED) and the RCGP. The new criteria are:

- training practices must demonstrate that the audit process is being taught.
- training practices must have in place an active programme of audit which demonstrates the full audit cycle and the application of both standards and criteria.

Verification was to be achieved for the JCPTGP by asking its visiting teams to "actively inspect the new criteria" which would be considered as part of the overall report on a deanery inspection.

In its three-yearly visit to the west of Scotland in September 2001 the JCPTGP commented (JCPTGP, 2001) on "the ethos of audit present in both primary and secondary care in the deanery" as a highlight of the visit. The three visitors spent two full days in four training practices and three hospital departments and reported that "none of the registrars or the trainers that we interviewed had experienced any difficulties with the implementation of the eight criteria audit cycle".

In its updated versions of Good Medical Practice the General Medical Council (1998a & 2001) upgraded its advice on the duties and responsibilities of doctors. It stated that "you must take part in regular and systematic medical and clinical audit and where necessary respond to the results of audit to improve your practice, for example by undertaking further training; you should also take part in confidential enquiries and adverse event recognition and reporting to help reduce risk to patients".

The integration of the audit programme into the teaching process using a combination of qualitative and quantitative methods with an external assessment of competence based on project work ensured that both the JCPTGP and GMC principles were being upheld.

Background Issues - evolving views on audit

In 1980 two editorials highlighted the importance of personal responsibility for medical audit with a need "to convince the sceptics and the silent, indifferent majority that the effort is worthwhile". In its editorial the Lancet (1980a) placed the responsibility for effective audit "above all squarely on the shoulders of all those responsible for postgraduate training in general practice. Without a willing spirit of enquiry, audit is worthless". In a slightly more optimistic vein the British Medical Journal (1980b) stated that: "many general practitioners will want to move at their own pace". By doing so the majority would become convinced of

the cost effectiveness and practical value of audit which "will become routine within a generation".

Fifteen years later more questions were being asked about the value of audit and the lack of evidence base on which its widespread introduction was promoted. Barton *et al* (1995) explored the research questions still requiring to be answered in justifying the time and money being increasingly spent on clinical audit. Further research was described as being both "essential and urgent".

Using more emotional language a short article in the British Medical Journal (Farrell, 1995) stated "the audit cycle has become a vicious circle, a noose to strangle any chance of it ever being a practical everyday tool". In similar terms Professor Sherwood (1992), Dean at the University of Cambridge, wondered whether it was conceivable that "the audit rage will one day look like mass mini chest radiography - worthy, high-minded and useless". Both of these articles prompted the suggestion that perhaps it was time to reappraise clinical audit. In a lecture given to a conference of Chairmen of Health Authorities and Trusts held at the Royal College of Physicians Professor Anthony Hopkins (1996) described audit as "failing to win the hearts and minds of the medical profession". Although the arguments in his lecture were largely based on hospital practice their relevance to general practice was also alluded to. He concluded that the conflict at that time was between clinical audit as a tool for education and professional development and for monitoring contract performance. If the former was to become integrated into daily practice much more attention would need to be paid to inadequate research and the social structures in hospitals and medical schools.

The last five years have been dominated by a discussion on where clinical audit fits in what has been described as "the gritty world of doctors and patients". The importance of a learning culture in an organisation for supporting participation in

audit was recognised by Wedderburn (1998). This theme of organisational change has been described as the key to quality improvement. Berwick *et al* (1992) described four barriers to quality improvement - time, territory, tradition and trust. The quantitative and qualitative aspects of the audit programme addressed these issues in a variety of ways. The importance of time management as part of an efficient and effective appointment system is taught; the importance of teamwork as part of chronic disease management challenges traditional areas of territory jealously guarded by professions - both medical and nursing; all traditions are questioned by the increasing importance of and difficulty with measuring patient satisfaction; and finally the opportunities offered by analysing significant events in engendering trust.

Frustration at the slow progress of change, increasing expenditure, and increasing need for transparency and accountability has focused minds on increasingly complex methods either of delivering or as an alternative to clinical audit in the management of change. Shared values which form the culture of an organisation such as in a practice team are increasingly recognised as important in promoting the quality of health care and improving performance (Davies *et al*, 2000). Clinical audit is described as an example of "single loop learning" with change being suggested but rarely implemented. Higher levels of "double loop learning" or "meta learning" where errors can be detected and corrected in ways that involve modification of an organisation's underlying norms and objectives involve a sharing of values which is rarely seen in many teams in general practice.

Emerging theories of "complexity science" have been suggested (Plsek & Greenhalgh, 2001) as more accurately reflecting the complex world of practice. It recognises that many systems in general practice, such as those responding to patient demand, exist in what has been called the "zone of complexity" where managing uncertainty, the hallmark of much of general practice, will always create

tensions between criterion audit with its relatively rigid structure and the intuitively more attractive "holistic view" such as significant event audit and narrative based practice. The difficulties of implementing such models, however, in a training environment should not be under-estimated as exemplified in one example (Greenhalgh & Eversley, 1999) described as a "post modern approach" to measuring quality of care. The authors acknowledged that the preliminary work in this study was overshadowed by structural re-organisation and "organisational politics within local groups" such as contested or absent leadership and overall confusion.

In addition the increasing reliance on computer systems for delivering quality reports based on valid and reliable data is increasingly seen as crucial but, in Scotland, is relatively under-developed. Despite the perceived advantages of having 82% of practices in Scotland all using GPass software, a recent report (GPass, 2002) stated that "GPass must improve its clinical functionality and usability as a matter of urgency". Part of the review was an external assessment by a team led by Professor Mike Pringle which concluded that "GPass falls short of both the strategic expectations of the Health Service and the needs of general practice". Until these shortcomings are addressed a very limited interpretation of "integrated audit" can be expected for practices - training and non-training - in Scotland.

The last decade began with four standard setting bodies – the GMC, JCPTGP, RCGP and the Government – trying to balance quality and accountability. Each body had its own definition for measuring the quality of care provided by doctors whose competence to do so was assumed.

The decade ended with clinical governance (current government), re-validation (General Medical Council), practice accreditation (RCGP Scotland) and a national

summative assessment system for quality assuring the competence of registrars entering general practice (JCPTGP).

Over the 10 years each institution has moved from leaving it to the profession to ensure that its mandates were implemented to relying on legislation. This is at variance with much of the literature on encouraging and implementing change.

The inability to ensure that the JCPTGP criterion for involving a registrar in audit could be implemented was the instigation for this thesis. At this time medical audit was assumed, without much prior evidence, to be a useful method for assessing quality of care. How it should be done and by whom had not been evaluated. Whether a method was being taught had not been addressed but again was assumed to be taking place in training practices throughout the United Kingdom and quality assured by the JCPTGP. What was being taught and to what standard were unknown.

The strategic use of assessment and a curriculum is well recognised in encouraging specific learning processes. The addition of an audit project submission as part of summative assessment offered the opportunity to both define a quantitative method of audit for doctors in training and to identify those doctors who were not able to analyse a specific aspect of their work.

The move from five to eight criteria has been led by those registrars who would appear to be more confident in going beyond the assessment task expected. Although difficult to quantify the change has occurred on approximately a three-year cycle with the next step being to include an assessment of the ability to calculate proportions and confidence intervals. Initial surveys of both trainers and registrars in the west of Scotland have shown a complete lack of confidence and knowledge base of statistical method despite all receiving undergraduate teaching

some years ago. Successful implementation of a more formal and rigorous approach to criterion audit will finally address the advice of many commentators on audit method in the early 1990s on the importance of statistical input. The more pragmatic approach taken however has ensured that training practices have had the on-going support required in response to the wide variation in their own confidence in teaching audit methodology.

The implementation of the five-year audit programme offered an opportunity to carry out a randomised control trial of whether audit methodology did offer an adequate method of assessing quality of care. Small individual audits have undoubtedly shown this but no formal RCT on such a scale has been found in the literature. The implication of carrying out such an exercise, however, was that approximately half the practices would be acting as controls and their responsibility in ensuring that the JCPTGP criterion for audit was being implemented could not be assured. Those practices chosen as controls would thus be vulnerable to a JCPTGP visit and in addition the workload involved in evaluating such a trial would have been beyond the remit of the author.

The programme has resulted in a more balanced mix of quantitative and qualitative audit of practical significance for a doctor in training and the initial hostility to such a programme has largely dissipated. The increasing importance of adequate IT systems in supporting audit has been highlighted and the possible implementation of a new general practitioner contract has ensured that training practices in the west of Scotland are well placed to deliver on the quality criteria and standards which will be expected of them.

Approximately two thirds of the practices in the west of Scotland are non-training and their doctors' competence in understanding audit method has not been tested. Some of the outcomes of the research in this thesis, in particular a peer review

system for assessing a quantitative and qualitative audit project, are being accepted by such doctors. Early evidence from the results of assessing these doctors' understanding of criterion and significant event audit gives some cause for concern. Feedback however can be provided and problems addressed where necessary. With the addition of practice accreditation for all practices by 2004 and an appraisal process forming the main constituent of revalidation in Scotland, each of which places a heavy emphasis on an understanding of criterion and significant event audit, a culture of peer review is slowly being accepted and encouraged.

The next phase of this work is to explore methods of collecting and analysing significant events in real practice time and to explore their subsequent use as a form of needs assessment for criterion audit. Involving the practice team will be vital and the model has already been taught to nurses in the Greater Glasgow area.

In conclusion, the integration of a defined quantitative and qualitative audit method into training practices in the west of Scotland has had implications beyond the training environment. Increasing discussion through local health care co-operatives and an encouragement to increase educational assessment through an external review process should ensure that doctors in the west of Scotland are at least addressing the expectations of the government, General Medical Council, RCGP and the JCPTGP in having to demonstrate their competence in two important methods of clinical audit.

REFERENCES

ANONYMOUS. 1980a. Editorial. *The Lancet* (i) p 23-24.

ANONYMOUS. 1980b. Editorial. *British Medical Journal*; 281: p1375-1376.

BAKER, R. (1985) Comparison of standards in training and non-training practices, *Journal of the Royal College of General Practitioners*, 35, pp.330-332.

BAKER, R. (1988) Practice assessment and quality of care, *Occasional Paper* 39, (London, Royal College of General Practitioners).

BAKER, R. (1990a) Experience in medical audit: general practice, *Postgraduate Medical Journal*, 66 (Supplement 3), pp.S14-S16.

BAKER, R. (1990b) Problem solving with audit in general practice. *British Medical Journal*, 300, pp.378-380.

BAKER, R. (1990c) Development of a questionnaire to assess patients' satisfaction with consultations in general practice, *British Journal of General Practice*, 10, pp.487-490.

BAKER, R. (1991) The reliability and criterion validity of a measure of patients' satisfaction with their general practice, *Family Practice*, 8, pp.171-177.

BAKER, R. (1992) General practice in Gloucestershire, Avon and Somerset: explaining variations in standards, *British Journal of General Practice*, 42, pp.415-418.

BAKER, R. & GREEN, L. (1990) Quality assurance for family practice trainees in the USA and the UK: too little too late? *Medical Education*, 24, pp.258-263.

BAKER, R. & PRESLEY, P. (1990) *The Practice Audit Plan*, (Bristol, Severn Faculty of The Royal College of General Practitioners).

BAKER, R. & WHITEFIELD, M. (1992) Measuring patient satisfaction: a test of construct validity, *Quality in Health Care*, 1, pp.104-109.

BAKER, R. & FRASER, R. C. (1995) Development of review criteria: linking guidelines and assessment of quality, *British Medical Journal*, 311, pp.370-373.

BAKER, R. & FRASER, R. C. (1996) Evidence of effectiveness of audit in improving care is growing, Letter, *British Medical Journal*, 313, p.1405.

BAKER, R. & STREATHFIELD, J. (1995) What type of general practice do patients prefer? Exploration of practice characteristics influencing patient satisfaction, *British Journal of General Practice*, 45, pp.654-659.

BAKER, R. & THOMPSON, J. (1995) Innovation in general practice: is the gap between training and non-training practices getting wider? *British Journal of General Practice*, 45, pp.297-300.

BAKER, R., FRENCH, D., LAKHANI, M. & KHUNTI, K. (1994) Audit protocols, in: *Patients' Charter Protocols*. Eli Lilly National Clinical Audit Centre, University of Leicester.

BAKER, R., ROBERTSON, L. & FAROOQI, A. (1995) Audit in general practice: factors influencing participation, *British Medical Journal*, 311, pp.31-34.

BALINT, M. (1957) The doctor, his patient and the illness. (London, Tavistock).

BARBACCIA, J. C. (1976) Introducing quality assurance and medical audit into the UCSF Medical Centre curriculum, *Journal of Medical Education*, 51, pp.386-391.

BARTON, A. G. & THOMSON, R. G. (1993) Is audit bad research? *Audit Trends*, 1, pp.51-53.

BARTON, A., THOMSON, R. & BHOPAL, R. (1995) Clinical audit: more research is required, *Journal of Epidemiology and Community Health*, 49, pp.445-447

BENNETT, J. & WALSHE, K. (1990) Occurrence screening as a method of audit, *British Medical Journal*, 300, pp.1248-1251.

BENETT, J. J. & HAYDEN, J. (1995) Audit as part of summative assessment, *British Journal of General Practice*, 45, pp.47-49.

BERK, R. A. (1980) Criterion referenced measurement: the state of the art, *John Hopkins University Press*.

BERK, R. A. (1986) Minimum competency testing: Status and potential, in: **P. S. PLAKE & J. C. WITT (Eds)** *The future of testing*. (New Jersey, Erlbaum).

BERWICK, D. M. (1989) Continuous improvement as an ideal in health care. *New England Journal of Medicine*, 320, pp.53-56.

BERWICK, D. M., ENTHOVEN, A. & BUNKER, J. (1992) Quality Management in the NHS: The Doctor's Role – II, *British Medical Journal*, 304, pp.304-308.

BHOPAL, R. S. & THOMSON, R. (1991) A form to help learn and teach about assessing and evaluating medical audit papers, *British Medical Journal*, 303, pp.1520-1522.

BIGGS, J. S. G. (1995) Medical audit and medical education, *Medical Education*, 29, pp.99-100.

BOWIE, P., GARVIE, A., OLIVER, J. & LOUGH, M. (2002) Impact on non-principals in general practice of the summative assessment audit project, *Education for Primary Care*, 13, pp.356-361.

BRADLEY, C. (1992) Turning anecdotes into data - the critical incident technique, *Family Practice*, 9, pp.98-103.

BRITISH MEDICAL ASSOCIATION. (1977) Evidence of Council to the Royal Commission of the National Health Service, *British Medical Journal*, i., pp.299-334.

BRITISH MEDICAL ASSOCIATION. (1979 & 1989) *Reports of the Conference of Representatives of Local Medical Committees*, (London, BMA).

BROOKFIELD, S. D. (1986) *Understanding and facilitating adult learning*. (Buckingham, Open University Press).

BUCKLEY, G. (1990) Clinically significant events, in: **M. MARINKER (Ed)**, *Medical Audit in General Practice*, (London, British Medical Journal).

BUCKLEY, G. (1995) Projects revisited, *Education for General Practice*, 6, pp.7-9.

BUCKNALL, C. E., ROBERTSON, C., MORAN, F. & STEVENSON, R. D. (1992) Improving management of asthma: Closing the loop or progressing along the audit spiral, *Quality in Health Care*, 1, pp.15-20.

CALMAN, K. C. & DOWNIE, R. S. (1988) Education and training in medicine, *Medical Education*, 22, pp.488-491.

CAMPBELL, L. M., HOWIE, J. G. R. & MURRAY, T. S. (1993) Summative Assessment: A Pilot Project in the West of Scotland, *British Journal of General Practice*, 43, pp.430-434.

CAMPBELL, L. M., HOWIE, J. G. R. & MURRAY, T. S. (1995) Use of video taped consultations in summative assessment of trainees in general practice, *British Journal of General Practice*, 45, pp.137-141.

CAMPION, P., STANLEY, I. & HADDLETON, M. (1992) Audit in general practice: students and practitioners learning together, *Quality in Health Care*, 1, pp.114-118.

CARMINES, E. G. & ZELLER, R. A. (1979) *Reliability and validity assessment*, (London, Sage Publications).

CHAMBERS, R. & BOWYER, S. (1993) *An evaluation of the quality of medical audit in primary care* (University of Keele).

CHAMBERS, R., BOWYER, S. & CAMPBELL, I. (1996) Investigation into the attitudes of general practitioners in Staffordshire to medical audit, *Quality in Health Care*, 5, pp.13-19.

CHASTONAY, P., BRENNER, E., PEEL, S. & GUILBERT, J - J. (1996) The need for more efficacy and relevance in medical education, *Medical Education*, 30, pp.325-328.

CHEATER, F. M. & KEANE, M. (1998) Nurses' participation in audit: a regional study, *Quality in Health Care*, 7, pp.27-36.

COHEN COMMITTEE. (1950) *General practice and the training of the general practitioner*, (London, British Medical Association).

COLES, C. (1990) Making audit truly educational, *Postgraduate Medical Journal*, 66 (Supplement 3), p.S32-S36.

COLES, C. (1998) How students learn: The process of learning, in: **B. JOLLY, L REES** (Eds), *Medical Education in the Millennium*, pp.63-82 (Oxford University Press).

COLLEGE OF GENERAL PRACTITIONERS. (1953) *First Annual Report*, (London, College of General Practitioners).

COLLEGE OF GENERAL PRACTITIONERS. (1965) Special Vocational Training for General Practice, *Report from General Practice 1*, (London, Journal of the College of General Practitioners).

COLLEGE OF GENERAL PRACTITIONERS. (1966) Evidence of the College of General Practitioners to the Royal Commission on Medical Education, *Report from General Practice 5*, (London, Journal of the College of General Practitioners).

COOKE, L. & HURLOCK, S. (1999) Education and training in the senior house officer grade: results from a cohort study of United Kingdom medical graduates, *Medical Education*, 33, pp.418-423.

CRAWLEY, H.S. & LEVIN, J.B. (1990) Training for General Practice: A National Survey, *British Medical Journal*, 300, pp.911-915.

CROMBIE, I.K. & DAVIES, H.T.O. (1992) Towards good audit, *British Journal of Hospital Medicine*, 48, pp.182-185.

CROMBIE, I.K. & DAVIES, H.T.O. (1993) Missing link in the audit cycle, *Quality in Health Care*, 2, pp.47-48.

CRONBACH, L.J. (1971) Test validation, in: **R.L.THORNDIKE (Ed)**, *Educational Measurement*, pp.443-507. (Washington, American Council on Education).

DAUPHINEE, D. (1994) Determining the content of certification examinations, in: **D. NEWBLE, B. JOLLY & R WAKEFORD (Eds)**, *The certification and re-certification of doctors: issues in the assessment of clinical competence*, pp.92-104 (Cambridge University Press).

DAVIES, H.T.O, NUTLEY, S.M. & MANNION, R. (2000) Organisational culture and quality of health care, *Quality in Health Care*, 9, pp.111-119.

DEPARTMENT OF HEALTH AND SOCIAL SECURITY. (1987) Welsh Office; Department of Health and Social Services, Northern Ireland; Scottish Home and Health Department. *Promoting better health: The Government's programme for improving primary health care.* Cmnd 249, (London, HMSO).

DEPARTMENT OF HEALTH. (1989a) *Working for patients, NHS White Paper*, Cmnd 555, (London, HMSO).

DEPARTMENT OF HEALTH. (1989b) *Medical Audit*, NHS Review Working Paper No.6, (London, HMSO).

DEPARTMENT OF HEALTH. (1989c) *Medical Audit*, Scottish Working Paper 2, (Edinburgh, HMSO).

DEPARTMENT OF HEALTH. (2000) *The NHS Plan: A Plan for Investment, A Plan for Reform*, (Norwich, Department of Health).

DERRY, J., LAWRENCE, M., GRIEW, K., ANDERSON, J., HUMPHREYS, J. & PANDHER, K.S. (1991) Auditing audits: the method of Oxfordshire Medical Audit Advisory Group, *British Medical Journal*, 303, pp.1247-1249.

DERRY, J. (1993) Sample size for audit. *Managing audit in general practice*, 1(2), pp.17-20.

DEWEY, J. (1916) *Democracy and Education*, (New York, McMillan).

DIFFORD F. (1990) Defining essential data for audit in general practice, *British Medical Journal*, 300, pp.92-94.

DONABEDIAN, A. (1980) The measurement of quality: a conceptual explanation, in: *Explorations in Quality Assessment and Monitoring*. Volume 1: The Definition of Quality and Approaches to Assessment, ANN ARBOR, (Michigan, Health Administration Press).

DONABEDIAN, A. (1981) Advantages and limitations of explicit criteria for assessing the quality of health care, *Milbank Memorial Fund Quarterly*, 59, pp.99-105.

DRESSEL, P.L. & THOMPSON, M.M. (1973) *Independent Study*, (San Francisco, Jossey-Bass).

EBEL, R.L. (1961) Must all tests be valid? *American Psychologist*, 16, pp.640-647.

EDITORIAL. (1979) Trainee Projects, *Journal of the Royal College of General Practitioners*, 29, pp.453-455.

ENTWISTLE, N.J. (1987) *Styles of learning and teaching: an integrated outline of educational psychology for students, teachers and lecturers*, 2nd Edition, (Chichester, John Wiley & Sons).

FALCHIKOV, N. & BOUD, D. (1989) Student self assessment in higher education: a meta-analysis, *Review of Educational Research*, 59, pp.395-430.

FARRELL, L. (1995) Audit my shorts, *British Medical Journal*, 311, p.1171.

FIRTH-COZENS, J. (2001) Multi-disciplinary teamwork: the good, bad and everything in between, *Quality in Health Care*, 10, pp.65-69.

FLANAGAN, J.C. (1954) The critical incident technique, *Psychological Bulletin*, 51, pp.327-358.

FLEISS, J.L. (1981) *Statistical methods for rates and proportions*. 2nd Edition. pp.217-234, (Wiley, New York).

FOWKES, F.G.R. (1982) Medical audit cycle, *Medical Education*, 16, pp.228-238.

FRASER, R.C. (1982) Trainee projects and their relationship to audit and research, *Trainee*, 2, pp.18-21.

FULTON, R.A. (1996) Goals and methods of audit should be re-appraised, Letter, *British Medical Journal*, 312, p.1103.

GAMBRILL, D.E. (1985) Projects, Audit and Research, *Journal of the Association of Course Organisers*, 10, pp.1621-1644.

GENERAL MEDICAL COUNCIL. (1993) *Tomorrow's Doctors: recommendations on undergraduate medical education*, (London, General Medical Council).

GENERAL MEDICAL COUNCIL. (1995) *Good Medical Practice*, Para 6, (London, GMC).

GENERAL MEDICAL COUNCIL. (1998a) *Good Medical Practice*, Para 7, (London, GMC).

GENERAL MEDICAL COUNCIL. (1998b) *Maintaining Good Medical Practice*, P.8, (London, GMC).

GENERAL MEDICAL COUNCIL. (1998c) *The Early Years: Recommendations on Senior House Officer Training*, (London, GMC).

GENERAL MEDICAL COUNCIL. (2001) *Good Medical Practice*, P.12, (London, GMC).

GIPPS, C.V. (1994) *Beyond testing: towards a theory of educational assessment*, (London, Falmer Press).

GLASER, R. (1963) Instructional technology and the measurement of learning outcomes, *American Psychologist*; 18, pp.519-521.

GPASS REVIEW GROUP. (2002) *Report for GPass Programme Board*. May.

GRANT, J.A., NAPIER, A., STEPHEN, S., FAIRFIELD, G., CAMPBELL, L. & GRIFFIN, E. (1998) Integrating audit into primary care: a Tayside initiative, *Health Bulletin*, 56(5), pp.822 – 827.

GRAY, D.J.P. (1984) Selecting general practitioner trainers, *British Medical Journal*, 288, pp.195-198.

GRAY, D.J.P. (1985) Trainees who publish projects, in: Trainee Projects. *Occasional Paper 29*, (London, RCGP).

GRAY, D.J.P. (1990) The College Plan - Priorities for the Future, *Occasional Paper 49*, (London, RCGP).

GREENHALGH, T. & EVERSLEY, J. (1999) Quality in General Practice: Towards a holistic approach, (London, Kings Fund).

HAYDEN, J. & MCKINLAY, D. (2001) Monitoring standards of training, *Medical Education*, 35, pp.68-72.

HEANEY, D.J., HOWIE, J.G.R. & PORTER, A.M.D. (1991) Factors influencing waiting times and consultation times in general practice, *British Journal of General Practice*, 41, pp.315-319.

HEARNshaw, H., REDDISH, S., CARLYLE, D., BAKER, R. & ROBERTSON, N. (1998) Introducing a quality improvement programme to primary health care teams, *Quality in Health Care*, 7, pp.200-208.

HEATH, R. (1964) *The reasonable adventurer*, (Pittsburgh, University of Pittsburgh).

HENRY, J. (1994) *Teaching through projects*, Open and Distance Learning Series. (London, Kogan Page).

HILGARD, E.R. (1962) *Theories of learning*, (New York, Methven).

HONIGSBAUM, F. (1972) Quality in General Practice. A commentary on the quality of care provided by general practitioners, *Journal of the Royal College of General Practitioners*; 22, pp.429-451.

HOPKINS, A. (1996) Clinical audit: time for a reappraisal? *Journal of the Royal College of Physicians of London*, 30, pp.415-424.

HOWIE, J.G.R. (1985) Research Protocols, in: Trainee Projects, *Occasional Paper 29*, (London, Royal College of General Practitioners).

HUTCHINSON, L., AITKEN, P. & HAYES, T. (2002) Are medical postgraduate certification processes valid? A systematic review of the published evidence, *Medical Education*, 36, pp.73-91.

INSTITUTE OF MEDICINE. (1990). **K. LOHR (Ed)**, *Medicare: A strategy for quality assurance*, Volume 1, (Washington, National Academy Press).

INSTITUTE OF MEDICINE. (1992) *Guidelines for Clinical Practice: From development to use*, **M.FIELD & K.N.LOHR (Eds)**, (Washington, National Academy Press).

IRBY, D.M. (1994) What clinical teachers in medicine need to know, *Academic Medicine*, 69, pp.333-342.

IRVINE, D.H. (1983) Quality of care in general practice: our outstanding problem, *Journal of the Royal College of General Practitioners*, 33, pp.521-523.

IRVINE, D.H. (1990) Standards in General Practice: The Quality Initiative Revisited, *British Journal of General Practice*, 40, pp.75-77.

IRVINE, D. (1993) General Practice in the 1990s: A personal view on future developments, *British Journal of General Practice*, 43, pp.121-125.

IRVINE, D.H. & JEFFREYS, M. (1971) British Medical Association planning unit survey of general practice, 1969. *British Medical Journal*, 4, pp.535-543.

IRVINE, D.H., GRAY, D.J.P. & BOGLE, I.G. (1990) Vocational Training for General Practice: The meaning of "Satisfactory Completion", Letter, *British Journal of General Practice*, 40, pp.434.

IRVINE, D.H. & IRVINE, S. (Eds), (1991) *Making sense of audit*, (Radcliffe Medical Press).

JOINT COMMITTEE ON POSTGRADUATE TRAINING FOR GENERAL PRACTICE. (1976) *Criteria for the selection of trainers in general practice*, (London, JCPTGP).

JOINT COMMITTEE ON POSTGRADUATE TRAINING FOR GENERAL PRACTICE. (1980) *Criteria for the selection of trainers in general practice*. (London, JCPTGP).

JOINT COMMITTEE ON POSTGRADUATE TRAINING FOR GENERAL PRACTICE. (1982a) *The training year in general practice*, in: *Training for General Practice*, (London, JCPTGP).

JOINT COMMITTEE ON POSTGRADUATE TRAINING FOR GENERAL PRACTICE. (1982b) *Accreditation of deaneries and schemes for vocational training in general practice: General Guidance*, (London, JCPTGP).

JOINT COMMITTEE ON POSTGRADUATE TRAINING FOR GENERAL PRACTICE. (1985) *Recommendations to regions for the establishment of criteria for the approval and re-approval of trainers in general practice*, (London, JCPTGP).

JOINT COMMITTEE ON POSTGRADUATE TRAINING FOR GENERAL PRACTICE. (1992a) *Recommendations to regions for the establishment of criteria for the approval and re-approval of trainers in general practice*, (London, JCPTGP).

JOINT COMMITTEE ON POSTGRADUATE TRAINING FOR GENERAL PRACTICE. (1992b) *Assessment Working Party. Interim Report*. (London, JCPTGP).

JOINT COMMITTEE ON POSTGRADUATE TRAINING FOR GENERAL PRACTICE. (1993) *Report of Summative Assessment Working Party*, (London, JCPTGP).

JOINT COMMITTEE ON POSTGRADUATE TRAINING FOR GENERAL PRACTICE. (1998) *Recommendations on the selection and re-selection of hospital posts for general practice training*, (London, JCPTGP).

JOINT COMMITTEE ON POSTGRADUATE TRAINING FOR GENERAL PRACTICE. (2001) *Review of Vocational Training. Visit Report on the West of Scotland Deanery, 18-20 September 2001*, (London, JCPTGP).

JOSS, R. & KOGAN, M. (1995) *Advancing Quality: A Study of Total Quality Management in the National Health Service*, (Buckingham, Open University Press).

KAISER, M. (1960) The application of an electronic computer to factor analysis. *Educational and Psychological Measurement*, 20, pp.141-151.

KELLY, D. & MURRAY, T.S. (1997) An assessment of hospital training for general practice in the west of Scotland, *Education for General Practice*, 8, pp.220-226.

KERRISON, S., BACKWOOD, T. & BUCKSTON, M. (1993) *Medical Audit - Taking stock*, (London, King's Fund Centre).

KILPATRICK, W.H. (1918) The Project Method. *Teacher's College Record*, 19, pp.319-335.

KIRKLEES METHOD. (1993) Supplement to Kirklees MAAG. *Audit Gazette*, 8 June.

KOLB, D. A. (1984) *Experiential learning: Experience as the source of learning and development*, (Englewood Cliffs, New Jersey, Prentice Hall).

KRATKY, A.P. (1977) An audit of the care of diabetics in one general practice, *Journal of the Royal College of General Practitioners*, 27, pp.536-543.

LANDIS, R.J. & KOCH, G.G. (1977) The measurement of observer agreement for categorical data, *Biometrics*, 33, pp.159-174.

LAWRENCE, M., GRIEW, K., DERRY, J., ANDERSON, J. & HUMPHREYS, J. (1994) Auditing audits: use and development of the Oxfordshire medical audit advisory group rating system, *British Medical Journal*, 309, pp.513-516.

LAWSON, J.A.R. (1992) Royal College of General Practitioners: Its growth and influence over the past 40 years, *British Journal of General Practice*, 42, pp.210-211.

LEMBCKE, P.A. (1956) Medical auditing by scientific methods illustrated by major female pelvic surgery, *Journal of the American Medical Association*, 162, pp.646-655.

LERVY, B., WAREHAM, K. & CHEUNG, W.Y. (1994) Practice characteristics associated with audit activity: A Medical Audit Advisory Group Survey, *British Journal of General Practice*, 44, pp.311-314.

LISTER, S., WORSWICK, L., NELIGAN, M. & FEARN, G. (1998) A new way to assess audit activity - the Camden and Islington MAAG "AAA tool", *Audit Trends*, 6, pp.122-125.

LOUGH, J.R.M. (1996) Regions should define audit strategy, Letter, *British Medical Journal*, 313, pp.497.

LOWRY, S. (1993) Teaching the teachers, *British Medical Journal*, 306, pp.127-130.

LYONS, C. & GUMPERT, R. (1990) Medical audit data: Counting is not enough, *British Medical Journal*, 300, pp.1563-1566.

MARINKER, M. (Ed). (1990) Standards, in: *Medical Audit in General Practice*, pp.12-13, (London, The MSD Foundation).

MARSON, W.S., MORRELL, D.C., WATKIN, C.J. & ZANDER, L.I. (1972) Measuring the quality of general practice, *Journal of the Royal College of General Practitioners*, 23, pp.23-31.

MAXWELL, R.J. (1984) Quality assessment in health, *British Medical Journal*, 288, pp.1470-1472.

MAXWELL, R.J. (1992) Dimensions of quality revisited: from thought to action, *Quality in Health Care*, 1, pp.171-177.

MCCOWAN, C., NEVILLE, R.G., CROMBIE, I.K., CLARK, R.A. & WARNER, F.C. (1997) The facilitator effect: results from a 4-year follow up of children with asthma, *British Journal of General Practice*, 47, pp.156-160.

MCINTYRE, N. & POPPER, H. (1983) The critical attitude in medicine: The need for a new ethics, *British Medical Journal*, 287, pp.1919-1923.

MCKAY, J., LOUGH, J.R.M. & MURRAY, T.S. (2002) Registrar audit and implementation of change, *Education for Primary Care*, 13, pp.336-339.

MERRISON COMMITTEE. (1975) *Report of the Committee of Enquiry into the regulation of the medical profession*, Cmd 6018, (London, HMSO).

MILLER, G.E. (1990) The assessment of clinical skills/competence/performance, *Academic Medicine*, 65, pp.S63-S67.

MILNE, R.M, TAYLOR, M.W. & DUNCE, R. (1991) An assessment of computing activities by GPass users in Scottish General Practice, *Health Bulletin*, 303, pp.152-160.

MORGAN, A. (1982) Project-based learning. IET: *Teaching and Consultancy Monograph No.8*, (Milton Keynes, Open University Press).

MORRISON, J.M. & SULLIVAN, F.M. (1993) Audit: teaching medical students in general practice, *Medical Education*, 27, pp.495-502

MOSS, P.A. (1992) Shifting conceptions of validity in educational measurement: Implications for performance assessment, *Review of Educational Research*, 62(3), pp.229-258.

MUKHERJI, P.S. (1979) Trainee Research in South East Scotland, *Journal of the Royal College of General Practitioners*, 29, pp.466-467.

MULHOLLAND, H. & TOMBLESON, P.M.J. (1990) Assessment of the general practitioner, *British Journal of General Practice*, 40, pp.252-254.

NATIONAL AUDIT OFFICE. (1994) *Auditing clinical care in Scotland*, (London, HMSO).

NATIONAL HEALTH SERVICE (Vocational Training) Regulations. (1979) (London, HMSO).

NATIONAL HEALTH SERVICE (Vocational Training for General Medical Practice), (1997) Regulation 9(2). *Statutory Instrument 3150*, (London, HMSO).

NELSON, A.R. (1976) Orphan data and the unclosed loop: a dilemma in PSRO and medical audit, *New England Journal of Medicine*, 295, pp.617-619.

NEUFELD, V.R. & NORMAN, G.R. (1985) *Assessing clinical competence: Volume 7*, (New York, Springer).

NEVILLE, R.G. & SOWERBY, R. (1988) The role of undergraduate project work in clinical audit in general practice, *Medical Teacher*, 9, pp.473-477.

NEWBLE, D. & JAEGER, K. (1983) The effect of assessments and examinations on the learning of medical students, *Medical Education*, 17, pp.165-171.

NHS CIRCULAR. (1989a) *Medical Audit in the Family Practitioner Services for England and Wales*.

NHS CIRCULAR. (1989b) *Guidance on Implementation of Medical Audit*. (GEN 29) Scottish Home and Health Department.

NYGREN, E., WYATT, J.C. & WRIGHT, P. (1998) Helping clinicians to find data and avoid delays, *The Lancet*, 352, pp.1462-1466.

PATRICK, J. (1992) *Training: Research and practice*, (London, Academic Press).

PENDLETON, D., SCHOFIELD, T., TATE, P., & HAVELOCK, P. (1984) *The consultation: an approach to learning and teaching*, (Oxford, Oxford Medical Publications).

PENDLETON, D., SCHOFIELD, T. & MARINKER, M. (Eds.) (1986) in: *pursuit of quality*, (London, Royal College of General Practitioners).

PETERS, R.S. (1967) *The concept of education*, (London, Routledge and Kegan Paul).

PETERSON, P. (1973) Teaching peer review, *Journal of the American Medical Association*, 224, pp.884-885.

PICKERING, G. (1978) *Quest for excellence in medical education*, (Oxford, Oxford University Press).

PLSEK, P.F. & GREENHALGH, T. (2001) The challenge of complexity in health care, *British Medical Journal*, 323, pp.625-628.

POPHAM, W.J., CRUSE, K.L., RANKIN, S.C., SANDIFER, P.D. & WILLIAMS, P.L. (1985) Measurement-driven instruction: It's on the road, *Phi Delta Kappa*, 66, pp.628-634.

POWSNER, S.M., WYATT, J.C. & WRIGHT, P. (1998) Opportunities for and challenges of computerisation, *The Lancet*, 352, pp.1617-1621.

PRESTON – WHYTE, M.E., FRASER, R.C. & MCKINLEY, R.K. (1993) Teaching and assessment in the consultation: A workshop for general practice clinical teachers, *Medical Teacher*, 15, pp.205-210.

PRINGLE, M., BRADLEY, C., CARMICHAEL, C., WALLIS, H. & MOORE, A. (1994) A survey of attitudes to and experience of medical audit in general practice: implications for MAAGs, *Audit Trends*, 2, pp.9-13.

PRINGLE, M., BRADLEY, C.P., CARMICHAEL, C.M., WALSH, H. & MOORE, A. (1995) Significant event auditing: a study of the feasibility and potential of case-based auditing in primary medical care, *Occasional Paper 70*, (London, RCGP).

PRINGLE, M. (1998) Preventing ischaemic heart disease in one general practice: from one patient, through clinical audit, needs assessment and commissioning into quality improvement, *British Medical Journal*, 317, pp.1120-1124.

RADEMOCHER, M. (1993) Medical audit: need for a positive commitment, *Clinical and Experimental Dermatology*, 18, pp.17-20.

ROGERS, B. (1983) *Diffusion of Innovations*, 3rd Edition, (New York, Free Press).

RONALDS, C., DOUGLAS, A., GRAY, D.J.P. & SELLY, P. (Eds.) (1981) Fourth National Trainee Conference, *Occasional Paper 18*, (London, RCGP).

ROTH, I. (Ed.) (1990) *Introduction to Psychology*, Volumes 1 and 2, (Hove, Lawrence Erlbaum Associates, with the Open University).

ROYAL COLLEGE OF GENERAL PRACTITIONERS. (1972) Teaching Practices, *Report from General Practice 15*, (London, RCGP).

ROYAL COLLEGE OF GENERAL PRACTITIONERS. (1977) Evidence to the Royal Commission of the National Health Service, *Journal of the Royal College of General Practitioners*, 27, pp.197-206.

ROYAL COLLEGE OF GENERAL PRACTITIONERS. (1978a) Practice activity analysis 5: referrals to specialists, *Journal of the Royal College of General Practitioners*, 28, pp.251-252.

ROYAL COLLEGE OF GENERAL PRACTITIONERS. (1978b) Practice activity analysis 6: visiting profiles, *Journal of the Royal College of General Practitioners*, 28, pp.316-318.

ROYAL COLLEGE OF GENERAL PRACTITIONERS. (1983) The Quality Initiative - Summary of Council Meeting, *Journal of the Royal College of General Practitioners*, 33, pp.523-524.

ROYAL COLLEGE OF GENERAL PRACTITIONERS. (1984-5) Quality in Practice Bulletins, Nos.1-16, (London, RCGP).

ROYAL COLLEGE OF GENERAL PRACTITIONERS. (1985a) Quality in General Practice, *Policy Statement 2*, (London, RCGP).

ROYAL COLLEGE OF GENERAL PRACTITIONERS. (1985b) What sort of doctor? *Report from General Practice 23*, (London, RCGP).

ROYAL COLLEGE OF GENERAL PRACTITIONERS. (1985c) Towards Quality in General Practice, *Council Consultation Document*, (London, RCGP).

ROYAL COLLEGE OF GENERAL PRACTITIONERS. (1985d) Trainee Projects, *Occasional Paper 29*, (London, RCGP).

ROYAL COLLEGE OF GENERAL PRACTITIONERS. (1987) The Front Line of the National Health Service, *Report from General Practice 25*, (London, RCGP).

ROYAL COLLEGE OF GENERAL PRACTITIONERS. (1989) *Summary Statement on Working for Patients*, Press statement, 9 April.

ROYAL COLLEGE OF GENERAL PRACTITIONERS. (1990) Fellowship by Assessment, *Occasional Paper 50*, (London, RCGP).

RUSSELL, I.T. & WILSON, B.J. (1992) Audit: the third clinical science? *Quality in Health Care*, 1, pp.51-55.

RUST, J. & GOLOMBOKS, S. (1989) *Modern Psychometrics. The science of psychological assessment*, (London, Routledge).

SAS INSTITUTE. *Statistical Analysis Software*, (Cary, North Carolina, USA).

SCHOFIELD, T.P.C. & HASLER, J.C. (1984a) Approval of trainers and training practices in the Oxford Region: Criteria, *British Medical Journal*, 277, pp.538-540.

SCHOFIELD, T.P.C. & HASLER, J.C. (1984b) Approval of trainers and training practices in the Oxford Region: Assessment, *British Medical Journal*, 277, pp.612-614.

SCHOFIELD, T.P.C. & HASLER, J.C. (1984c) Approval of trainers and training practices in the Oxford Region: Evaluation, *British Medical Journal*, 277, pp.688-689.

SCHON, D.A. (1988) *Educating the reflective practitioner*, (San Francisco, Jossey-Bass).

SCOTTISH OFFICE. (1995) A Strategic Framework for Clinical Audit in Scotland. MEL.52, (Edinburgh, The Scottish Office).

SHAW, C.D. (1980) Aspects of audit: 1. The background, *British Medical Journal*, ii, pp.1256-1258.

SHAW, C.D. (1990) Criterion based audit, *British Medical Journal*, 300, pp.649-651.

SHAW, C.D. & COSTAIN, D. (1989) Guidelines for medical audit: seven principles, *British Medical Journal*, 299, pp.498-499.

SHELDON, M.G. (1982) Medical Audit in General Practice. *Occasional Paper 20*, (London, RCGP).

SHEPHERD, D.J. (2000) Supporting practice-based audit, Letter, *British Journal of General Practice*, 50, p61.

SHERWOOD, T. (1992) Exitus auditus - no fun, *The Lancet*, 340, pp.37-38.

SHORTER OXFORD ENGLISH DICTIONARY ON HISTORICAL PRINCIPLES. (1985) 3rd Edition, (Oxford, University Press).

SMITH, R. (1992a) Audit in action, (London, BMJ Publications).

SMITH, R. (1992b) Audit and Research, *British Medical Journal*, 305, pp.905-906.

SPENCER, J.A. (1992) Audit and academic departments of general practice: a survey in the United Kingdom and Eire, *British Journal of General Practice*, 42, pp.333-335.

SPENCER, J.A. & BARTON, A.G. (1994) The need for teaching in medical audit: a survey in one medical school, *Medical Education*, 28, pp.323-330.

SPENCER, J., SCHOFIELD, P. & GREGORY, D. (1989) Audit projects for medical students, Letter, *Journal of the Royal College of General Practitioners*, 39, p520.

STANDING COMMITTEE ON POSTGRADUATE MEDICAL EDUCATION. (1989) Response to the White Paper, *British Medical Journal*, 299, p509.

STANDING MEDICAL ADVISORY COMMITTEE FOR THE SECRETARIES OF STATE FOR HEALTH AND WALES. (1990) *The Quality of Medical Care*, (London, HMSO).

STEVENS, J.L. (1977) Quality of care in general practice. Can it be assessed? *Journal of the Royal College of General Practitioners*, 27, pp.455-466.

STEVENS, R. (1990) Humanistic Psychology, in: **I.ROTH (Ed)**, *Introduction to Psychology*, (Hove, Lawrence Erlbaum Associates, with the Open University).

SWANSON, D.B., NORMAN, G.R. & LINN, R.L. (1995) Performance-based assessment: lessons learned from the Health Professions, *Educational Research*, 24, pp.5-11.

TODD REPORT. (1965-8) Royal Commission on Medical Education, Cmnd 3569, (London, HMSO).

UNITED KINGDOM CONFERENCE OF REGIONAL ADVISERS IN GENERAL PRACTICE. (1995) *Summative Assessment for Registrars completing vocational training after 1st September 1996*, (Southampton, UKCRA).

VAN DER VLEUTEN, C.P.M. (1996) The assessment of professional competence: developments, research and practical implications, *Advances in Health Sciences Education*, 1, pp.41-67.

VAN DER VLEUTEN, C.P.M., VAN LUYK, S.J., VAN VALLEGOOIJEN, A.M, & SWANSON, D.B. (1989) Training and experience of examiners, *Medical Education*, 23, pp.290-296.

VIMPANY, M. & NIXON, K. (1995) Reflections on education and audit, *Audit Trends*, 3, pp.137-139.

- WAGNER, E.H., AUSTIN, B.T. & VON KORFF, M.** (1996) Organising care for patients with chronic illness, *The Milbank Quarterly*, 74 (4), pp.511-544.
- WALSHE, K.** (1995) Opportunities for improving the practice of clinical audit, *Quality in Health Care*, 4, pp.231-232.
- WALSHE, K. & TOMALIN, D.** (1993) Raincheck: Selecting, planning, implementing and evaluating audit projects, *Health Services Journal*, 103, pp.28-29.
- WARNOCK, M.** (1973) in: **R S PETERS (Ed)**, *Towards a definition of quality in education*, (Oxford, Oxford University Press).
- WATERS, W.H.R., KELLY, J. & LUNN, J.E.** (1983) Attitudes to audit, *Journal of the Royal College of General Practitioners*, 33, pp.263-266.
- WATKINS, C.J.** (1981) The measurement of the quality of general practitioner care, *Occasional Paper 15*, (London, RCGP).
- WEBB, S.J., DOWELL, A.C. & HAYWOOD, P.** (1991) Survey of general practice audit in Leeds, *British Medical Journal*, 302, pp.390-392.
- WEDDERBURN, D.** (1998) Where does clinical audit fit in? *Quality in Health Care*, 7, p2.
- WEST, M.A. & WALLACE, M.** (1991) Innovation in health care teams, *European Journal of Social Psychology*, 21, pp.303-315.
- WHICH?** (1995) *What makes a good GP?* Consumers Association, June 18th.

WILSON, D.M. (1993) Education and training of pre-registration junior house officers: the consultants' viewpoint, *British Medical Journal*, 306, pp.194-196.

WOLF, A. (1995) *Competence-based assessment*, (Buckingham, Open University Press).

WORLD HEALTH ORGANISATION. (1983) The Principles of Quality Assurance, *Report on a WHO meeting*, (Copenhagen, WHO).

